UNIVERSITÉ SORBONNE PARIS NORD

ÉCOLE DOCTORALE ERASME
Université Sorbonne Paris Nord

UNIVERSITÉ PARIS XIII – SORBONNE PARIS NORD

École doctorale Érasme

Analyse contrastive de l'industrialisation des médias d'apprentissage adaptatif en France et aux États-Unis

A contrastive analysis of the industrialization of adaptive learning media in France and the United States

---

THÈSE DE DOCTORAT
présentée par

Hubert BOËT
LabSIC – LabEx ICCA

pour l'obtention du grade de
DOCTEUR EN SCIENCES DE L'INFORMATION ET DE LA COMMUNICATION

soutenue le 10 janvier 2025 devant le jury d'examen constitué de :

BONFILS Philippe, Université de Toulon, Président du jury
SAEMMER Alexandra, Université Paris 8, Rapportrice
SEURRAT Aude, Université Paris-Est Créteil, Rapportrice
MŒGLIN Pierre, Université Sorbonne Paris Nord, Examinateur
BOUQUILLION Philippe, Université Sorbonne Nouvelle, Directeur de thèse
PETIT Laurent, Sorbonne Université, Co-directeur de thèse

Résumé

Cette recherche doctorale aborde le déploiement de médias algorithmiques visant à adapter la pédagogie aux besoins de chaque apprenant, en France et aux États-Unis. L'idéal techno-éducatif de l'adaptativité est certes ancien dans ces deux pays, qui ont connu au fil du XX^{e} siècle des innovations technologiques nombreuses dans cette direction. Cependant, les développements récents de l'intelligence artificielle et des techniques de fouille de données relancent cette ambition sur l'ensemble des marchés éducatifs, sous l'étendard de l'« *adaptive learning* ». Cette recherche interroge dans quelle mesure les médias adaptatifs contribuent au remplacement de la forme scolaire traditionnelle par un paradigme singulier, cohérent et transversal aux cultures nationales et aux secteurs éducatifs. Pour cerner les enjeux techniques, pédagogiques et socio-économiques soulevés par ces médias, la thèse se divise en trois parties. La première aborde les recherches académiques portant sur l'adaptativité, via un travail documentaire et des entretiens semi-directifs menés avec des chercheurs en *learning analytics* et en *educational data mining*. La deuxième partie traite de la socio-économie des médias adaptatifs, au moyen d'analyses socio-économiques et d'entretiens semi-directifs concernant les politiques publiques, les stratégies des firmes et les réseaux d'acteurs *EdTech*. La troisième partie porte sur les médias adaptatifs codifiant la pédagogie, à travers des analyses techno-sémiotiques et sémio-discursives réalisées sur un corpus de sept produits. Cette recherche met au jour un *modus operandi* industriel visant l'engagement de l'apprenant à des fins pédagogiques et commerciales, tant dans le développement d'algorithmes que dans la conception d'interfaces. S'il prolonge la convergence entre industries éducatives et culturelles, sa cohérence est rendue friable par des limitations techniques et des différences de positionnement selon les firmes. Ce projet en évolution soulève ainsi des questionnements politiques transversaux pour l'éducation et la formation.

Mots-clés : industrialisation de l'éducation ; adaptive learning ; apprentissage adaptatif ; adaptativité ; intelligence artificielle ; systèmes de tutorat intelligent ; environnements informatiques pour l'apprentissage humain ; educational data mining ; learning analytics ; forme scolaire ; algorithmes ; EdTech ; engagement de l'apprenant

Abstract

This doctoral research examines the deployment of algorithmic media designed to adapt the pedagogy to individual needs in France and the United States. The techno-educational ideal of adaptivity is not new in these two countries, which have seen numerous technological innovations in this direction throughout the 20th century. However, recent developments in artificial intelligence and data mining have reignited this ambition across educational markets under the banner of "adaptive learning". This research examines the extent to which adaptive media contribute to replacing the traditional school model with a singular, coherent paradigm that transcends national cultures and educational sectors. To identify the technical, pedagogical, and socio-economic issues raised by these media, this thesis is divided into three parts. The first part addresses academic research on adaptivity through document analysis and semi-structured interviews with researchers in *learning analytics* and *educational data mining*. The second part examines the socio-economics of adaptive media through socio-economic analysis and semi-structured interviews conducted at the levels of public policies, corporate strategies, and EdTech networks. The third part focuses on adaptive media codifying pedagogy, through techno-semiotic and semio-discursive analyses conducted on a corpus of seven products. This research reveals an industrial modus operandi focused on learner engagement for both pedagogical and commercial purposes, evident in the development of algorithms and the design of interfaces. While it extends the convergence between educational and cultural industries, its coherence is undermined by technical limitations and differentiated positioning among firms. This evolving project thus raises broad political questions for education and training.

Keywords: educational industrialization; adaptive learning; adaptivity; artificial intelligence; intelligent tutoring systems; educational data mining; learning analytics; algorithms; EdTech; student engagement

Remerciements

Je tiens à remercier mes deux directeurs de thèse, Laurent Petit et Philippe Bouquillion, pour leur soutien continu et bienveillant durant ces quatre années de thèse. Ils ont su composer un tandem complémentaire et inspirant pour moi.

Je tiens à remercier les membres du jury Philippe Bonfils, Alexandra Saemmer, Aude Seurrat et Pierre Mœglin qui ont accepté de lire et juger mon travail.

Je remercie également les divers professionnels en France et aux États-Unis avec qui j'ai pu échanger, et en particulier Peter Brusilovsky, Carolyn Rosé, Bruce McLaren, Laurent Jolie et Gérard Puimatto qui ont accepté de témoigner dans cette thèse.

J'ai également profité de mon insertion au sein du LabSIC et du LabEx ICCA pour mener à bien cette thèse, et je tiens à remercier à ce titre Yanita Andonova, Philippe Bouquillion, Christine Chevret-Castellani, Pascaline Adonaï et Johanna Chollet pour leur aide précieuse.

Je remercie aussi mes camarades doctorants, pour les échanges joyeux et inspirants que l'on a eus au cours de ces quatre années.

Je tiens enfin à remercier Valentine, mes parents, mon frère et mes amis pour le soutien constant qu'ils m'ont apporté.

Table des matières

INTRODUCTION. L'adaptive learning, étape ou tournant dans l'histoire de la personnalisation éducative ? 10

Questionnements techniques 11
Une histoire tumultueuse de l'adaptativité 11
Le lien avec l'intelligence artificielle 13

Questionnements pédagogiques 14
Distinguer des paradigmes 14
L'industrialisation d'un projet politique 17

Questionnements socio-économiques 18
D'industries éducatives revisitées à la constitution d'une socio-économie nouvelle 19
Des facteurs de reconfiguration 21
Une reconfiguration symptomatique de la convergence avec les industries culturelles ? 22

Problématique et hypothèses 23

La nécessité d'un éclectisme méthodologique 23

PREMIÈRE PARTIE. La constitution d'un projet scientifique 27

Chapitre I. L'émergence tumultueuse de l'adaptativité 29

1. L'enseignement programmé : l'individualisation à l'heure du behaviorisme 30
Le sillage de la cybernétique et du behaviorisme 30
Considérations paradigmatiques 32
Le temps des critiques 34

2. Au-delà du *drill and practice,* les « tuteurs intelligents » 37
L'avènement des sciences cognitives et la survivance du behaviorisme 37
Le modèle pédagogique 39
Le modèle du domaine 40
Le modèle de l'élève 41
Des obstacles poussant à une convergence avec les micromondes 42

3. Les « environnements informatiques pour l'apprentissage humain » (EIAH), des perspectives nouvelles d'adaptativité malgré des contraintes substantielles 43
Une opposition intestine entre anthropocentrisme et technocentrisme 44
De nouvelles avancées en matière d'adaptativité 47
L'arrivée des hypermédias adaptatifs dans la formation à distance 49
L'arrivée du « web social » et les défis de l'industrialisation 50

Chapitre II. Les apports du Big Data et de l'IA connexionniste aux recherches sur l'adaptativité 54

1. Une tension épistémologique nouvelle 54
Une épistémologie de la donnée au service de l'éducation 54
Le connectivisme, résurgence du socio-constructivisme 56

2. L'*Educational Data Mining,* ou la poursuite étendue de l'adaptativité 57
Perfectionner les traitements algorithmiques pour dépasser les limites des enseignants 57
La prédiction de l'apprentissage 58

3. Les *Learning Analytics,* ou les limites de la technologisation additive 61
Perfectionner les visualisations pour repousser les limites des enseignants 61
Une place ambiguë réservée à l'enseignant 61

Chapitre III. Des limitations scientifiques en question 64

1. Le dilemme entre données massives et épaisses, ou l'illusion d'une technologisation substitutive 64
La question du volume de données 64
L'enjeu de l'épaisseur des données 65

2. La question des contenus 66

Les opportunités des grands modèles de langage 66
Des doutes et des velléités de « *fine-tuning* » 68

3. L'explicabilité des décisions algorithmiques 70
L'émergence des enjeux d'explicabilité dans l'éducation 70
Les biais des modèles de *machine learning* 71
Des méthodes de *computer science* et interdisciplinaires réinterrogeant l'adaptativité 72
Conclusion provisoire 72

DEUXIÈME PARTIE. De l'expérimentation à l'industrialisation : morphologies des industries éducatives revisitées par l'adaptive learning 74

Chapitre IV. Une socio-économie des médias éducatifs en suspens 75

1. Des questionnements *dans* et *entre* les industries éducatives 75
Cinq filières revisitées 75
Des enjeux transversaux 80

2. Des politiques du numérique éducatif questionnées par l'adaptativité 82
Des spécificités organisationnelles 82
Des velléités et des incohérences à actualiser 82

3. Les stratégies des firmes 84
Cinq modèles en question 84
La thèse de la convergence 86
Penser une filière intégrée 88

4. Des professionnalités dynamisées par l'*EdTech* 89
Observer la morphologie et les orientations des réseaux 89
Des compétences communes à questionner 90

Chapitre V. Des industries éducatives revisitées aux États-Unis 91

1. L'appel des acteurs publics au *personalized learning* 91
Les modes d'organisation de l'appareil éducatif et de formation 91
La construction d'un projet officiel 93
Les contraintes logistiques des districts et des établissements d'enseignement public 107

2. Les stratégies des firmes 111
Le développement des algorithmes d'apprentissage adaptatif 111
Sur les contenus : des fédérations entre structures de production informatique et éditeurs de contenus encore timides 117
La distribution des médias adaptatifs 121

3. Derrière l'*EdTech* états-unienne, des communautés renforçant l'influence des acteurs des technologies 134
Des réseaux diversifiés et dynamiques 135
La prescription de compétences professionnelles 137
Conclusion provisoire 140

Chapitre VI. La différenciation pédagogique française 142

1. L'idéal ministériel de différenciation pédagogique, une vocation culturelle et politique 142
Les modes d'organisation de l'appareil éducatif et de formation 142
Contours d'un projet officiel 144
Une stratégie ministérielle de recherche-action préalable aux généralisations 147
Des obstacles financiers et scientifiques 159

2. Les stratégies des firmes 163
Le développement des algorithmes d'apprentissage adaptatif 163
Sur les contenus : des fédérations timides entre structures de production informatique et éditeurs de contenus 172
La distribution des médias adaptatifs 175

3. Des professionnalités dynamisées par l'*EdTech* 184
Des réseaux diversifiés et dynamiques 184
La construction de professionnalités 187
Conclusion provisoire 189

Chapitre VII. Des permanences et des spécificités dans les implications socio-économiques de l'adaptativité ___ *191*

TROISIÈME PARTIE. Les modalités d'industrialisation de l'éducation à l'œuvre dans les médias adaptatifs : une convergence vers les industries de la communication ? ___ *197*

Chapitre VIII. L'industrialisation par l'engagement, un idéal-type pour la conception de médias hybrides ___ *203*

1. L'engagement de l'utilisateur, indicateur de l'efficience des médias adaptatifs ___ **203**
L'*engagement* de l'apprenant, indicateur de l'efficacité pédagogique ___ 204
L'engagement de l'utilisateur, indicateur de l'efficacité économique ___ 210

2. Les algorithmes d'apprentissage adaptatif comme gouvernail de la *learning experience* ___ **216**
La sélection algorithmique de contenus au service de la motivation des apprenants ___ 217
Des algorithmes de *data visualization* pour permettre aux enseignants de différencier ___ 223

3. Un agencement de dispositifs multimédias dans une perspective behavioriste renouvelée ___ **226**

Chapitre IX. Des tensions posées par le déploiement de l'industrialisation par l'engagement ___ *232*

1. Des alternatives algorithmiques à rebours de l'« épistémologie de la donnée » ___ **232**

2. Les limitations techniques du paradigme ___ **237**

3. Des positionnements contradictoires entre constructivisme et behaviorisme ___ **240**
La gamification, sujet d'ambivalences ___ 240
Les formes d'agentivité de l'enseignant face aux algorithmes ___ 242
Conclusion provisoire ___ 248

CONCLUSION. Vers une adaptativité au service de l'individuation ___ *249*

Principaux enseignements ___ **249**

La cohérence et ses hypothèques ___ **252**

Une singularité relative vis-à-vis de la forme scolaire ___ **253**

La question de la convergence avec les industries culturelles ___ **255**

L'adaptativité algorithmique, renouvellement de la méritocratie ? ___ **256**

Pour une éthique de l'engagement ___ **258**

ENTRETIENS ___ *262*

Entretien 1. Carolyn Rosé, le 28 septembre 2023 ___ **262**

Entretien 2. Peter Brusilovsky, le 27 septembre 2023 ___ **268**

Entretien 3. Bruce McLaren, le 24 octobre 2023 ___ **274**

Entretien 4. Laurent Jolie, le 7 mars 2023 ___ **281**

Entretien 5. Gérard Puimatto, le 30 mai 2024 ___ **285**

ANNEXES ___ *294*

Annexe 1. Tableau récapitulatif des *AI Institutes* liés à l'éducation ___ **294**

Annexe 2. Tableau récapitulatif des acquisitions de start-ups par des groupes de l'édition scolaire et universitaire en France et aux États-Unis ___ **295**

Annexe 3. Tableau comparatif des deux GTnum liés à l'IA dans l'éducation ___ **296**

BIBLIOGRAPHIE ___ *297*
Introduction ___ 297

Chapitre I 298
Chapitre II 302
Chapitre III 304
Chapitre IV 305
Chapitre V 308
Chapitre VI 318
Chapitre VII 326
Chapitre VIII 326
Chapitre IX 331

« Lorsqu'un tailleur fait un vêtement, il l'ajuste à la taille de son client et, si celui-ci est gros ou petit, il ne lui impose pas un costume trop étroit sous prétexte que c'est la largeur correspondant dans la règle à sa hauteur. Au contraire, l'école habille, chausse, coiffe tous les esprits de la même façon. Elle n'a que du tout-fait et ses rayons ne contiennent pas le moindre choix. Pourquoi n'a-t-on pas pour l'esprit les égards dont on entoure le corps, la tête, les pieds ? »

Claparède, É. (1920). *L'école sur mesure*, p.37

« Si, par miracle de l'ingéniosité mécanique, un livre pouvait être agencé d'une telle façon que seulement pour celui qui aurait fait ce qui était demandé à la première page, la page deux devienne visible, et ainsi de suite, beaucoup de ce qui requiert actuellement de l'instruction personnelle pourrait être assuré par le livre ».

Thorndike, E. (1912), *Education*, cité par Bruillard, E. (1997), pp. 33-34

INTRODUCTION. L'*adaptive learning*, étape ou tournant dans l'histoire de la personnalisation éducative ?

Personalized, individualized ou *differentiated learning* aux États-Unis, individualisation, personnalisation ou différenciation pédagogique en France : quelles que soient ses dénominations, l'idéal d'une « éducation sur-mesure » est ancien dans ces deux pays. Il a accompagné l'industrialisation croissante de l'éducation dès le début du XX^{e} siècle, alors que les systèmes éducatifs se sont massifiés et ont accueilli des publics croissants et sociologiquement diversifiés. Les deux citations présentées en exergue montrent comment l'aspiration émerge à cette époque chez les penseurs de la pédagogie occidentale. Se nourrissant de la critique d'un enseignement traditionnel standardisé censément inadapté aux besoins cognitifs individuels, ces auteurs témoignent cependant de spécificités culturelles selon les espaces francophone et anglophone. Derrière la focale individualiste, la citation du psychologue suisse Édouard Claparède fait référence à une forme d'artisanat, comparant l'éducation au secteur de l'habillement sur mesure. La seconde citation, d'Edward Thorndike, maître à penser du behaviorisme états-unien, place cependant cet espoir dans une voie technique et industrielle, envisageant un livre s'adaptant aux réponses de l'élève.

Les États-Unis seront tout au long du siècle le berceau d'innovations permettant des formes d'individualisation par la technologie, comme les « machines à enseigner » des années 1950 ou les programmes informatiques des années 1970. À l'inverse des pédagogues contestataires de « l'Éducation nouvelle » en France, qui envisagent l'enseignant au centre de pratiques de différenciation, les penseurs états-uniens des « technologies éducatives » remettent en question l'avenir du métier d'enseignant.

Si ces questionnements ont été rapidement tranchés devant les limites techniques et les échecs commerciaux des innovations du XX^{e} siècle, les développements récents de l'intelligence artificielle et des techniques de fouille de données dans l'éducation les relancent de façon puissante. Ces dernières années, les logiciels éducatifs et sites d'apprentissage promettant une adaptation de la prestation éducative à l'apprenant, selon les données numériques qu'il laisse en ligne, se sont multipliés sur l'ensemble des marchés éducatifs, de la maternelle à la formation professionnelle. Ce type d'outils a d'abord émergé aux États-Unis dans les années 2000 sous l'étendard de l'« *adaptive learning* », avant de germer en France et dans d'autres pays industrialisés à partir des années 2010. Les producteurs justifient leurs offres par la même critique d'un enseignement traditionnel standardisé – ou « *one size fits all* » outre-Atlantique. La pénurie de professeurs qualifiés, le passage forcé au distanciel lors de la pandémie de Covid-19 ou les résultats jugés alarmants des élèves aux enquêtes internationales sont autant de facteurs conduisant les gouvernements à s'intéresser à ces technologies.

Les médias intégrant ces algorithmes seraient-ils en passe de relancer le projet d'un « sur-mesure de masse » et ainsi de refonder l'éducation et la formation, au-delà des spécificités culturelles et organisationnelles liées aux systèmes éducatifs nationaux ? Sur cette vaste interrogation, il convient de détailler trois points de vue – technique, pédagogique et socio-économique – ayant trait aux potentialités et limites de ces algorithmes dans le domaine de l'éducation.

Questionnements techniques

L'adaptativité (*adaptivity* en anglais) est définie par l'informaticien Reinhard Oppermann comme la capacité d'un système à « modifier automatiquement ses propres caractéristiques selon les besoins de l'utilisateur »[1] (Oppermann, 1994, p.456). Par la recherche des besoins de l'apprenant à son insu, elle se distingue de l'adaptabilité, « forme d'adaptation initiée et mise en œuvre par l'utilisateur seul »[2] (ibid.). Plus de vingt ans après, dans son rapport de 2016 intitulé *Decoding adaptive*, le média Edsurge définit cet idéal comme la capacité d'une technologie à « répondre aux interactions d'un élève en temps réel en lui offrant automatiquement un soutien personnalisé »[3] (Edsurge, 2016, p.15). Se jouerait désormais, dans les environnements algorithmiques, la collecte en continu des données des apprenants, au-delà de tests ponctuels, pour adapter la pédagogie « en temps réel ».

Nous proposons pour notre part une définition de l'adaptativité reposant sur trois éléments :

- La présence d'un média, outil reproductible permettant l'accès à des contenus éducatifs, selon l'acception proposée par Pierre Mœglin (2005, p.72) ;
- La captation des besoins de l'élève en continu, autrement dit la capacité du média à enregistrer les diverses interactions de l'apprenant avec l'interface pour en déduire ses besoins pédagogiques, sans que celui-ci ne les formalise ;
- La flexibilité de la pédagogie, c'est-à-dire la capacité du média à modifier un aspect pédagogique en fonction de ces besoins individuels.

Cette grille de lecture doit être restituée dans une histoire scientifique de l'adaptativité. Il s'agit de comprendre quelles ont été les avancées techniques et les controverses scientifiques ayant émergé dans les laboratoires de recherche et les organisations éducatives, et conduit aux dispositifs actuels.

Une histoire tumultueuse de l'adaptativité

Tout d'abord, les velléités d'automatisation de l'éducation sont millénaires, ainsi que l'écrit Pierre Mœglin : « Si contestable soit-il, le préjugé qui fait de la pédagogie l'art de la répétition est à l'origine de nombre d'inventions depuis l'Antiquité mais qui ne sont pas industrielles ». Au premier critère de l'adaptativité, qui suppose la reproduction industrielle d'un média, répond le livre imprimé du XV^{e} siècle. Selon Marshall McLuhan, il s'agit « littéralement [d']une machine à enseigner là où le manuscrit n'avait été qu'un outil grossier » (ibid., p.267). Pour être adaptatif, il lui manque cependant les deux autres critères, la capacité à enregistrer en continu les caractéristiques de l'usager pour faire évoluer son contenu. Les « *teaching machines* » inventées aux États-Unis durant le XIX^{e} siècle, du *Mode of Teaching to Read* de H. Chard (1809) à l'*Apparatus for Teaching Spelling* de H. Skinner (1866), ne présentent que peu d'avancées dans ces directions. Elles mettent en œuvre une adaptation sommaire, assimilable à de l'adaptabilité – une réponse mécanique à la volonté d'avancer de l'apprenant, qui actionne des manivelles pour faire défiler des mots ou des images (Watters, 2015).

Une première avancée survient en 1926 lorsque, toujours aux États-Unis, Sidney Pressey présente une machine qui exauce le vœu de Thorndike d'individualiser l'avancement des leçons. L'élève

[1] *« A system is called adaptive if the system is able to change its own characteristics automatically according to the users' needs ».*
[2] *« The adaptation is initiated and performed by the user alone (adaptability) ».*
[3] *« [W]e define digital adaptive learning tools as education technologies that can respond to a student's interactions in real-time by automatically providing the student with individual support ».*

répond à une question affichée sur un rouleau en actionnant une touche correspondant à une réponse, et la question suivante s'affiche uniquement si la réponse est correcte. Si elle demeure pédagogiquement très basique, la *Pressey Testing Machine* s'adapte en déduisant d'une erreur un besoin de répétition. Elle ouvre la voie au courant industriel de l'« enseignement programmé ». Celui-ci culmine au tournant des années 1950 et 1960 à travers des « machines à enseigner » permettant aux élèves de s'exercer seuls dans des programmes pré-écrits.

Viennent ensuite, avec l'avènement de l'ordinateur personnel, les programmes informatiques tâchant de « modéliser » numériquement la cognition de l'apprenant. Ils fonctionnent à partir d'algorithmes, c'est-à-dire de « suites d'instructions à exécuter », ainsi que les définit l'informaticien Yann Le Cun (2021, p. 381). Les années 1970 et 1980 voient émerger dans diverses matières des « systèmes-experts ». À partir de règles écrites par des programmeurs (du type « Si… alors…»), ceux-ci adaptent leurs stratégies pédagogiques non seulement aux bonnes ou aux mauvaises réponses de l'élève, mais aussi à des aspects diversifiés de leurs interactions avec les interfaces (temps de réponse, demande d'indices…).

À partir des années 1990, avec la démocratisation du web ainsi que l'accroissement des capacités de stockage des serveurs informatiques, des algorithmes de *machine learning* sont développés. L'apprentissage-machine est, selon Le Cun, « l'ensemble de méthodes pour entraîner un système, au lieu de le programmer explicitement » (ibid.). En éducation, de tels algorithmes sont entraînés, à partir de données nombreuses et de natures diverses, à diagnostiquer et prédire les progrès d'un individu, pour lui assigner des contenus adéquats. L'avènement du *Big Data* au tournant des années 2000 et 2010 donne un coup d'accélérateur à ces recherches, jusqu'au développement de « réseaux de neurones » et l'intégration de grands modèles de langage dans l'éducation.

Les médias adaptatifs actuels reposent, pour certains, sur des techniques anciennes liées aux systèmes-experts, ou se fondent sur des modèles « apprenants ». Collectant quoiqu'il en soit des données « en continu », ils opèrent des adaptations sur trois aspects selon Edsurge (2016) :

- L'« *adaptive content* » concerne les aménagements à l'intérieur de l'exercice, à la suite d'une erreur commise par l'élève, comme le découpage de l'exercice en micro-actions, ou l'affichage d'une vidéo ou d'une précision grammaticale.
- Toujours à l'échelle de l'exercice, des algorithmes peuvent également adapter la teneur et la difficulté d'une question en fonction du résultat obtenu à la précédente (« *adaptive assessment* »).
- Le troisième mode d'intervention, l'« *adaptive sequencing* », se situe au niveau plus vaste du programme : des algorithmes ré-agencent le parcours de l'apprenant en modifiant la chronologie des notions et formats, pour s'adapter à son niveau – là où les deux modes précédents conservent un programme unique pour tous.

Dans tous ces cas, les matériaux pédagogiques sont supposés s'adapter de façon optimale aux performances des apprenants. Ainsi, un système d'enseignement de la lecture peut ralentir le rythme d'affichage d'une phrase pour un enfant ayant commis une erreur. Un environnement web d'enseignement des mathématiques au collège peut adapter la difficulté des questions et la quantité de choix proposés dans des questionnaires. Un logiciel de formation aux techniques de vente peut proposer des parcours adaptés aux connaissances de chaque commercial, évalué en continu. L'adaptation peut aussi varier en termes de modalités communicationnelles, d'un contenu imposé

de manière indolore pour l'apprenant à un autre seulement recommandé, que l'usager consultera s'il le souhaite.

De cette brève histoire de l'adaptativité ressortent des lignes de crête techniques qui mériteront d'être approfondies au cours de notre thèse. Les médias adaptatifs actuels doivent premièrement être mis en perspective avec ce qu'ont été les limites des dispositifs précédents de différenciation. Dans quelle mesure les environnements algorithmiques dépassent-ils la rigidité pédagogique des machines et des premiers programmes informatiques qui n'enseignaient que par répétition ?

Il s'agit également de distinguer, derrière l'étiquette de l'*adaptive learning* en vogue dans l'offre médiatique actuelle, les dispositifs proprement adaptatifs de ceux opérant une collecte et un traitement de données seulement de façon ponctuelle (lors de tests de positionnement par exemple). L'*adaptive learning* s'apparente en effet à ce que Gaston Bachelard nommait une « expression obstacle » (Bachelard, 1938) : sa popularité absorbe le récepteur dans un imaginaire empêchant de discerner les méthodes algorithmiques qui le définissent. Nous mettons ainsi à distance cette expression en désignant plutôt par *médias adaptatifs* et *algorithmes d'apprentissage adaptatif* les technologies qui répondent à la définition en trois temps que nous avons apportée plus haut.

Parmi celles-ci, les modèles de *machine learning* en plein essor depuis l'avènement du *Big Data* doivent en outre être distingués des systèmes-experts traditionnels, pour saisir la rupture scientifique qu'ils représentent, via notamment leurs derniers avatars (les « réseaux de neurone » et les grands modèles de langage). Dans quelle mesure les méthodes inductives centrées sur les données éclipsent-elles, dans les protocoles de recherche, les enquêtes hypothético-déductives basées sur des théories pédagogiques ? Quelles en sont les implications en termes de fonctionnalités techniques ?

Enfin, bon nombre de ressources commerciales et académiques situent les médias adaptatifs dans le domaine de l'intelligence artificielle. Quelle est la part de réalité scientifique et d'opportunité communicationnelle dans cette association, à l'heure où l'IA s'apparente elle-même à une « expression obstacle », suscitant des discours promotionnels et des financements très importants ?

Le lien avec l'intelligence artificielle

Ce questionnement nécessite de partir d'une définition scientifique de l'IA. La tâche est malaisée, étant donnée l'opposition traditionnelle entre deux approches dans le courant. D'un côté, les tenants d'une IA dite « symbolique » considèrent l'intelligence comme un phénomène logique et intelligible, qu'il s'agirait de reproduire dans des machines. De l'autre, les partisans du courant « connexionniste » situent l'intelligence dans la liaison d'unités élémentaires semblables aux neurones, indépendamment du sens des opérations qu'elles effectuent. Cette opposition philosophique intestine est susceptible d'éclairer la distinction entre les systèmes-experts et les modèles de *machine learning* qui segmentent actuellement les offres adaptatives.

Cependant, pour dépasser cet antagonisme et partir d'une définition consensuelle, nous reprenons la proposition de Jean-Gabriel Ganascia, spécialiste des humanités numériques, qui définit l'IA comme une discipline scientifique visant à « décomposer l'intelligence en fonctions élémentaires » et « à construire des machines pour les simuler une à une » (Ganascia, 2018). Il liste ces fonctions comme suit :

> « 1. Perception des signaux envoyés par nos organes des sens, c'est-à-dire construction d'une représentation qui agrège les informations qu'ils nous transmettent ; 2. mémoire, à savoir représentation – au sens étymologique de re-présentation, c'est-à-dire de restitution de la présence d'une chose en son absence –, et exploitation de ces informations avec l'apprentissage ; 3. pensée, c'est-à-dire calcul sur les représentations ; 4. fonction communicative, à savoir capacité à échanger entre machines ainsi qu'entre hommes et machines ; 5. enfin, fonctions exécutives pour qu'une machine prenne des décisions et les mette en œuvre » (ibid.).

On retrouve des simulations de ces cinq activités dans les médias adaptatifs, aussi divers soient-ils. La perception est assumée par des dispositifs métriques (des *cookies* aux capteurs oculométriques) visant à traiter des apprentissages sous formes de données. La mémoire est simulée par les capacités de stockage des serveurs informatiques. La pensée est répercutée par des traitements algorithmiques, qui varient entre les systèmes-experts et les modèles de *machine learning*. Les fonctions communicatives sont assurées par la complémentarité entre l'adaptativité et l'interaction humain-machine (IHM) – témoin les recherches en IA générative et en traitement du langage naturel pour instrumenter des agents conversationnels. Enfin, les fonctions exécutives sont variées comme on l'a vu dans la trilogie d'Edsurge, de l'adaptation de feedbacks à la recommandation de parcours d'apprentissage.

Toutefois les systèmes adaptatifs mobilisent également des techniques qui ne sont guère assimilées à de l'intelligence artificielle, comme des traitements statistiques à des fins descriptives – notamment pour actualiser des indicateurs dans des tableaux de bord à destination des apprenants et des enseignants. Il importe donc de jauger la parenté entre les recherches sur l'adaptativité et l'IA pour dissiper les imaginaires technologiques qui entourent l'*adaptive learning*.

Il importe aussi de replacer ces divers dispositifs techniques dans une histoire des projets pédagogiques poursuivis par les chercheurs et les industriels.

Questionnements pédagogiques

En tant que *dispositifs* – dont « les normes et les formats qu'ils comportent influencent les formes de la communication » (Jeanneret, 2005) –, les algorithmes projettent certaines manières d'apprendre et d'enseigner. Notre hypothèse exposée plus haut d'une refondation des systèmes éducatifs et de formation suppose que ces dispositifs rompent avec les approches pédagogiques traditionnelles, notamment celles perçues comme standardisées et collectives, pour instituer une pédagogie nouvelle et cohérente.

Distinguer des paradigmes

Pour tenter de cerner des propositions antinomiques, nous reprenons une approche dite « paradigmatique ». Celle-ci permet d'expliquer le devenir éducatif d'innovations techniques par des tendances sociétales, auxquelles serait soumise l'éducation comme tout autre secteur social. Cette méthode se fonde sur le concept englobant de *paradigme*. Le philosophe Thomas Kuhn l'a proposé pour désigner un « problème admissible », une « solution légitime » à un stade donné du développement scientifique. Dans cette logique, les évolutions des connaissances et des représentations dans la société feraient progressivement advenir des critiques envers le paradigme, jusqu'à une révolution scientifique marquant son renversement et l'installation d'un nouveau paradigme (Kuhn, 2008).

Nous reprenons cette approche, qui va nous permettre de penser l'opposition d'un modèle éducatif traditionnel à un nouveau auquel se rattacheraient les médias adaptatifs. Comment définir le premier, au-delà des niveaux d'enseignement et des cultures éducatives nationales ? La sociologie française a proposé le concept de *forme scolaire*, à partir des travaux de Guy Vincent ou de Bernard Lahire. Il est difficile d'en donner une définition unique car ses caractéristiques ne font pas consensus, étant donné son statut d'*idéal-type*, c'est-à-dire d'« abstraction construite à partir d'une grande diversité de pratiques et d'organisations concrètes » comme le remarquent Olivier Maulini et Philippe Perrenoud (2005). Les deux chercheurs proposent tout de même huit traits distinctifs : un contrat didactique entre un formateur et un apprenant, une organisation centrée sur les apprentissages, une pratique sociale distincte et séparée du reste de la société, une planification de l'apprentissage en curriculum, une transposition didactique des savoirs enseignés, un temps didactique, une discipline intellectuelle et corporelle, et des normes d'excellence (ibid.). Ces caractéristiques ont l'avantage de pouvoir s'appliquer à des institutions très diverses et autres que les écoles *stricto sensu*, comme les universités ou les formations professionnelles.

Claire Peltier, Daniel Peraya, Philippe Bonfils et Laurent Heiser proposent quant à eux de désigner ce qu'ils nomment une *forme éducative* à partir de quatre dimensions interdépendantes, où s'enchevêtrent les caractéristiques de Maulini et Perrenoud : un espace, une temporalité, des règles et rapports sociaux, et un rapport au savoir (Peltier et al., 2021). Nous reprenons pour notre part le concept de *forme scolaire* au sens large, en tant que modèle transversal qui « tolère des "institutionnalisations" différenciées » selon la formule de Laurent Petit (2018), comme l'université ou l'école prise au sens strict. Nous alternerons entre une approche par les quatre dimensions de Peltier et al., et une autre se référant aux huit caractéristiques de Maulini et Perrenoud.

Pour mesurer l'avènement d'un nouveau paradigme, il s'agit de jauger trois éléments :

- la singularité d'un projet pédagogique « adaptatif » par rapport à la forme scolaire ;
- la cohérence des caractéristiques de ce projet, leur permettant de prendre ensemble « forme » ;
- la transversalité de ce projet, qui lui permettrait d'être valable au-delà des démarcations entre secteurs éducatifs et des particularismes nationaux.

Sur la singularité, il s'agit d'abord de resituer le phénomène marketing de l'*adaptive learning* dans un contexte de crise pour la forme scolaire, et d'identifier les parentés du projet adaptatif avec des propositions paradigmatiques antérieures. La forme scolaire est en effet remise en cause depuis plusieurs décennies, en France comme aux États-Unis, par des appels à l'individualisation. Devant l'afflux de publics sociologiquement diversifiés depuis le début du XX^{e} siècle, cette pression prend des formes multiples comme on l'a vu, de la réorganisation des cours permettant une différenciation par l'enseignant à l'intégration de technologies personnelles.

Pour de nombreux chercheurs, cette tendance s'est cependant accélérée avec l'avènement des technologies numériques, à rebours du manuel standardisé et représentatif de la forme scolaire. Les pratiques personnelles d'autoproduction et de partage de ressources se diffusent depuis de nombreuses années dans les institutions éducatives, notamment avec la démocratisation des terminaux personnels (ordinateurs, smartphones et tablettes), remettant en cause la configuration collective et présentielle.

Alain Chaptal a notamment souligné trois implications paradigmatiques des technologies numériques. Premièrement « les outils de la pédagogie sont ceux grâce auxquels l'ensemble de la société crée désormais l'essentiel de sa richesse » (Chaptal, 2003, p.120). Deuxièmement, ces technologies sont à la fois présentes au domicile de l'enseignant et en classe. Troisièmement « ces technologies sont fondamentalement des outils de maîtrise personnelle de l'information » (ibid., p.130). Elles s'adapteraient ainsi à tous les styles sans obliger les enseignants innovateurs à changer leurs méthodes. Dans cette lignée, Pierre Mœglin a proposé une liste non-exhaustive de tendances opposables entre d'hypothétiques « nouveau » et « ancien » paradigmes éducatifs : « apprentissage contre enseignement, appropriation contre transmission, sur-mesure contre standard, autonomie contre hétéronomie, individualisation contre massification, pédagogie constructiviste contre pédagogie behavioriste, validation ou brevet d'aptitude contre diplôme, etc » (Mœglin, 2005, p.188). Laurent Petit ajoute d'autres tendances réformatrices comme la pression nouvelle des parents « qui réclament une adaptation de la prestation éducative » ou la formation tout au long de la vie, « injonction d'organisations internationales multiples (UNESCO, OCDE, Commission européenne, etc.) qui responsabilise l'individu dans son rapport à la formation » (Petit, 2018).

À l'avènement du numérique et du web dans les années 1990 s'ajoute un autre événement sociotechnique qui a accompagné l'émergence de l'*adaptive learning* dans l'espace public, à savoir le déploiement du *Big Data* et de techniques afférentes de fouille de données. Certains, comme Bruno Bachimont, le tiennent pour une « révolution scientifique », par sa manière d'encoder le social via la donnée numérique (Bachimont, 2015). On peut alors se demander dans quelle mesure les médias adaptatifs qui en découlent prolongent-ils les dimensions paradigmatiques du numérique exposées plus haut, ou au contraire s'en détournent-ils.

Concernant la cohérence d'un « projet adaptatif », les algorithmes d'apprentissage adaptatif soulèvent de potentiels bouleversements pour la forme scolaire, sur chacune des caractéristiques données par Maulini et Perrenoud. Les médias adaptatifs mettent d'abord au défi la clôture de l'institution éducative par rapport au reste de la société. Ils s'inscrivent en cela dans le sillage des environnements d'apprentissage en ligne, permettant déjà des usages au croisement de la vie professionnelle et privée, dans la lignée de la tendance ATAWAD (« *Any Time, Any Where, Any Device* »/« N'importe quand, n'importe où, sur n'importe terminal »). Désormais les médias adaptatifs, reposant *a fortiori* sur le *machine learning*, ne supposeraient-ils pas une porosité encore accrue entre sphères éducatives et privées, si, pour améliorer leurs adaptations, ils doivent collecter des données sur les facteurs psychologiques et contextuels de l'apprentissage ?

D'autre part, le contrat didactique entre un formateur et un apprenant est susceptible d'être revisité en profondeur. Quelle est la répartition des tâches pédagogiques prévue entre l'enseignant et la machine ? Celle-ci joue-t-elle le rôle d'un tuteur voire d'un précepteur, à même de se substituer au formateur humain pour des usages à distance ? Si, dans une vision plus complémentaire, l'enseignant conserve son pouvoir de décision et jouit d'une assistance de la machine, verrait-il ses capacités de différenciation augmentées, ou devrait-il s'en remettre aux recommandations inexplicables de « boîtes noires » ?

Ces considérations revisitent l'opposition traditionnelle entre behaviorisme et constructivisme qu'opérait Pierre Mœglin plus haut, chacune plaçant dans l'individualisation des enjeux différents. D'un côté, le behaviorisme, une théorie psychologique née au début du XX^{e} siècle sous l'influence

du susnommé Edward Thorndike, vise l'acquisition de comportements chez un individu, en agissant sur les conditions d'apprentissage en entrée, sans s'intéresser à ses états mentaux intermédiaires. L'individualisation y est promue en tant que moyen pour observer et conditionner les comportements.

À l'inverse, le constructivisme hérité de Jean Piaget ainsi que ses ramifications (cognitivisme, socio-constructivisme) postulent que l'individu construit ses connaissances en observant et en agissant sur le monde qui l'entoure. Appliquée aux environnements numériques, cela donne lieu à une focalisation sur l'agentivité de l'apprenant au sein de l'outil et d'une communauté, et l'individualisation est ignorée sinon visée à travers l'autonomie de l'élève.

Du côté de l'apprenant, y aurait-il dorénavant une focalisation accrue sur ses besoins – ce que promet la dénomination d'*adaptive learning* – ou, au contraire, une simplification de l'apprentissage selon les traitements algorithmiques ? La question se pose alors que la focalisation sur les comportements d'entrée et ceux désirés en sortie – principe du behaviorisme – est communément imputée aux techniques de fouilles de données massives (Rouvroy, 2013).

D'autres questionnements pèsent sur le principe de « discipline » dans la forme scolaire, supposant « que l'apprentissage soit d'une certaine façon « laborieux », qu'il ne se fasse pas spontanément mais au prix d'une volonté et d'efforts » (Maulini et Perrenoud, 2005, p.152). Dans quelle mesure l'aménagement automatisé du contenu selon les besoins – notamment en vue de maintenir la motivation de l'apprenant – peut-il remettre en cause la primeur de l'effort ?

Enfin la référence à des normes d'excellence et des critères d'évaluation permettant de définir et de mesurer une progression des apprentissages est également en question. L'évaluation des compétences des apprenants, fondamentale dans l'adaptation des contenus pédagogiques, obéit nécessairement à des normes collectives. Elle poursuit ainsi l'idéal de mesure des progressions individuelles, conformément à la *quantophrénie*, tendance moderne des sociétés occidentales à mettre en chiffres le social et l'individu (Supiot, 2014). Pourtant l'évolution de chaque élève dans des parcours plus ou moins cloisonnés, invisibilisant le collectif, ne risque-t-elle pas de porter atteinte à l'idée normative et méritocratique d'excellence ?

L'industrialisation d'un projet politique

Quant à la transversalité du « projet adaptatif », celui-ci soulève des enjeux politiques et potentiellement un projet de société concurrent qu'il s'agit de déterminer. Ces enjeux sont liés à la massification et à la standardisation de l'éducation, que Pierre Mœglin et les chercheurs du SIF dépeignent comme un processus d'*industrialisation* de l'éducation, défini par trois marqueurs (Mœglin, 1998, 2016) :

- La *technologisation*, indiquant « la présence et l'utilisation de dispositifs techniques », qui « ajoutent la dimension des usages qu'ils prescrivent et des pratiques qu'ils modèlent à la dimension matérielle des outils et médias qui leur est intrinsèque » (Mœglin, 2016, p.41).
- La *rationalisation*, signalant la diffusion d'une « rationalité en finalité » (Weber, 1963) entre les parties prenantes du produit, c'est-à-dire d'une conception instrumentale et organisée de la pédagogie.
- L'*idéologisation*, ou « le processus par lequel […] des acteurs rationalisent la rationalité de leurs stratégies » (Mœglin, 2016, p.45), indiquant ainsi la diffusion de récits mobilisateurs visant à convaincre des publics différents de coopérer dans l'utilisation de l'outil.

La forme scolaire, par son espace-temps cloisonné et sa conception verticale de l'enseignement se prête à des formes d'industrialisation de l'éducation organisées par une autorité politique. Témoin son institutionnalisation étatique en France au XIX^{e} siècle, qui reste globalement en vigueur. Celle-ci était traversée par une forte rationalisation à travers l'enseignement collectif et simultané, l'enseignant dispensant un cours devant une vingtaine d'élèves. Elle était travaillée par une technologisation poussée à travers le recours au manuel scolaire, organisant le temps de cours et le travail à la maison des élèves et des enseignants (Mœglin, 2005). Cette institutionnalisation devait ainsi servir, sur le plan de l'idéologisation, des idéaux méritocratiques et une uniformisation linguistique et culturelle du pays.

De nombreux chercheurs ont mis au jour la contribution des dispositifs numériques personnels à un modèle industriel renouvelé. À rebours de la forme scolaire, celui-ci mettrait les usagers en position non plus passive mais de producteurs autonomes du service (*self-service*), dans une logique parfois qualifiée de « postfordiste » (Bates, 2000). Celle-ci profiterait notamment à des acteurs des technologies, comme les plateformes de *Massive Open Online Courses* (Bullich, 2018).

Qu'en est-il des médias adaptatifs ? Supposant une mise en activité de l'usager en vue de collecter des données, ces médias sont susceptibles d'étayer ce mode industriel. En même temps, par le cloisonnement de l'apprenant dans des parcours pédagogiques « sur-mesure », ils sont susceptibles de favoriser une industrialisation de l'éducation plus coercitive. Ce faisant, ils pourraient faire advenir un projet industriel différent, recélant des rapports de pouvoir nouveaux.

Les médias adaptatifs semblent en outre porter à première vue les germes d'une annihilation de la méritocratie républicaine, idéal de la forme scolaire devant permettre une ascension de chaque individu selon ses mérites. Dans ce nouveau modèle, le potentiel échec scolaire pèserait non plus sur l'élève mais sur la machine qui aurait pêché dans l'adaptation des contenus d'apprentissage et de remédiation. En outre, la logique de compétition entre les individus serait abolie par des parcours d'apprentissage sur-mesure. Ces médias pourraient ainsi apporter une réponse à la critique adressée par le sociologue britannique Michael Young en 1958 qui inventa le terme pour dénoncer l'*hubris* nouvelle des sociétés occidentales. La discorde sociale en serait le résultat, par orgueil des vainqueurs, qui croiraient ne devoir leurs privilèges qu'à eux-mêmes, et l'humiliation des perdants, qui se sentiraient responsables de leur position inférieure dans l'échelle sociale (Young, 1958). Le projet politique des médias adaptatifs serait-il plus inclusif et soutenable pour les individus ?

Questionnements socio-économiques

La capacité d'un modèle industriel à s'établir dans une société dépend également de sa viabilité sur le plan économique. Or les difficultés des technologies numériques à supplanter la forme scolaire lors des dernières décennies ont procédé en partie sur ce niveau : insuffisance des moyens financiers engagés dans les réformes des organisations éducatives, difficultés de production de ressources numériques en nombre suffisant, et faiblesse des structures de production. La contribution du courant industriel de l'*adaptive learning* à l'établissement d'un paradigme nouveau tient dès lors à la nouveauté des conditions économiques du déploiement des médias adaptatifs.

Notre hypothèse est que les algorithmes revisitent les rapports de force *dans* et *entre* les industries éducatives existantes. Pierre Mœglin en a dénombré cinq (2010) : l'édition de manuels, l'informatique pédagogique, la formation à distance, les produits pour l'éducation non-formelle et les systèmes numériques de gestion et d'information. Ces dernières se définissent par la mise à disposition de médias spécifiques, projetant des situations pédagogiques propres, où interviennent des consommateurs, acheteurs, prescripteurs et offrants divers.

D'industries éducatives revisitées à la constitution d'une socio-économie nouvelle

Pour appréhender ces spécificités, chaque industrie peut être définie comme une « filière ». Dans la théorie des industries culturelles, ce concept désigne un « ensemble structuré, dynamique et finalisé des relations entre les acteurs d'une branche ou d'un secteur […] organis[ant] donc la complémentarité des contributions de ces acteurs, en vue de la production et de la valorisation d'un produit » (Mœglin, 2005, p.212). Gaëtan Tremblay en fait le lieu de la compétition et de la coopération, de la concurrence et du partenariat où se fixent les statuts des acteurs, la position dominante revenant à celui qui parvient à imposer son modèle de synergie (Tremblay, 2003). Les filières éducatives sont en outre soumises à des interventions différentes des pouvoirs publics, qu'ils soient des États ou des collectivités territoriales, pouvant prendre la forme de financements de l'offre et/ou de la demande, de régulations ou de simples prescriptions. Ces interventions varient en outre selon les pays et les échelons éducatifs.

La plus ancienne filière est l'édition de manuels, entretenant un lien ténu avec les systèmes éducatifs nationaux, pour lesquels elle fournit ce média représentatif de la forme scolaire aux échelons primaires et secondaires. Cette industrie concentre des oligopoles puissants des deux côtés de l'Atlantique, et la place des acteurs étatiques et des collectivités territoriales y est forte, selon des modalités toutefois bien différentes : alors que le système scolaire français demeure centralisé autour du Ministère de l'Éducation nationale pour la prescription, partageant l'achat avec les collectivités, le fonctionnement est décentralisé aux États-Unis, les États fédérés opérant des prescriptions et la multitude de districts passant des achats pour leurs établissements.

Une deuxième industrie dont les produits tendent à intégrer directement des algorithmes d'apprentissage adaptatif est l'informatique pédagogique. Elle se distingue par l'ambition de fournir un environnement informatique d'apprentissage à un apprenant, pour une utilisation personnelle lors d'un cours ou d'une formation en présentiel. Elle s'organise également en oligopoles comprenant sur leurs franges une myriade de petites structures fragiles. La place des acteurs publics y est également forte aux échelons primaire et secondaire, suivant des modalités de prescription et d'achat semblables à celles des manuels.

De nombreux médias adaptatifs sont issus d'une autre industrie, la formation à distance, du fait de leurs usages hors des classes et des amphithéâtres. Ces outils ciblent des publics plus âgés et autonomes, notamment dans l'enseignement supérieur et la formation professionnelle. Des formes d'oligopoles se sont constitués entre des entreprises aux bases d'utilisateurs mondialisées (à l'image des plateformes de MOOC) et une multitude de petites structures fragiles. La place des acteurs publics est moins forte dans cette industrie que dans l'informatique pédagogique mais reste notable, *a fortiori* en France, du fait du poids des régulations et des aides nombreuses.

Des environnements utilisables à la fois à distance et en présentiel intègrent des algorithmes d'apprentissage adaptatif. Ils forment une sous-industrie au croisement de l'informatique pédagogique et de la formation à distance : le *e-learning*. Vendus le plus souvent aux institutions éducatives du primaire et du secondaire, ils obéissent à la même chaîne de commande que l'informatique pédagogique, sous l'influence là encore des pouvoirs publics.

Les algorithmes d'apprentissage adaptatif revisitent une filière très différente de celles vues jusqu'ici, dans la mesure où elle passe outre les institutions éducatives pour une distribution directement auprès du consommateur, via les magasins d'applications mobiles : celle des « produits pour l'éducation non-formelle ». La frontière y est mince entre divertissement et éducation chez ces produits qui, en s'industrialisant via le cinéma, la radio, la télévision, ou le jeu vidéo, capitalisent sur la double référence aux industries éducatives et aux industries culturelles. Les interventions publiques se limitent à des réglementations pesant sur les éditeurs d'applications en termes de gestion des données personnelles.

Les dispositifs algorithmiques touchent enfin la filière récente des systèmes numériques de gestion et d'information. Celle-ci recouvre des environnements numériques de travail (ENT) pour l'administration de l'enseignement, des outils-auteurs permettant à des enseignants d'éditer des cours, et des plateformes d'intermédiation donnant aux élèves et enseignants un accès à un catalogue de ressources numériques. Le paysage est composé d'acteurs aux agendas variés, entre des plateformes issues d'initiatives gouvernementales, de gestionnaires privés, ou de multinationales des technologies.

Ainsi spécifiées, les filières sont néanmoins réunies dans une « socio-économie » des médias éducatifs. Ce concept également tiré de la théorie des industries culturelles désigne une configuration donnant à voir, de façon transversale aux filières, des complémentarités et des clivages structurants, de même que des stratégies économiques codifiées et, derrière celles-ci, des projets de société concurrents. Les cinq industries éducatives forment une socio-économie animée par des logiques contradictoires, les poussant tantôt à se spécifier les unes des autres, et tantôt à s'allier. À la spécification fait écho la prétention de chaque industrie à l'hégémonie et à l'affaiblissement de ses concurrentes. Du côté des alliances, on peut recenser, par exemple, les collaborations entre éditeurs de manuels et structures de production informatique pour concevoir des environnements d'apprentissage.

Notre hypothèse est que le déploiement accru d'algorithmes d'apprentissage adaptatif refaçonnerait les positions des acteurs dans l'ensemble des industries éducatives existantes. Il contribuerait ainsi à leur fédération au sein d'une socio-économie rénovée, qui ferait advenir certaines stratégies économiques ainsi que des complémentarités et clivages nouveaux *dans* et *entre* les filières.

Pour observer les stratégies des firmes éditrices et distributrices de médias adaptatifs ainsi que les projets de société qu'elles recouvrent, la théorie des industries culturelles nous offre un autre outil conceptuel pertinent : le « modèle socio-économique ». Celui-ci dépeint des modes d'organisation idéal-typiques formalisés *a priori*, régissant les conditions de production des produits, la structuration de leurs processus de valorisation ainsi que les modes de rémunération des diverses parties prenantes. Du point de vue des filières, les modèles socio-économiques sont des stratégies collectives visant à atténuer l'incertitude sur la rentabilité des produits – risque consubstantiel aux industries où la valeur symbolique excède la valeur d'usage. Le consensus autour de ces modèles tient à l'efficacité

des représentations sociales sous-jacentes, susceptibles d'amener des acteurs différents à sceller des collaborations, pour non seulement rechercher la rentabilité mais aussi accomplir des projets de société nouveaux. Traitant d'idéologies, les modèles socio-économiques sont en cela des outils conceptuels inscrits dans une « économie politique de la communication », au sens de Gaëtan Tremblay, associant théorie de la production des biens matériels et théorie des relations de pouvoir (Tremblay, 2007).

L'intégration croissante de l'adaptativité dans les filières éducatives et la centralité stratégique désormais accordée à la collecte de données sont susceptibles de faire évoluer les rapports de force et de favoriser des *leaderships* dans la définition des médias à destination des publics. Les stratégies d'acteurs dominants, incluant les projets de société qu'ils parviendraient à imposer, leur permettraient d'influer, dans l'offre médiatique, sur le déploiement de certaines formes d'adaptivité et *in fine* la définition d'un paradigme.

Des facteurs de reconfiguration

En l'occurrence, les équilibres socio-économiques bougent. Du point de vue des pouvoirs publics des deux pays, l'attrait stratégique pour l'intelligence artificielle et pour l'émergence de « champions » nationaux est susceptible de lever certaines barrières, à commencer par la faiblesse des aides gouvernementales pour l'implémentation des technologies numériques et l'inertie des institutions éducatives. Si les lignes bougent, des vents contraires comme l'attachement des pouvoirs publics français et états-uniens à la confidentialité des données à caractère personnel, ou leur lutte contre le temps passé par les jeunes générations sur les écrans affleurent comme des obstacles politiques.

Du côté des acteurs privés, un facteur de reconfiguration est l'arrivée de petites entreprises innovantes au croisement des filières de l'informatique pédagogique, de la formation à distance, de l'éducation non-formelle voire des systèmes de gestion de l'information. Communément désignées sous l'étiquette médiatique « *EdTech* » et souvent identifiées comme des « start-ups », elles bénéficient de la popularité de l'intelligence artificielle pour lever des fonds privés et toucher des subventions publiques pour des montants inédits. Concurrençant des oligopoles très concentrés comme ceux de l'édition scolaire, elles constituent cependant pour la plupart des structures de production informatique fragiles du fait de barrières à l'entrée fortes : difficulté d'accès aux institutions éducatives centralisées en France, flotte de commerciaux insuffisantes pour vendre leurs solutions dans les districts dispersés des États-Unis, coûts élevés de production de contenus pédagogiques… De nouvelles opportunités commerciales, et de nouveaux modèles économiques liés notamment à l'intelligence artificielle pourraient-ils offrir des avantages concurrentiels à ces nouveaux entrants innovants et flexibles ?

L'adaptativité pose également des questions sur le positionnement d'acteurs plus anciens. D'abord, les structures de production informatique anciennes considèrent-elle ces nouvelles opportunités industrielles pour revoir leurs offres ? Ensuite, quelles stratégies adoptent les éditeurs scolaires, et notamment ceux qui dominent la distribution des médias éducatifs en France et aux États-Unis, aux niveaux primaire, secondaire, et – dans une moindre mesure – supérieur ? Face à des technologies susceptibles d'affaiblir leur industrie historique, l'édition de manuels, adoptent-ils des stratégies d'obstruction, ou poursuivent-ils au contraire plus fortement leur transition ?

Un dernier aspect structurant est l'influence des géants des technologies sur les offres adaptatives. Les *Big Tech* – multinationales des technologies dont les GAFAM (Google, Amazon, Facebook devenu Meta, Apple et Microsoft) sont les représentants les plus identifiés du côté états-unien – sont à la pointe du traitement algorithmique et de la distribution informatique des plateformes. Ils adoptent de longue date une stratégie de diversification touchant notamment l'éducation. Leur contribution à l'industrialisation des médias adaptatifs doit ainsi être également jaugée.

Une reconfiguration symptomatique de la convergence avec les industries culturelles ?

Ces mutations socio-économiques doivent être jaugées à l'aune du phénomène de convergence entre les industries culturelles et éducatives. Celle-ci est attestée de longue date par les chercheurs s'intéressant à l'industrialisation de l'éducation, *a fortiori* à travers les technologies numériques. La question est de savoir dans quelle mesure les médias adaptatifs l'approfondissent.

Un premier facteur de convergence est la diversification des formes de marchandisation dans l'éducation, au-delà de la traditionnelle logique de service public – supposant la mise à disposition gratuite par les pouvoirs publics de médias éducatifs auprès d'usagers. Des voies privatives se développent depuis plusieurs décennies, telles que les transactions entre éditeurs et usagers privés, dans le cadre de prescriptions institutionnelles (Mœglin, 2005, p.222). C'est par ailleurs en vertu de cette convergence marchande que des concepts de la théorie des industries culturelles (la *socio-économie*, la *filière*, le *modèle socio-économique*) sont applicables à l'éducation. Il s'agira d'observer dans quelle mesure la marchandisation des médias adaptatifs répond à une logique de service public spécifique à l'éducation, ou fait écho à des stratégies commerciales à l'œuvre dans les industries culturelles.

Un second facteur de convergence est la multiplication des innovations technologiques dans ces deux industries, *a fortiori* dans le sens d'une individualisation des services. Dans la musique enregistrée sont apparues notamment des consommations individuelles avec l'avènement du transistor dans les années 1960, culminant aujourd'hui avec les plateformes de streaming musical installées sur des smartphones personnels, adaptant les recommandations à un auditeur selon ses écoutes passées. L'émergence de filières éducatives comme l'informatique pédagogique, supposant un apprentissage individuel face à l'écran, est également symptomatique de cette tendance de fond. À ce titre, l'analogie entre la personnalisation algorithmique des consommations culturelles sur des plateformes numériques et l'adaptativité dans les médias éducatifs mérite d'être questionnée.

Une autre tendance rapprochant l'éducation et la culture, à la fois à l'origine et à la retombée des mouvements précédents, est la financiarisation, ou le « développement de l'actionnariat financier, notamment transnational » selon Philippe Bouquillion (2012). Des mouvements de concentration et de fusion-acquisitions, permis par la financiarisation, ont ainsi conduit ces dernières décennies à la constitution de groupes multimédias réunissant des filiales dans les industries de la culture, des télécommunications, de l'informatique et de l'éducation. Le développement de l'actionnariat est ainsi devenu un moyen de financement central jusque dans les industries éducatives, où les pressions concurrentielles se sont accrues. Par exemple, avec les coûts de recherche et développement induits par la multiplication des innovations technologiques, les petites entreprises et les plus grands groupes doivent solliciter des financements externes accrus. L'émergence des médias adaptatifs semble

à ce titre favoriser un peu plus la financiarisation de l'éducation, dans des directions restant à délimiter.

Une dernière tendance symptomatique de la convergence est l'augmentation des coûts marketing. Devant des offres médiatiques de plus en plus nombreuses et variées, les enjeux de *branding* se sont affirmés dans les industries de la culture, de l'information et de la communication. L'avènement des plateformes numériques a accentué cette tendance avec la quête d'effets de réseau (Katz et Shapiro, 1985), c'est-à-dire d'une masse critique d'usagers devant attirer par sa densité de nouveaux utilisateurs. En éducation, le développement de plateformes de formation à distance, notamment de type MOOC, a répliqué cette même logique (Bullich, 2018). Les algorithmes d'apprentissage adaptatif, reposant en particulier sur le *machine learning,* supposent en outre la quête d'une audience importante pour pouvoir être entraînés sur des jeux de données suffisamment grands. L'ampleur des budgets alloués et la spécificité des rhétoriques marchandes mobilisées par les acteurs des médias adaptatifs restent là encore à questionner.

Les industries éducatives conservent toutefois des spécificités fortes, notamment sur le plan des visées politiques et des modalités d'intervention des pouvoirs publics. Ces spécificités devront être circonscrites dans le cas des médias adaptatifs pour jauger correctement la contribution de ces derniers à la convergence des industries éducatives et culturelles.

Problématique et hypothèses

Notre problématique générale revient donc à déterminer dans quelle mesure le déploiement de médias adaptatifs, reposant *a fortiori* sur le *machine learning*, contribuerait à l'établissement d'un paradigme éducatif singulier, cohérent et transversal aux cultures françaises et états-uniennes et aux secteurs éducatifs.

Notre hypothèse est que ces technologies véhiculent le dépassement d'une forme scolaire centrée sur le balisage spatio-temporel et l'autorité des professionnels sachants, et promeuvent un projet désinstitutionnalisé, centré sur les données, recelant de nouveaux rapports de pouvoir au profit des acteurs des technologies, et faisant avancer la convergence entre les industries culturelles et éducatives.

Pour éprouver cette hypothèse et tenter de trancher les questionnements techniques, pédagogiques et socio-économiques évoqués plus haut, une méthodologie diversifiée doit être mise en place.

La nécessité d'un éclectisme méthodologique

Nous appliquons ces questionnements à la France et aux États-Unis, suivant une méthode contrastive héritée de l'anthropologie structurale (Descola, 2008). Celle-ci consiste à mettre en regard des sociétés différentes sur des aspects communs (comme les normes sociales ou les croyances religieuses), afin de faire émerger des spécificités et des permanences. À la différence d'une démarche comparative qui viserait à jauger les réussites et échecs d'un pays par rapport à l'autre, le but est double : saisir les constantes dans la formulation et le déploiement d'un projet pédagogique

« adaptatif », et au contraire les particularismes conservés selon les cultures nationales et les niveaux d'enseignement.

Nous avons choisi d'appliquer ces questions à la France et aux États-Unis pour quatre raisons principales. Premièrement, ce sont des terrains d'observation heuristiques sur le plan des valeurs politiques attenantes à l'éducation. Ils sont réunis par un idéal démocratique. Confrontée à des défis socio-politiques nombreux depuis l'après-guerre, cette vocation commune est propice au déploiement de l'adaptativité, dans la mesure où les gouvernements cherchent à faire accéder le plus grand nombre de citoyens à des formes d'apprentissage individualisé. Ces pays se distinguent néanmoins par leurs conceptions traditionnelles de l'École au sens large, perçue comme un vecteur d'émancipation culturelle en France, et comme un espace de préparation des individus à la vie sociale et économique aux États-Unis.

En découlent, deuxièmement, des déclinaisons particulières de la forme scolaire sur le plan pédagogique, offrant des perspectives d'industrialisation de l'éducation différentes. Par exemple l'autonomie des enseignants est plus forte en France, où la garantie de la « liberté pédagogique » joue en faveur de formes d'artisanat, alors qu'aux États-Unis leurs pratiques sont bien davantage cadrées – aspect plus propice à l'industrialisation. Ces facteurs sont de nature à favoriser diverses formes d'adaptativité.

Troisièmement, ces pays organisent très différemment leurs systèmes éducatifs – entre une tradition jacobine de centralisation en France, et une autre, fédérale, de décentralisation aux États-Unis. Ces variations concourent à une diversité des configurations socio-économiques, particulièrement intéressante pour observer les stratégies des acteurs des médias adaptatifs.

Enfin, les deux pays sont des terrains intéressants pour observer les recherches en matière d'adaptativité. La prééminence des États-Unis dans la recherche sur les technologies éducatives est attestée de longue date : Marc Scholer indiquait dès les années 1980 que la « technologie éducative » prise dans son ensemble était « très certainement de nature et de facture américaines » (1983, p.77). Dans les années 2000, Alain Chaptal complétait ce constat en notant la contribution essentielle des chercheurs états-uniens à la pensée théorique sur ces techniques : « Les Américains ont joué – et continuent de jouer – un rôle sans équivalent dans la réflexion théorique en matière d'utilisation de technologies éducatives ». (Chaptal, 2003, p.137). Les recherches menées en France, certes moins dotées, sont pourtant très dynamiques, *a fortiori* en intelligence artificielle, épousant des directions propres qu'il conviendra d'expliciter.

Notre approche contrastive se décline dans des méthodologies différentes selon les parties de notre thèse. Dans la première d'entre elles, nous retraçons une histoire du projet scientifique d'adaptativité, afin de faire émerger les lignes de crête scientifiques qui influencent l'offre médiatique actuelle. Il s'agit d'observer l'évolution des acceptions de l'adaptativité selon ses avatars techniques d'une part – des « machines à enseigner » d'après-guerre aux « *intelligent tutoring systems* » d'aujourd'hui. Il convient d'autre part de relever les principaux consensus et controverses liant les chercheurs de disciplines et cultures différentes. Notre méthode consiste en un travail documentaire à partir d'écrits et d'anthologies consacrés aux technologies éducatives, que nous avons filtrées pour retenir les évolutions techniques de l'adaptativité. Afin de saisir les crédos et aspirations des chercheurs actuels, nous avons mené une série d'entretiens, à la fois avec des chercheurs français et états-uniens. Nous avons notamment rencontré ces derniers lors d'une mission de trois jours en septembre 2023

à Pittsburgh, ville abritant les universités de Pittsburgh et Carnegie Mellon, deux des principaux centres de recherche mondiaux en technologies éducatives.

Dans une deuxième partie, nous confrontons le projet scientifique d'adaptativité aux paysages socio-économiques de l'éducation et de la formation aux États-Unis et en France. Il s'agit d'analyser les stratégies des acteurs industriels concourant au déploiement de ces technologies, ainsi que leurs rapports de force, conditionnant leurs contributions respectives à la formalisation d'un paradigme éducatif. L'étude des paysages propres à chaque pays s'effectue en trois temps, correspondant aux niveaux où se prennent les décisions dans les industries éducatives selon Pierre Mœglin (2010) : les politiques publiques, les stratégies de firmes et les professionnalités.

Du point de vue des acteurs publics, il s'agit de saisir les manières dont la demande publique projette la quête d'adaptativité en éducation. Nous comptons mettre en exergue l'évolution rhétorique des discours officiels soutenant le déploiement de techniques de personnalisation éducative, pour cerner des continuités ainsi que des nouveautés dans les arguments politiques en faveur des médias adaptatifs et de leur étiquette médiatique, l'*adaptive learning*. Il convient ensuite d'expliciter les politiques industrielles récentes en matière d'IA éducative, de nature à soutenir et encadrer le développement de filières, par une étude sémio-discursive des appels à projets et programmes de subvention. Nous devons ainsi saisir les contours d'un « filtre officiel » privilégiant certaines formes d'adaptativité plutôt que d'autres. Ces visions doivent cependant être nuancées par une veille documentaire des difficultés principales que rencontrent les échelons territoriaux et les établissements publics dans leur mise en œuvre.

Nous analysons ensuite les points de vue des acteurs privés concourant à l'industrialisation des médias adaptatifs selon trois principaux domaines de valeur : la conception d'algorithmes, la production de contenus pédagogiques et la distribution des médias adaptatifs. Cinq principaux types d'acteurs sont considérés : start-ups *EdTech,* entreprises anciennes du numérique éducatif, éditeurs scolaires, gestionnaires privés de plateformes d'intermédiation et *Big Tech*. Pour chacun d'entre eux, nous menons une analyse socio-économique de « trajectoires » d'entreprises spécifiques. Cette démarche méso consiste dans des recherches documentaires, des entretiens menés avec des dirigeants d'entreprises, et, quand celles-ci sont cotées en bourse, l'analyse comptable de leurs publications financières. Nous interprétons leurs modèles d'affaires à l'aune de la théorie des modèles socio-économiques. Celle-ci nous permet de mesurer la dimension sociale des stratégies marchandes, pour saisir notamment les manières dont ces entreprises projettent des rôles pour l'usager, à la fois en tant qu'apprenant et consommateur. Notre ambition est de souligner ainsi les enjeux concurrentiels entourant l'émergence de l'*adaptive learning* sur les marchés éducatifs, et les rapports de force nouveaux à même de conditionner l'établissement d'un paradigme.

Nous observons enfin la constitution de professionnalités transverses, c'est-à-dire de compétences et d'intérêts partagés par des professionnels variés – haut-fonctionnaires, entrepreneurs, pédagogues, enseignants, *designers*, chefs d'établissements, investisseurs – notamment issus de diverses filières industrielles, de l'édition scolaire à l'informatique pédagogique. Fréquentant les mêmes salons et s'informant aux mêmes médias spécialisés, ils constituent une communauté technophile, souvent rassemblée sous le syntagme « *EdTech* », à même de soutenir et d'influencer l'industrialisation des médias numériques, tant dans les débats politiques que dans les stratégies des firmes. Nous observons d'abord par une veille documentaire la typologie de ces réseaux d'acteurs, leur diversité sociologique et professionnelle. Par l'analyse discursive d'entretiens menés avec

leurs membres et de publications de ces communautés auprès du grand public, nous comptons mettre au jour leurs positionnements paradigmatiques.

Notre troisième partie consiste à saisir, dans un corpus de médias adaptatifs transversal aux divers types de firmes, aux deux pays et aux différents niveaux d'enseignement, les contours et apories d'un projet techno-éducatif dominant. Nous nous tenons ici au croisement de la théorie de l'industrialisation de la formation et de celle de la *trivialité* formulée par Yves Jeanneret, pour mettre au jour les manières dont les visions pédagogiques et industrielles des firmes sont codifiées dans les médias.

Nous cherchons dans un premier temps les traits saillants de ce *modus operandi* industriel à la lumière des trois marqueurs de l'industrialisation de l'éducation formulés par Pierre Mœglin (Mœglin, 1998, 2016). Nous observons en particulier la technologisation et la rationalisation des médias sur deux niveaux interdépendants :

- D'abord en *back-end*, dans les architectures algorithmiques de ces sites, afin d'en savoir plus sur les formes d'adaptativité en présence. Celles-ci doivent être identifiées suivant la triade *adaptive content/assessment/sequencing* proposée par Edsurge. Notre méthodologie consiste dans l'analyse des publications de chaque firme traitant de leur ingénierie, ainsi que dans l'étude de leurs politiques de confidentialité, pour saisir les finalités de la collecte de données.
- En *front-end*, au niveau des interfaces des publics éducatifs (apprenants, enseignants, administrateurs). Nous menons une analyse techno-sémiotique (Souchier et al., 2019) pour saisir d'une part les modes d'existence de ces dispositifs algorithmiques à l'écran, et d'autre part les dispositifs traduisant la capacité des firmes à modeler des habitus, ou des « systèmes de disposition à la pratique » (Bourdieu, 1987, p.94) chez les usagers. Lorsque ce fut possible, nous nous sommes inscrit en tant qu'utilisateur pour mener une « observation participante médiatisée » (Bullich, 2018), sinon, nous avons analysé des ressources documentaires publiées par les firmes.

Nous observons l'idéologisation en traquant les promesses qui parent ces dispositifs algorithmiques dans les discours publics des entreprises ou de leurs représentants. Nous effectuons pour cela une analyse sémio-discursive de documents promotionnels publiés par les firmes ainsi que d'entretiens semi-directifs menés avec leurs dirigeants.

Nous interrogeons en parallèle les enjeux socio-économiques que cristallise ce modèle, à l'aune des stratégies des firmes vues en deuxième partie. Il s'agit ainsi de saisir la contribution de ce *modus operandi* à la convergence des industries éducatives et culturelles.

Dans un dernier chapitre, nous mesurons les limites et incohérences de ce modèle à plusieurs égards. D'abord sur un plan technique, les partis-pris de ces médias doivent être mis en regard avec les angles morts de l'adaptativité observés en première partie. Ensuite sur un plan pédagogique, les traits saillants de cet idéal-type doivent être questionnés à l'aune des positionnements différenciés des firmes. Il s'agit d'identifier les dissensions pouvant exister sur des aspects paradigmatiques, comme par exemple l'agentivité de l'enseignant face aux algorithmes. À travers les trois focales (techniques, pédagogiques, et socio-économiques) adoptées au cours de notre thèse, la prétention d'un « projet adaptatif » au statut de paradigme éducatif peut ainsi être éclairée.

PREMIÈRE PARTIE. La constitution d'un projet scientifique

Interroger le lien entre le déploiement de médias adaptatifs et l'établissement d'un nouveau paradigme éducatif nécessite de s'intéresser en amont aux projets des chercheurs qui développent, depuis des décennies, des machines capables de s'adapter à l'apprenant. Cette première partie vise à saisir les évolutions saillantes des recherches, aux États-Unis, où elles ont débuté dans la première moitié du XX^{e} siècle, et en France, où elles ont été approfondies dans des directions propres.

Deux aspects font l'objet d'attention. D'une part, les évolutions techniques de l'adaptativité doivent être explicitées, par un travail documentaire à partir d'écrits et d'anthologies traitant des technologies éducatives. Nous étudions au fil des époques la façon dont des artefacts ont fait avancer l'idéal d'adaptativité vers ses trois critères définitionnels détaillés en introduction : médiatisation, captation des besoins individuels en continu et flexibilité de la pédagogie. Doivent apparaître ainsi la diversification des aspects et des méthodes d'adaptation, ainsi que l'évolution des usages pédagogiques – d'une utilisation en présentiel ou à distance, seul face à l'outil ou en groupe, avec ou sans l'assistance du formateur.

D'autre part, ces innovations doivent être présentées à l'aune des principales théories pédagogiques qui ont inspiré les chercheurs au fil du temps. Doit émerger ainsi une histoire des controverses pédagogiques ayant fait évoluer l'adaptativité, tant dans ses avatars techniques que dans ses usages éducatifs. Outre notre travail documentaire, les crédos actuels de la communauté scientifique ont été investigués via une série d'entretiens semi-directifs avec des chercheurs français et états-uniens.

Nous observons dans cette partie que ces médias s'inscrivent dans une histoire conflictuelle, entamée par les tenants du behaviorisme, et consacrant progressivement des préceptes constructivistes. Ce dualisme théorique renseigne les configurations pédagogiques projetées dans les outils : un glissement s'opère du « *drill and practice* » behavioriste – privilégiant l'entraînement de l'élève par des batteries d'exercices, avec ou sans le secours du formateur – à des situations d'apprentissage diversifiées propices aux préceptes constructivistes – d'un face à face avec l'outil inspiré du tutorat, capable d'orienter et conseiller l'apprenant dans son apprentissage, à des usages collaboratifs intégrant élèves et enseignants grâce aux possibilités du web.

Cette histoire scientifique n'est cependant pas celle d'un effacement du behaviorisme au profit du constructivisme, car il est rare que les axes de recherche s'éteignent totalement. Les survivances « sous les radars » sont légion dans le domaine des technologies éducatives. À propos des théories ayant motivé leur conception au XX^{e} siècle, Alain Chaptal note :

> « Le vingtième siècle va voir se développer divers courants qui exerceront tour à tour une influence déterminante. Il est […] intéressant de noter que cette évolution ne se déroule nullement selon une logique linéaire […]. En réalité, les mouvements de la pensée sont infiniment plus riches et complexes et les influences se croisent, s'interprètent et continuent à s'exercer plus ou moins souterrainement même dans les périodes d'éclipse » (Chaptal, 2003, p.144).

Il nous faut notamment expliciter la distinction commune, au sein des *Learning Sciences* contemporaines, entre deux disciplines, l'*Educational Data Mining* et les *Learning Analytics,* ayant émergé dans les années 2000 avec l'essor du *Big Data* en éducation. La première applique les méthodes de fouille de données aux environnements informatiques pour accroître les capacités

descriptives et prédictives des algorithmes, quand la seconde vise à assister les publics éducatifs par des algorithmes d'aide à la décision et des systèmes de visualisation. Ce sont *a priori* deux courants proches projetant différemment l'apprentissage – en vase clos, face à la machine, pour l'un, et en présence d'autres apprenants et de formateurs pour l'autre. Cet antagonisme réactualise-t-il l'opposition entre behaviorisme et constructivisme ? Dans quelle mesure affecte-t-il la formulation d'un projet éducatif cohérent dans la sphère académique ?

Nous traitons en parallèle les apports épistémologiques extérieurs. Nous nous efforçons notamment de circonscrire les domaines où l'IA est mobilisée, et suivant quelle acception de l'apprentissage. Nous avons en effet décrit la lutte d'influence entre deux approches internes au courant, dites « symbolique » et « connexionniste ». Dans quelle mesure la voie connexionniste dominante depuis l'avènement du *Big Data*, « considér[ant] que penser s'apparente à un calcul massivement parallèle de fonctions élémentaires » (Cardon et al., 2018, p.178) et réduisant ainsi l'explicabilité des décisions algorithmiques, est-elle partagée au sein des communautés de chercheurs ? Quelles en sont les implications sur les usages pédagogiques projetés ?

Notre argumentation se partage entre les époques antérieures et postérieures à l'avènement du *Big Data*, communément daté au tournant des années 2000 et 2010. Le premier chapitre questionne l'état de la recherche antérieure aux données massives sur les deux rives de l'Atlantique, des machines à enseigner de l'entre-deux-guerres aux environnements informatiques pour l'apprentissage humain des années 2000. Nos deux chapitres suivants traitent de l'état de la recherche « post-*Big Data »*. Le premier est consacré à la distinction entre *Educational Data Mining* et *Learning Analytics,* à l'aune de leurs préceptes théoriques et méthodologiques. Le second aborde les principales limites techniques et épistémologiques des dispositifs expérimentaux, conditionnant l'impulsion d'un projet éducatif cohérent vers les organisations éducatives.

Chapitre I. L'émergence tumultueuse de l'adaptativité

Au début du XX^{e} siècle, *adaptativité* n'a pas encore pris son sens pédagogique, et n'est usité que dans les écrits de biologie darwinienne, pour signifier la capacité d'un être vivant à s'adapter à son milieu. En revanche, le début du siècle est le théâtre, outre-Atlantique, d'une ébullition théorique décisive pour les développements futurs des technologies éducatives.

Comme le dépeint l'historien Paul Saettler dans son anthologie de 1967 sur les technologies éducatives, une concurrence débute entre deux philosophies : d'un côté le connectionisme d'Edward Thorndike et de l'autre le courant progressiste de John Dewey (Saettler, 1967, p.52). Les deux penseurs de l'éducation ont certes des points communs : d'abord l'héritage d'un maître à penser, le philosophe pragmatiste William James, selon qui la vérité d'une idée réside dans la réussite de l'action ; ensuite une volonté de réformer la pédagogie.

Thorndike et Dewey se différencient néanmoins radicalement sur les moyens pour y parvenir, et notamment sur la question de l'individualisation pédagogique – prémisse de l'adaptativité, en ce qu'elle promeut une déclinaison de l'enseignement au niveau de l'individu. John Dewey perçoit l'éducation comme le moteur du progrès social. L'élève doit apprendre en expérimentant, notamment en groupe, sous l'œil de l'enseignant chargé d'orienter les expériences et éviter l'enfermement dans l'immédiateté et la répétition (Dewey, 1968, p.116). Précurseur du behaviorisme, Edward Thorndike rejette en revanche toute problématique liée au fonctionnement de l'esprit humain, se focalisant sur l'observation expérimentale du comportement à l'échelle individuelle. Selon lui l'élève apprend essentiellement par répétition, une réponse suivie de plaisir s'en trouvant renforcée. Considéré comme le père fondateur de la science de l'éducation, sa pensée va le plus imprégner les premiers développements des technologies éducatives (*educational technology*) au XX^{e} siècle.

Le livre tournant automatiquement ses pages selon la maîtrise de l'élève, tel que Thorndike l'imaginait dans la citation en exergue de notre thèse, équivalait à un espoir d'adaptativité sommaire, sans que le terme ne soit à l'époque posé. En 1926, Sidney Pressey présente un nouveau type de machines qui exauce son vœu dans le cadre d'évaluations, la *Pressey Testing Machine*. L'élève répond à une question à choix multiples affichée sur un rouleau en actionnant une des touches disponibles. Si sa réponse est correcte, une autre question apparaît, sinon le rouleau reste immobile et il lui faut réessayer.

On retrouve partiellement, et dans des termes simplistes, les trois éléments définitionnels de l'adaptativité : la présence d'un média automatisant une situation d'apprentissage, sa capacité à prendre en compte les besoins de l'apprenant, en l'occurrence sa maîtrise ou non de la notion sous-tendue par la question, et enfin sa capacité à adapter un aspect pédagogique, le passage à la question suivante. Manquent encore la collecte et le traitement en continu des besoins de l'élève.

Sommaire, l'invention n'en demeure pas moins séminale, et s'inscrit en faux contre le paradigme traditionnel états-unien, comme l'explique Éric Bruillard (1997, p.35) citant Arthur Lumsdaine et Robert Glaser :

> « Pressey [est] motivé par une volonté de réformer la traditionnelle récitation par un élève en classe, qu'il [juge] inefficace, et [espère] ainsi « construire des dispositifs pouvant poser une question et permettant à l'élève

d'y répondre et d'avoir une confirmation ou une information immédiate de la véracité de cette réponse » (Bruillard, 1997, p.35).

L'inventeur espère libérer l'enseignant de la contrainte d'éditer des évaluations, et lui dégager du temps pour se consacrer à des tâches plus stimulantes – rhétorique que l'on retrouvera dans les discours autour de l'*adaptive learning.* Il s'agit aussi d'éviter un délai entre la production d'un travail par l'élève et sa correction par un professeur. Cette machine annonce ainsi le courant industriel de l'« enseignement programmé ».

1. L'enseignement programmé : l'individualisation à l'heure du behaviorisme

Comme l'explique Éric Bruillard, l'enseignement programmé est avant tout un « courant de recherche », qui « utilise des machines à enseigner, mais elles ne sont que des instruments, elles n'enseignent pas ; ce sont les programmes supportés par les appareils qui contiennent les éléments pédagogiques et servent à instruire » (Bruillard, 1996, p.36). Ce courant théorisé par le psychologue états-unien Burrhus Skinner a pour ambition de mobiliser des méthodes scientifiques, empruntées en particulier à la psychologie comportementale, pour organiser au mieux le découpage de la matière enseignée sous la forme de programmes. Il convient d'expliciter ici les préceptes théoriques qui sous-tendent cette démarche, liés à deux courants de pensée dominants à l'époque : la cybernétique et le behaviorisme.

Le sillage de la cybernétique et du behaviorisme

Dans l'immédiat après-guerre, la formalisation par Norbert Wiener d'une nouvelle théorie de la communication, la cybernétique, donne de l'épaisseur à la thèse selon laquelle l'éducation pourrait être pilotée par des machines et favoriser certains comportements. Il formule une théorie de l'efficacité communicationnelle, supposant des systèmes vivants ou artificiels capables d'adapter leurs réactions à leur environnement par des échanges d'informations, au moyen de rétroactions ou *feedbacks*. Une lecture éducative de la cybernétique est ensuite proposée par le mathématicien allemand Helmar Frank, qui se nourrit d'une critique de la pédagogie conçue comme un art, et d'une volonté de mobiliser des méthodes scientifiques pour la contrôler efficacement. Comme le restitue Éric Bruillard, cette doctrine prétend que « l'enseignement traditionnel se caractérise par une insuffisance de rétroaction de l'élève au maître qui engendre un grave défaut dans le contrôle du processus d'enseignement » (Bruillard, 1996, p.29). L'application aux technologies éducatives de ce principe d'adaptation amène les chercheurs à penser des formes d'adaptativité avant l'heure.

Si le lien entre la machine et le cerveau n'est qu'analogique pour les fidèles de Wiener, la modélisation mathématique d'un réseau de neurones, en 1943 par le neurophysiologiste Warren McCulloch et le logicien Walter Pitts, ouvre une voie plus radicale. Comme le relève Jean-Pierre Dupuy, pour ces scientifiques, le lien est ontologique : le cerveau *est* une machine computationnelle (Dupuy, 2002, p.43). Cette radicalisation est un élément fondateur pour la future discipline de l'intelligence artificielle[4] et son courant dit « connexionniste », qui définit l'intelligence comme une propriété émergente de la connexion des neurones, naturels ou artificiels.

[4] Il faudra attendre la conférence de Dartmouth de 1956 pour qu'une communauté scientifique se forme autour de ce syntagme.

L'ambiguïté de la cybernétique sur le lien entre machine et cerveau entraîne une ambivalence de la recherche en technologies éducatives à propos du statut de l'enseignant : est-il un professionnel que la machine doit assister, suivant la conception de Wiener ? ou lui-même un type de machine capable d'enseigner, et en cela remplaçable artificiellement, selon la conception de McCulloch et Pitts ? Pour reprendre les termes de Pierre Mœglin, la technologisation doit-elle être additive ou substitutive (Mœglin, 2016, p.16) ? Si aujourd'hui, cette dernière position est, semble-t-il, rejetée par une grande majorité des chercheurs, l'héritage connexionniste du secteur, en termes de méthodes et de concepts, entretient le débat, comme on va le voir plus bas. Après-guerre, la cybernétique de Wiener influence grandement les partisans de l'usage de technologies éducatives en Europe comme le relève Éric Bruillard (1997, p.27 ; p.33), alors qu'aux États-Unis, une autre théorie innerve plus directement l'enseignement programmé : le behaviorisme.

Cette conception psychologique du comportement inspirée par Thorndike et théorisée dans les années 1910 par John Watson est dominante dans la sphère académique états-unienne d'après-guerre. Le maître à penser de l'enseignement programmé, Burrhus Skinner, place l'action de l'apprenant au centre des considérations. À partir d'expérimentations sur des rats et des pigeons, il propose le concept de « conditionnement opérant », stipulant qu'un être vivant apprend un comportement à mesure qu'il reçoit des conséquences positives lorsqu'il le reproduit (liées à une récompense alimentaire par exemple).

Skinner est animé par un idéal contestataire, comme le rappelle Maria Casas. Il voue aux gémonies l'enseignement standardisé et collectif et cherche à mener de front industrialisation et individualisation, grâce aux économies d'échelles réalisées par les machines. Il fait valoir, après avoir observé la classe d'arithmétique de sa fille, que les élèves devraient non pas attendre une correction après avoir effectué leurs devoirs mais obtenir des « renforcements immédiats » – en écho à Sidney Pressey –, et travailler non pas sur des exercices communs à tous mais adaptés à leurs niveaux (Casas, 2002, p.3). Il pose alors les bases de l'enseignement programmé autour de cinq règles :

- La nécessité de fractionner la matière en unités plus petites.
- Le fait pour l'apprenant de ne répondre qu'à une question à la fois.
- La vérification d'un savoir par la réalisation d'une tâche concrète.
- Le fait de ne passer à l'unité suivante qu'après avoir satisfait aux obligations de la précédente.
- Le fait d'avancer à son rythme.

Ces principes appellent une organisation rationnelle et scientifique de la pédagogie, découpée en micro-actions. Cette tendance est une référence au taylorisme, doctrine industrialiste exposée en 1911 par l'ingénieur Frederick Taylor, visant l'efficacité dans le travail en usine, appliquée ensuite par les tenants du *Social Efficiency Movement* aux services publics, dont l'école.

Ils induisent une individualisation semblable à celle pratiquée par la machine de Pressey. Cependant il ne s'agit plus seulement d'évaluer l'élève, mais aussi de lui enseigner une matière. La méthode d'apprentissage privilégiée est le « *drill and practice* », la répétition d'exercices jusqu'à ce que l'élève prouve sa maîtrise. Ces cinq propositions, peu innovantes (Mœglin, 2005, pp. 143-144), restent toutefois de « bon sens » et consensuelles. Elles vont permettre de rallier à partir des années 1960 des pédagogues, chercheurs en technologies éducatives, entrepreneurs et autres professionnels de l'éducation technophiles dans une science à vocation industrielle : l'*educational technology* aux États-Unis, traduite en « technologie éducative » en France. Ces acteurs

aux aspirations variées se retrouvent selon Guy Berger (1982) sur la volonté de faire bouger les lignes de systèmes éducatifs jugés sclérosés autour de l'enseignement collectif.

Ce mouvement scientifique et industriel gagne en crédit auprès des pouvoirs publics états-uniens, sur fond de crises éducatives. Témoin le *National Defense of Education Act* (NDEA), ratifié en 1958 en réponse notamment aux craintes de déclassement du pays suscitées par le lancement du premier satellite par les Soviétiques la même année. Ce texte prévoit, dans son *Title VII*, des subventions aux éditeurs d'outils d'enseignement programmé, en vue de relancer le système éducatif national. Skinner en perçoit notamment pour ses machines dans les années 1960 (Casas, 2002, p.17).

Considérations paradigmatiques

Toujours est-il qu'entre les courants français et états-unien, des différences importantes apparaissent, concernant notamment le rôle du professeur. Comme l'écrit Pierre Mœglin, « chez les Anglo-saxons et sous l'influence du taylorisme, l'*Educational Technology* n'attribue à l'enseignant aucun rôle prééminent. Ce sont l'organisation et les organisateurs qui constituent la cheville ouvrière du dispositif » (Mœglin, 2005, pp.146-147). En France en revanche, la place de l'enseignant est historiquement plus forte. « La référence fayolienne est vivace, quoique rarement nommée. L'enseignant exerce l'ensemble des responsabilités y compris celle de l'organisation » (Mœglin, 2005, p.147). Alors qu'aux États-Unis l'on attend de la technologie qu'elle adapte l'enseignement à chacun, en France l'on se félicite du temps qu'elle libère pour permettre à l'enseignant d'adapter lui-même son cours. Ces deux perceptions du potentiel technologique sont visibles aujourd'hui dans la conception des environnements numériques d'apprentissage, comme nous le verrons par la suite.

Dans le cas de l'enseignement programmé, la promesse de libérer du temps pour les enseignants apparaît rapidement illusoire. Les machines inspirées par les préceptes de Skinner, loin de permettre des économies d'échelle, rencontrent des échecs commerciaux dans les organisations éducatives. Témoin les *Didak 501* et *Min-Max*, lancées respectivement par Rheem Califone et Grolier. De fait, ces machines sont peu flexibles et peu adaptatives, du fait de leurs programmes linéaires.

L'ingénieur états-unien Norman Crowder apporte cependant une avancée décisive en termes d'adaptation avec un nouveau type de programme, dit « à branchement ». À la différence des parcours linéaires de Pressey et Skinner, celui-ci prend la forme d'une arborescence, dont l'apprenant suit des ramifications adaptées à l'avance aux erreurs qu'il est susceptible de commettre. Norman Crowder appartient à une autre école : à la différence de Skinner, il valorise l'erreur comme une composante de l'apprentissage, permettant de construire des parcours sur-mesure.

Inspirées des livres brouillés, les machines de Crowder, les *Autotutor Mark* (Figure 1) sont des terminaux contenant des bobines de films photographiques. Celles-ci affichent des questions à choix multiples auxquels correspondent les boutons à droite de l'écran. Novatrices, ces machines rencontrent en revanche un succès commercial très relatif, en raison de leurs coûts élevés. Au-delà de certaines institutions équipant leurs classes à titre expérimental comme la New York Institute of Technology, le déploiement est très limité. La prophétie formulée par le concepteur en 1962, selon laquelle il faudrait moins de trois ans pour que la moitié des étudiants américains travaillent sur des machines, s'avère illusoire (Booms, 2024).

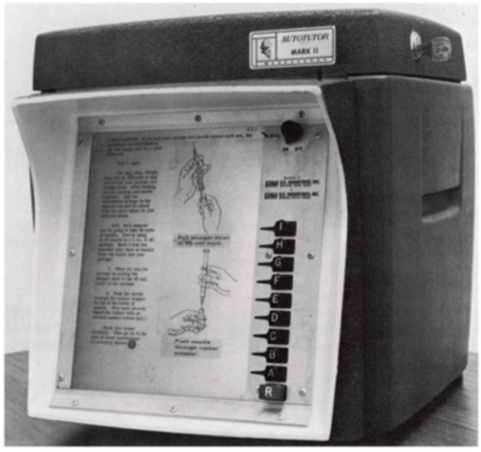


Figure 1. Autotutor Mark II *de Norman Crowder. Source : Etzwiler et al., 1972.*

Autre avancée technique rendant les machines plus flexibles : les systèmes dits « adaptatifs » du britannique Gordon Pask. L'inventeur du *Self-Adaptive Keyboard Instructor* (SAKI) est l'un des premiers à promouvoir explicitement l'idéal d'adaptativité (*adaptivity*), auquel il adjoint « *self* » dans ce qui apparaît aujourd'hui comme un pléonasme. Le SAKI vise à entraîner des élèves au maniement de la poinçonneuse Hollerith – un appareil qui a servi à enregistrer des valeurs numériques avant l'arrivée de l'informatique, via des cartes perforées à des emplacements spécifiques. Le SAKI doit alors faire apprendre à l'élève la correspondance entre les touches et les endroits à perforer. À l'écran apparaît un chiffre, et une diode indiquant la touche sur le clavier s'allume. L'idée est de faire intérioriser à l'apprenant la position des chiffres sur le clavier, pour qu'il les saisisse de plus en plus rapidement. Pendant ce temps, la machine enregistre les réponses et construit un modèle probabiliste du processus d'apprentissage. Si la machine détecte une difficulté pour retenir la position d'une touche, son affichage devient de plus en plus fréquent et le rythme ralentit, pour inciter l'opérateur à prendre son temps. Les nombres bien retenus arrivent au contraire plus vite et l'éclat des diodes devient moins fort (Figure 2).



Figure 2. Exemple de manipulation du SAKI de Gordon Pask. Source : Watters, 2015

Cette conception de l'adaptativité est plus complexe que les précédentes, et annonce les partis-pris des dispositifs actuels. Dans un article de 1958, Pask défend notamment le caractère relatif de la difficulté – un présupposé de nombreux médias adaptatifs actuels :

> « Le seul sens que l'on puisse donner à la « difficulté » est ce que l'apprenant trouve difficile. Il ne servirait à rien de construire notre propre idée de la difficulté dans une machine à enseigner, et encore moins une sorte de difficulté moyenne – par exemple une échelle obtenue en faisant la moyenne des résultats d'un certain nombre de tests utilisant ce dispositif sur des sujets différents. Cette mesure de la difficulté pourrait être parfaitement valable sur la moyenne, mais elle ne s'appliquerait presque jamais à un individu spécifique. En fait, même chez une seule personne, quelque chose jugé difficile à un moment donné sera jugé facile le lendemain »[5] (Pask, 1958).

Toutefois, le scénario de l'échec commercial se répète. En 1961, les droits de fabrication de SAKI sont cédés à Cybernetic Developments, qui ne vend qu'une cinquantaine d'appareils (Watters, 2015). Outre ces désillusions, et quelles que soient les avancées et les complexifications techniques, l'enseignement programmé apparaît de plus en plus sous le feu des critiques au tournant des années 1950 et 1960.

Le temps des critiques

Sur les préceptes théoriques, la critique virulente du linguiste Noam Chomsky rencontre un écho certain dans les sphères académiques. En 1959, il réfute les idées behavioristes de Skinner sur le comportement verbal, selon lesquelles les hommes et les animaux apprendraient de la même façon, et les résultats obtenus en laboratoire pourraient être transférés dans la réalité. Dans les salles de classe, ces machines qui ont se sont déployées grâce aux subventions publiques suscitent une déception à la hauteur des promesses formulées par les industriels.

Nombre de chercheurs affiliés au courant déplorent le manque de méthodes expérimentales. Comme l'écrit Éric Bruillard, « les recherches se heurtent à une forme de barrière de complexité qui se dresse dans les tâches les plus classiques de l'enseignement programmé : l'analyse de la structure du contenu à enseigner, les théories de l'apprentissage et la gestion de l'individualisation » (p.49). Sur la première limite, le secteur souffre d'un trop grand morcellement des contenus enseignés, ne permettant pas aux machines d'accompagner les apprenants dans leur compréhension synthétique des domaines. Concernant les théories de l'apprentissage, les angles morts de la psychologie comportementale, notamment sur la façon dont l'élève conceptualise le sujet et l'apprend, sont mis en cause. Dans une autocritique sans concessions, Richard Goodman, directeur de la section *Calcul et Cybernétique* au Collège de Brighton, écrit :

> « C'est une erreur flagrante […] de prétendre que les principes psychologiques sont connus depuis longtemps. Tout ce qu'on peut dire, c'est que nous avons là-dessus un vague agrégat de notions acquises empiriquement, pragmatiquement, ce qui ne constitue pas une base bien satisfaisante pour construire des machines à enseigner » (Goodman, 1964, traduit par Bruillard, 1997, p.50).

Quant à la troisième limite, le défaut d'individualisation, celui-ci découle d'une part des deux manques précédents, et d'autre part de la faiblesse technique des machines. Quels que soient la nature

[5] *« The only meaning which can be given to 'difficulty' is something which this particular trainee finds difficult. There would be little point in building our own idea of difficulty into a teaching machine, and still less an average difficulty scale, for example, a scale obtained by averaging the results of a number of tests using this exercise material presented to different subjects. This average measure of difficulty might be perfectly valid on the average, but it would almost certainly never apply to a specified individual. In fact, even for the same individual, something deemed difficult at one moment will be rated easy the next »*

des programmes et leur degré de complexité, les outils se révèlent incapables d'enregistrer durablement les résultats des élèves et d'opérer des arbitrages pédagogiques sur cette base. Le secteur se découvre alors un dilemme qu'expose Bruillard en ces termes : « soit on limite le nombre de chemins et on perd l'aspect individualisation, soit on le multiplie ce qui induit une grande complexité » (ibid., p.51).

L'arrivée des ordinateurs et de leurs puissances de stockage et de calculs donne lieu au développement de programmes plus puissants[6]. Débute ce qu'on appelle alors l'« enseignement assisté par ordinateur » (EAO). Les premiers avatars éducatifs poursuivent l'agenda behavioriste, opérant quelques ajustements par rapport aux préceptes des « pionniers ».

Les premiers programmes informatiques à usages éducatifs redonnent un second souffle à l'enseignement programmé. Des systèmes sophistiqués germent dans les laboratoires, à l'image du programme de géométrie de l'informaticien du MIT Richard Smallwood (1962). D'architecture crowderienne, il pratique une forme précurseure d'apprentissage automatique, via un arbre de décision lui permettant de modifier ses paramètres au fur et à mesure de son utilisation. Les décisions de branchement sont ainsi opérées par comparaison entre un élève et ses prédécesseurs selon leurs pourcentages de bonnes réponses passées. Sans théorie pédagogique filtrante, une explosion combinatoire des décisions pédagogiques a précipité l'échec du dispositif (Bruillard, 1997, p.119). Cependant, selon Éric Bruillard, ce programme demeure le premier à adapter de façon autonome sa stratégie pédagogique. Il a en effet ouvert la voie d'une adaptativité basée sur le *machine learning*. Dans un article de 1962, Richard Smallwood promouvait explicitement cet idéal pour décrire sa démarche :

> « Nous souhaitons que la décision pédagogique dépende non seulement des réponses passées de l'élève, mais aussi d'autres informations disponibles, y compris les réponses de tous les anciens étudiants ; en d'autres termes, le système de décision pédagogique doit être un système adaptatif »[7] (Smallwood, 1962, p.10).

De fait, cette adaptativité est semblable à de l'*adaptive assessment,* définie dans la trilogie d'Edsurge comme la capacité pour un système d'adapter la teneur et la difficulté d'une question en fonction du résultat obtenu par l'élève à la précédente (Edsurge, 2016).

Avec l'ordinateur sont également explorés d'autres formats d'exercices que le questionnaire à choix multiples. Il y a par exemple la résolution de problèmes, où arrivent les premiers programmes dits « génératifs ». Ceux-ci utilisent des algorithmes pour générer problèmes et solutions. Ils permettent un gain de temps dans la conception des exercices. En outre, leurs fonctions de résolution de problèmes, les « résolveurs », leur permettent de saisir les différentes étapes de résolution d'un problème. Ils commencent à être complétés par des algorithmes de sélection du concept adéquat à afficher, permettant une forme d'adaptativité.

Témoin les travaux de l'informaticien états-unien Elliott Koffman qui développe en 1972 un programme destiné à l'apprentissage de concepts logiques sur ordinateur. Il comporte un algorithme de sélection du concept à enseigner dans un arbre de notions figurant le domaine. Cela permet d'adapter dynamiquement le parcours de l'apprenant suivant son niveau. Koffman demandait

[6] Les années 1960 voient arriver les premiers circuits intégrés et microprocesseurs permettant de miniaturiser les immenses ordinateurs existant depuis les années 1940, incompatibles avec des usages pédagogiques individuels et massifiés.

[7] « *We should like for the next tutoring decision to depend not only on the present student's past responses, but also on other available information including the responses of all past students; in other words, the teaching decision system should be an adaptive system* »

au préalable au professeur de découper son cours en concepts élémentaires et d'en représenter l'enchaînement par un graphe. Chaque concept contient une batterie d'exercices de niveaux variables, générés automatiquement (Coulon et Kayser, 1975).

Aux États-Unis comme en France, le soutien politique est marqué : des bourses de recherche sont accordées outre-Atlantique, notamment à l'Université de l'Illinois où, de concert avec l'entreprise Control Data, est développé le projet *Plato*. Celui-ci débouche sur des ordinateurs pionniers pour le courant de la « *Computer-Assisted Instruction »* (CAI). En France, le Ministère de l'Éducation nationale lance en 1970 la première expérimentation autour de l'enseignement de l'informatique au lycée.

Cependant, à rebours des collaborations possibles entre enseignants et ordinateurs, l'essor du secteur inspire à de nombreux promoteurs des déclarations excessives vantant la substitution pure et simple de la machine au formateur. Louis Couffignal, chef de file de la cybernétique en France, préconise de « remplacer le professeur par le processeur ». Le philosophe américain Patrick Suppes écrit en 1966 que « l'on peut prévoir que dans peu d'années, des millions d'écoliers auront accès à ce qui était le privilège royal d'Alexandre, fils de Philippe de Macédoine : les services personnels d'un tuteur aussi bien formé et réactif que l'était Aristote » (Suppes, 1966, traduit par Mœglin, 2010, p.47). Quatre ans plus tard, il confesse pourtant l'échec de l'EAO à enseigner seul un domaine à l'apprenant, et reconnaît que celui-ci ne peut en tirer un bénéfice qu'à petites doses et en complément d'un cours ordinaire.

Les limites de l'enseignement programmé soulignées plus haut restent entières. De fait, comme le note le psychologue Pierre Oléron (1965) cité par Bruillard (1997, p.41), l'enseignement programmé n'adopte que le point de vue de l'enseignant. La capacité d'adaptation de ses décisions pédagogiques ne peut être améliorée que dans l'écriture de son programme, et non dans son aptitude à « comprendre » les besoins de l'apprenant. Les inconvénients sont de fait nombreux du point de vue de l'étudiant : celui-ci n'a guère la possibilité de s'exprimer lorsque la question ne lui semble pas claire, ou que la réponse qu'il souhaiterait apporter ne figure pas dans le choix offert. Pour les informaticiens, la programmation est en outre très fastidieuse, à cause de la combinatoire entraînée par la nécessité de prévoir chaque branchement. Commercialement, il en découle des coûts de conception très élevés qui dissuadent les jeunes industriels de l'informatique – Microsoft et Apple notamment – d'investir dans les contenus.

Dans l'espace scientifique, le behaviorisme a en outre perdu depuis longtemps ses lettres de noblesse. Comme l'écrit Alain Chaptal, « curieusement, le behaviorisme s'[est] développé dans le domaine éducatif au moment où son influence philosophique générale déclin[e] au profit de la psychologie cognitive » (Chaptal, 2003, p.154). L'emprise des sciences cognitives va se cristalliser dans l'avènement, en pédagogie, du constructivisme. Par sa volonté d'ouvrir la « boîte noire » de la cognition individuelle laissée fermée par le behaviorisme, cette théorie explore davantage les possibilités multiples offertes par l'ordinateur. L'implication de cet *aggiornamento* est cruciale pour les voies d'adaptativité investiguées par la suite. Deux avatars techniques émergent alors pour dépasser le *drill and practice* caractéristique des machines à enseigner : tuteurs intelligents et micromondes. Ils engagent une mutation dans le profil des inventeurs, d'ingénieurs commercialisant leurs « machines à enseigner » à des chercheurs en informatique s'en tenant davantage au monde académique. Nous étudions ci-après le premier type de dispositifs qui en découle, les « tuteurs intelligents ».

2. Au-delà du *drill and practice,* les « tuteurs intelligents »

Avec les développements de l'informatique à la fin des années 1960, les nouvelles possibilités d'interaction permettent de varier les modalités pédagogiques. Le tutorat devient un mode pédagogique de référence pour les informaticiens. Ceux-ci programment des systèmes capables d'interagir et de répondre aux questions de l'élève. L'adaptation des systèmes aux besoins non-formalisés par l'apprenant progresse dans l'affectation de certaines tâches de résolution à l'apprenant. Pour ce faire, des algorithmes opèrent des calculs mathématiques et des raisonnements logiques à partir de trois modules. Ceux-ci stockent des données numériques pour modéliser l'élève, les stratégies pédagogiques et le domaine à enseigner. L'adaptativité progresse ainsi dans une voie « symbolique » à travers cette architecture tripartite, qui devient la norme pour concevoir des programmes informatiques. Celle-ci découle de la prégnance des sciences cognitives, interdiscipline qui s'impose dans les *Learning Sciences* dès les années 1960.

L'avènement des sciences cognitives et la survivance du behaviorisme

À partir des années 1950, une certaine effervescence touche le champ scientifique, adoptant une manière nouvelle d'appréhender l'esprit. Loin d'une « boîte noire », celui-ci est dorénavant considéré comme un phénomène biologique se construisant chez l'individu au contact du monde extérieur. Dès lors, les sciences s'efforcent de décrire voire simuler des phénomènes tels que l'acquisition du savoir, le langage et l'intelligence. C'est ainsi que se forment les sciences cognitives, « nébuleuse de disciplines » selon les mots de Jean-Pierre Dupuy (1994), allant de la philosophie à l'intelligence artificielle en passant par la linguistique, et, en éducation, la psychologie constructiviste.

Cette dernière émerge à la suite de la traduction, aux États-Unis, des ouvrages de Jean Piaget sur le développement de l'intelligence. La théorie de l'apprentissage du psychologue suisse repose sur trois concepts fondamentaux : les schèmes – structures mentales par lesquelles l'individu comprend son environnement –, l'assimilation – processus par lequel l'individu intègre de nouvelles informations dans ses schèmes –, et enfin l'accommodation – la modification ou la création de schèmes par l'individu. La diffusion de cette théorie aux États-Unis est décisive pour la recherche sur les « tuteurs intelligents ». Comme l'écrit Alain Chaptal, « les travaux de Piaget fourniss[ent] une base conceptuelle théorique permettant de développer une technologie éducative fondée sur une évaluation individuelle du développement cognitif » (2003, p.157). En découle le constructivisme, qui devient la principale théorie concurrente du behaviorisme.

Autre discipline nouvelle émergeant avec les sciences cognitives et présentant des conséquences pour l'éducation, l'intelligence artificielle procède de la redécouverte de la logique booléenne à travers l'ordinateur. Dominique Cardon, Jean-Philippe Cointet et Antoine Mazières (2018) en retracent la naissance, sous le patronage de figures comme Marvin Minsky au MIT, ou John McCarthy à Stanford :

> « L'IA naît la même année que les sciences cognitives (1956) et les deux domaines vont façonner ensemble le projet de doter les ordinateurs d'une capacité de raisonnement (Gardner, 1985). Contre la psychologie behavioriste qui avait inspiré les « boîtes noires » adaptatives de la cybernétique, le projet des sciences cognitives est de donner à la machine des capacités abstraites et logiques » (Cardon et al., 2018, p.188).

Cette conception originelle, qualifiée plus tard de « symbolique », a pour ambition de répliquer la pensée humaine en « calcul[ant] des symboles qui ont à la fois une réalité matérielle et une valeur sémantique de représentation » (Cardon et al., 2018, p.178). Il s'agit alors d'ouvrir la « boîte noire » de la cognition pour tenter de programmer l'ordinateur en conséquence. En éducation, Herbert Simon et Alan Newell, de l'université Carnegie Mellon, investiguent la notion d'« états mentaux » chez l'apprenant, afin de complexifier la pédagogie embarquée dans la machine. Comme l'écrit Éric Bruillard, « l'apport de l'intelligence artificielle apparaît comme celui d'une ingénierie de la connaissance » dont il formule ainsi les questionnements phares : « Quelles connaissances peut-on représenter dans une machine et quelles améliorations peuvent-elles apporter dans l'enseignement avec l'ordinateur ? » (Bruillard, 1996, p.70).

En linguistique, les préceptes behavioristes dominants laissent la place à l'analyse des structures sous-tendant la compréhension et l'expression chez l'individu, sous l'influence des structuralistes comme Noam Chomsky. La rencontre de ce paradigme avec l'intelligence artificielle sera au cœur des recherches en « traitement automatique du langage naturel » (TALN), ou « *natural language processing* » (NLP). Visant à rendre les machines capables de générer du discours, ce domaine aura des implications décisives pour le développement des « tuteurs intelligents ».

Le behaviorisme n'a cependant pas dit son dernier mot. À la fin des années 1960, deux psychologues états-uniens renouvellent ses préceptes en formulant des théories pédagogiques qui approfondissent l'objectif d'individualisation. Pour l'enseignement supérieur, Fred Keller propose une doctrine revisitant le conditionnement opérant, nommée « *Personalized System of Instruction* » (PSI). Celle-ci suppose une pédagogie différenciée où la motivation de l'élève est un objectif central, chacun progressant à son rythme au sein d'un programme (Keller, 1967).

L'impératif de validation d'une notion avant de passer à la suivante y est réactualisé. Idem dans la théorie du « *mastery learning* » proposée par Benjamin Bloom. Visant plutôt l'enseignement primaire et secondaire, celle-ci impose aux élèves d'atteindre un certain niveau de maîtrise (par exemple 90% à un test de connaissances) sur des connaissances préalables avant de passer à l'apprentissage d'informations plus poussées sur un sujet. Les tests réguliers permettent de déterminer cette maîtrise, et les élèves n'ayant pas atteint le niveau nécessaire reçoivent un soutien pédagogique supplémentaire avant un autre test (Bloom, 1968).

Ces théories imputent l'échec d'un élève à une instruction défectueuse plutôt qu'à une déficience personnelle. Plaçant l'individualisation de l'apprentissage au premier plan, elles inspireront de nombreux concepteurs de technologies éducatives ultérieures, notamment adaptatives, en contribuant à renouveler la flamme d'un behaviorisme déclinant.

Les sciences cognitives jouent cependant le rôle de langue conceptuelle commune aux courants scientifiques dominants, qui sont alors en passe de s'incarner dans l'essor de l'informatique pédagogique. Le concept de modèle, « forme abstraite qui vient à s'incarner ou à se réaliser dans les phénomènes » selon la définition donnée par Jean-Pierre Dupuy (Dupuy, 1994, p.18), irrigue puissamment les développements informatiques. L'architecture canonique des « tuteurs intelligents » consiste en trois modélisations, que John Self, chercheur britannique en informatique, fait correspondre à trois interrogations : « *What* » (« Quoi »), pour le domaine à enseigner, « *Who* » (« Qui »), pour la cognition de l'élève, et « *How* » (« Comment »), pour l'enseignant et ses stratégies pédagogiques (Self, 1974). De nombreux systèmes de tutorat sont développés aux

États-Unis et en France à partir de cette norme, en poursuivant des usages pluriels en termes d'individualisation.

Le modèle pédagogique

Concernant les stratégies pédagogiques, le chantier des stratégies pédagogiques donne à voir la progression de questionnaires de type « *drill and practice* » à des interactions variées sur le mode du tutorat. Comme le rappellent Christian Depover et Jean-Jacques Quintin, citant Viviane Glikman (2011), le tutorat est un référentiel anglo-saxon ancien, apparu au XIII^{e} siècle dans les écoles et au XVIII^{e} dans les universités. « Dans les deux cas, il est employé dans le sens d'enseignant-conseiller, chargé du suivi d'un élève ou d'un petit groupe d'élèves » (Depover et Quintin, 2011). Les deux auteurs rapportent que ce n'est qu'à la fin des années 1990 que le terme prend en français cette signification éducative, appliquée à l'enseignement supérieur. Les modalités du suivi tutoral sont plurielles et peuvent être distinguées selon l'antagonisme entre behaviorisme et constructivisme, comme le remarquent les deux auteurs (2011) :

> « Ainsi, dans un dispositif centré sur un mode de transmission des savoirs s'inspirant des idées behavioristes, le tutorat prend la forme d'interventions directives orientées vers le contrôle strict et la régulation des apprentissages. Par contre, les modalités de tutorat mises en œuvre sont généralement plus souples et moins invasives lorsque le dispositif s'inscrit dans une perspective constructiviste » (Depover et Quintin, 2011).

Les « tuteurs intelligents » qui adviennent outre-Atlantique dès les années 1970 oscillent entre ces deux visions. La veine behavioriste transparaît dans les systèmes affichant des suites de problèmes ou de questions, et envoyant de simples rétroactions à l'apprenant selon la validité de ses réponses. Pour des domaines où les savoirs consistent dans la restitution de faits (comme l'histoire ou la géographie), ce que les didacticiens appellent « des connaissances déclaratives », des agents conversationnels permettent des retours ponctuels. En témoigne *Scholar,* développé par Jaime Carbonell, chercheur en informatique à l'université Carnegie Mellon en 1970. Généralement considéré comme le premier « tuteur intelligent », ce système permet à l'usager de converser au sujet de la géographie de l'Amérique du Sud (Figure 3). Particulièrement innovant pour l'époque, le système n'en reste pas moins rigide et directif en termes de retours pédagogiques, corrigeant ou validant les réponses de l'élève.



Figure 3. Aperçu d'un dialogue entre un élève et le programme Scholar. *Source : Carbonell, 1970*

Dans une veine constructiviste, les possibilités d'interaction offertes par l'ordinateur permettent de varier les interactions tutorales, en laissant davantage d'autonomie à l'apprenant. Le contexte d'apprentissage étant central dans le constructivisme, certains chercheurs investiguent des espaces de découverte nouveaux, notamment ludiques. Selon Ira Goldstein et Brian Carr du MIT (Golstein et Carr, 1977), le jeu électronique apparaît dans les années 1970 comme un format d'apprentissage informel intéressant, à condition que l'apprenant soit guidé dans son expérience d'apprentissage. Les deux auteurs théorisent alors le modèle du « *computer coach* »[8] (« moniteur informatique discret »). Il en va par exemple de *West* de Burton et Brown (1982, Figure 4). Jeu de l'oie inspiré de la conquête de l'Ouest, il met au défi l'apprenant-joueur d'arriver en premier dans une ville, en tirant à chaque tour trois chiffres qu'il doit combiner dans une expression arithmétique (addition, multiplication, division). *West* génère la liste des déplacements jugés meilleurs que celui opéré par l'apprenant et décide, en fonction de principes pédagogiques prédéfinis, d'intervenir ou non pour lui conseiller l'un d'eux. Ces principes – guider de manière discrète, ou ne pas intervenir deux fois de suite – suivent un agenda constructiviste, devant permettre à l'apprenant de construire des connaissances à partir de ses erreurs.

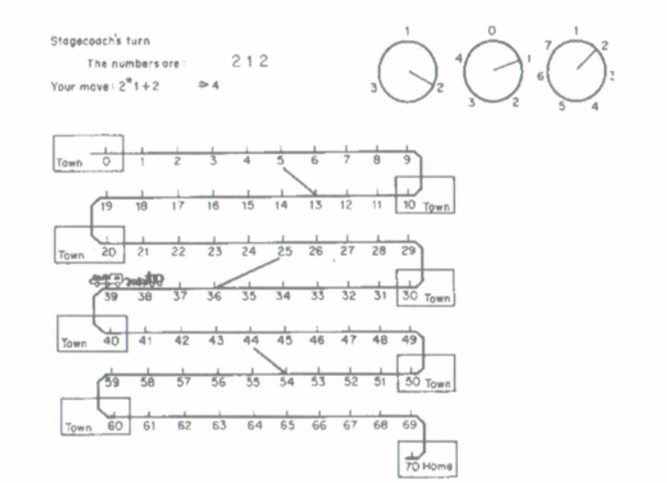


Figure 4. Capture d'écran de West. *Source : Solomon, 1986 ; répliqué par Bruillard, 1997, p.79*

Suivant la trilogie d'Edsurge, les premières formes d'*adaptive content* émergent ainsi, à travers l'affichage de remédiations modulables à l'intérieur de problèmes, à la suite d'erreurs commises par des élèves (Edsurge, 2016). Ces sélections automatiques, qu'elles soient behavioristes ou constructivistes, sont opérées en fonction des deux autres modèles qui conditionnent en amont l'adaptativité : ceux de l'élève et du domaine.

Le modèle du domaine

Concernant le domaine dans lequel le système doit assister l'élève, la contribution de l'IA symbolique à la fin des années 1970 conduit les chercheurs à coder des règles « expertes », à partir desquelles des moteurs d'inférence produisent des connaissances de manière autonome. Dans de nombreux secteurs

[8] Bruillard note à raison que « le terme anglais « *coach* » désigne à la fois un entraîneur sportif et quelqu'un qui donne des leçons particulières pour aider un élève à préparer un examen » (1997, p.78). On verra par la suite que ce double-sens sera réinvesti par les acteurs industriels de l'*adaptive learning.*

industriels, les recherches autour des systèmes experts se multiplient, comme le relatent Cardon, Cointet et Mazières :

> « *MYCIN*, la première réalisation d'une base de connaissances de 600 règles destinées à diagnostiquer les maladies infectieuses du sang sera le point de départ, dans les années 1980, du développement d'une ingénierie des connaissances qui s'applique essentiellement à des contextes scientifiques et industriels : *XCON* (1980) aide les clients des ordinateurs de DEC à les configurer, *DELTA* (1984) identifie les pannes des locomotives, *PROSPECTOR* repère des gisements géologiques, etc » (Cardon et al. 2018, p.192).

En éducation, les recherches approfondissent le sillon creusé par les programmes génératifs, visant à donner à la machine la capacité de « comprendre » ce qu'elle enseigne en résolvant elle-même des problèmes, pour ensuite pouvoir expliquer sa démarche à l'élève. À l'université de Stanford, William Clancey réalise en 1983 le système tutoriel *GUIDON* à partir de *MYCIN,* pour enseigner aux étudiants en médecine le diagnostic des maladies infectieuses. Comme le relate Éric Bruillard, ce dispositif marque un tournant, car il montre que « des systèmes experts, avant tout conçus pour leurs capacités d'expertise, ne sont pas *a priori* de bons systèmes d'enseignement » (Bruillard, 1997, p.117). En cause, le manque de pédagogie des connaissances produites auprès de débutants, qui ont besoin de davantage de relations de causalité pour saisir du sens.

Loin d'être une impasse, ce système motive des recherches nouvelles dans les systèmes-experts, afin d'accroître leurs capacités didactiques. En France, ce chantier connaît une fortune importante. Au début des années 1980 un courant se forme sous l'influence du pionnier français de l'intelligence artificielle symbolique, Jacques Pitrat. Des chercheurs contribuent ainsi à approfondir les capacités de résolveurs au sein de systèmes à base de connaissances (parmi lesquels des anciens doctorants de Pitrat comme Martial Vivet, Monique Grandbastien, Monique Baron, ou Dominique Pastre).

La pertinence de ces systèmes demeure cependant limitée en éducation, manquant une prise en compte suffisante des besoins de l'élève. Comme le reconnaît Nicolas Balacheff, « en fait, les projets IA de cette époque prennent l'enseignement comme un terrain d'application, plutôt que comme un domaine de recherche à part entière » (Balacheff, 1994). Une des rares initiatives françaises poursuivant une forme d'adaptativité pédagogique est celle d'Alain Bonnet, Dominique Kayser et Marie-Odile Cordier, de l'université d'Orsay, développant en 1981 un agent conversationnel couplé à un résolveur dans le domaine du calcul de dérivées. Sur le modèle de *Scholar*, cet agent sélectionne les types de dérivations à demander à l'apprenant et entretient un dialogue de type socratique (Bonnet et al, 1981 ; Bruillard, 1997, pp.115-116).

Le modèle de l'élève

Pour prendre en compte les besoins de l'élève dans leurs décisions pédagogiques, nombre de systèmes modélisent ses connaissances à partir de données numériques comme ses réponses correctes et incorrectes, ou le temps pris pour répondre. Le principal chantier de recherche consiste alors à trouver une représentation informatique qui serait fidèle à l'état des connaissances de l'apprenant. Deux voies sont principalement explorées. Une première est le modèle « d'expertise partielle » ou « de superposition » (« *overlay* ») dans lequel la connaissance de l'élève est un sous-ensemble de la connaissance cible. Il suffit alors de situer les connaissances de l'apprenant vis-à-vis de celles du « système expert », en termes de maîtrise plus ou moins grande, comme dans *West*. Une seconde voie est le modèle « différentiel », qui incorpore des « connaissances fausses » que peuvent avoir les apprenants. Ce genre de modèle appelle alors des stratégies de remédiation pour corriger ces

conceptions erronées « typiques », comme dans le système *Buggy* de John Seely Brown et Richard Burton (1978).

Les recherches en psychométrie se développent en vue d'évaluer plus finement les élèves. Il s'agit non plus de leur poser des questions identiques pour tous comme dans des tests standardisés, mais des questions différentes selon leurs réponses, devant permettre de jauger plus finement leur maîtrise. Se développent le courant de recherche « *computer adaptive testing »* (CAT) et, avec lui, les modèles statistiques affiliés à l'« *Item-Response Theory* » (IRT). Cette théorie vise à modéliser mathématiquement la probabilité de répondre correctement à un « item » à partir des « construits latents » de l'individu, comme sa maîtrise d'un concept. Le modèle le plus simple et répandu dans les années 1970 est celui de Rasch (du nom du mathématicien danois Georg Rasch). Celui-ci considère deux variables : le niveau de l'apprenant et la difficulté de l'item. Dès lors, si un élève répond correctement à une question, la suivante sera plus difficile, et à l'inverse, si sa réponse est incorrecte, une question plus facile suivra.

Ces techniques creusent le sillon de l'*adaptive assessment* dans une direction non pas « autonome » comme le faisait le programme de Smallwood, mais par des algorithmes opérant des calculs de probabilité formulés en amont par les chercheurs. La finalité initiale du CAT est l'évaluation plutôt que l'enseignement, avec l'ambition de jauger plus finement le niveau de l'élève par l'adaptation de la difficulté des questions. Néanmoins des applications pédagogiques apparaissent rapidement pour les « tuteurs intelligents », et en particulier la sélection automatique de questions dans les exercices d'entraînement.

Des obstacles poussant à une convergence avec les micromondes

Il apparaît rapidement que de tels systèmes ne parviennent que très imparfaitement à prendre en compte simultanément les trois modèles informatiques et à répondre aux problématiques de l'EAO présentées plus haut. En 1982, Derek Sleeman et John Seely Brown font état de quatre limites principales touchant directement à l'idéal d'adaptativité (Sleeman et Brown, 1982) :

- Le manque de « granularité » des contenus de remédiation, appréhendant à un trop faible niveau de détail les besoins de l'apprenant.
- L'obligation qu'a l'apprenant de travailler selon la conceptualisation du domaine, sans pouvoir exprimer ses propres conceptions.
- Le manque de théories de l'apprentissage pour étayer les stratégies d'enseignement, reposant surtout sur les intuitions des chercheurs.
- Les restrictions nombreuses pesant sur les interactions entre l'usager et l'interface, empêchant de cerner tout un éventail de facultés ou de besoins de l'élève.

Ces problèmes mettent en lumière des continuités entre les « tuteurs intelligents » et l'enseignement programmé, même si des ruptures ont eu lieu concernant les fondements théoriques et les utilisations pédagogiques. Le découpage du domaine est toujours opéré de façon verticale, et la place laissée à l'agentivité de l'apprenant est pauvre. S'y ajoute le manque de transfert des logiciels expérimentaux vers des marchés éducatifs, plus marqué que pour les machines à enseigner. Rares sont les produits utilisés dans des contextes d'apprentissage à grande échelle, pouvant donner lieu à des retours d'expérience significatifs.

En outre les « tuteurs intelligents » souffrent de la comparaison avec des technologies éducatives concurrentes, ayant émergé depuis l'arrivée des micro-ordinateurs dans les années 1960 : les « micromondes ». Robert Lawler caractérise ceux-ci comme « des mondes virtuels pour l'action créative, dont les objets ont des propriétés communes à la fois avec les objets formels de la science et les objets plus concrets de l'expérience sensible » (Bruillard, 1997, p.141). S'opposant frontalement à l'enseignement programmé, les pionniers Ted Nelson et Seymour Papert remettent en cause l'idée même de programme découpé en unités, ne permettant selon eux que la mémorisation et la formalisation par écrit chez l'apprenant. Au contraire, les tenants des micromondes soutiennent la prise de contrôle des élèves sur des environnements de programmation où ils exprimeraient leurs idées et mèneraient eux-mêmes des expérimentations. Ce courant profite des possibilités d'interaction humain-machine apportées par le développement des interfaces graphiques dans les années 1970 et 1980.

Cette perspective, marquée par la pensée de John Dewey et antinomique de toute conception verticale (et donc tutorale) de l'adaptativité, connaît une grande fortune chez les partisans des pédagogies actives. Le langage de programmation *Logo*, développé par Seymour Papert et Wallace Feurzeig à partir du *Lisp* – le langage de prédilection pour les recherches en intelligence artificielle – connaît une popularité internationale. C'est le cas en France à partir des années 1970, où des pédagogues spécialistes des technologies promeuvent leur déploiement, à l'image de Jacques Perriault.

Leur commercialisation est en outre plus poussée que celle des « tuteurs intelligents ». En 1989, Seymour Papert passe notamment un accord avec la firme Lego pour la conception et la commercialisation d'un jouet éducatif tiré de *Logo*, intitulé *Robolab*. Il ouvre la voie à nombre de « micromondes » : *Smalltalk, Sketchpad, Thinglab* aux États-Unis, ou *Cabri-Géomètre* en France. Si ces produits sont critiqués pour leurs prix jugés excessifs et leur manque de pédagogie, la popularité des micromondes au sein de la communauté scientifique va déteindre sur la recherche en technologie éducative.

Nombre de chercheurs s'unissent dès lors sous l'étendard des « environnements d'apprentissage », plus conformes à l'agenda constructiviste. Les « tuteurs intelligents » n'en deviennent qu'une déclinaison parmi d'autres, remise en cause par les partisans de modes éducatifs moins coercitifs et plus interactifs.

3. Les « environnements informatiques pour l'apprentissage humain » (EIAH), des perspectives nouvelles d'adaptativité malgré des contraintes substantielles

Depuis les recherches sur les machines à enseigner au XX^{e} siècle, les positions des chercheurs divergent sur le statut de la connaissance : est-elle un savoir détenu par des experts qu'il faudrait transmettre, ou un phénomène émergent que doit construire l'apprenant, seul ou en groupe, avec l'aide éventuelle d'une autorité extérieure ? Un collectif d'auteurs mené par Marlene Jones oppose dans un article de 1992 deux conceptions du chercheur et de son rapport à la connaissance : celle de l'ingénieur manipulant la connaissance comme un objet, et celle de l'anthropologue analysant les manifestations des processus cognitifs sous-jacents (Jones et al., 1992). En découlent selon eux des modèles éducatifs variés : transmission, négociation, exploration, construction guidée... Aux deux extrêmes du nuancier, il y aurait d'une part la « transmission », avec la métaphore ancienne

du récipient à remplir – critiquée dès le XVI^{e} siècle par Montaigne –, et d'autre part l'« engagement » de l'apprenant, « artisan » de son propre apprentissage (Bruillard, 1997, p.220).

Dans les années 1990, à mesure que s'accroissent les capacités d'interactivité offertes par les interfaces graphiques et qu'émergent de nouveaux formats avec le web, la quête constructiviste de l'engagement de l'apprenant trouve des avatars techniques variés. De nouvelles dénominations pour les produits de la recherche traduisent la focalisation sur l'agentivité de l'apprenant. Certains chercheurs états-uniens parlent dès lors d'« *intelligent computer-assisted instruction »* (ICAI), traduit en français par « enseignement intelligemment assisté par ordinateur » (EIAO), puis rebaptisé, en gardant le même sigle, en « environnement interactif d'apprentissage avec ordinateur » par Monique Baron, Régis Gras, Jean-François Nicaud (Baron et al., 1991). Se répandent également outre-Atlantique les étiquettes « *interactive learning environment »* (« environnement d'apprentissage interactif ») ou « *situated learning environment* » (traduit en « environnement d'apprentissage situé »).

Une expression populaire en français que nous privilégions par la suite est « environnement informatique pour l'apprentissage humain » (EIAH). Cette dénomination donne à voir l'influence nouvelle des micromondes par l'usage du terme « environnement » – métaphore d'un dispositif non-contraignant où l'apprenant doit pouvoir naviguer librement –, la préférence donnée à l'« apprentissage » plutôt qu'à l'enseignement, et la volonté de prendre de la distance vis-à-vis de l'intelligence artificielle, en qualifiant l'apprentissage d'« humain » par opposition implicite à l'apprentissage automatique.

Marie Lefèvre, informaticienne au CNRS, range alors les « tuteurs intelligents » parmi les EIAH, aux côtés d'outils ayant émergé dans les années 1990 et 2000 : les environnements interactifs, les environnements de simulation, les hypermédias pour l'apprentissage, la réalité virtuelle ou augmentée, les « *serious games* », les environnements d'apprentissage collaboratif, ou encore les plateformes de formation à distance (Lefèvre, 2021). Les « tuteurs intelligents » subsistent donc en tant que sous-domaine ancien des EIAH, cristallisant à ce titre des oppositions entre chercheurs.

Une opposition intestine entre anthropocentrisme et technocentrisme

Dans son ouvrage *Les machines à enseigner* de 1997, Éric Bruillard constate que, si le constructivisme est alors dominant chez les chercheurs, les positions varient sur la pertinence de concevoir des « tuteurs intelligents » :

> « Si presque tous sont favorables au développement des simulations, certains semblent abandonner l'idée même de tuteur intelligent, du fait de la difficulté de leur conception (Elshout) ou parce qu'ils considèrent qu'elle est basée sur des principes erronés (Shank). D'autres, tels Anderson, défendent un point de vue opposé. Pour eux, les fondements sont corrects, il est important de développer des tuteurs intelligents. La difficulté est réelle, mais le développement de bons outils devrait rendre leur conception possible et intéressante. Van Lehn, quant à lui, pense que le développement des tuteurs intelligents est souvent trop coûteux » (Bruillard, 1997, p.219).

Se dessine ainsi une opposition dans la recherche sur les EIAH. D'un côté, des chercheurs radicalisent les principes constructivistes pour investiguer les formats interactifs et collaboratifs nouveaux qu'offrent le web. D'un autre côté, les héritiers des « tuteurs intelligents », tout en étant redevables aux sciences cognitives, reconnaissent également l'héritage de l'enseignement programmé et tentent

d'approfondir des modes pédagogiques verticaux et individualisés. Cet antagonisme est formulé par Marie Lefèvre comme une opposition de point de vue entre anthropocentrisme et technocentrisme.

L'anthropocentrisme serait représenté par un point de vue d'abord porté sur les publics éducatifs (élèves, formateurs), en vue de les soutenir dans leurs usages des outils informatiques. Importe un apprentissage « collectif », les apprenants, tuteurs et enseignants étant amenés à interagir de façon synchrone ou asynchrone. Ce courant se fonde sur un affermissement du constructivisme et l'émergence de variantes théoriques mettant l'accent sur les dimensions sociales de l'apprentissage. Ainsi le socio-constructivisme avance l'idée de construction des fonctions mentales dans l'interdépendance du psychisme et de l'environnement – idée empruntée à Lev Vygotsky, psychologue soviétique des années 1920 et 1930. Son œuvre a été redécouverte aux États-Unis dans les années 1970, et en France dans les années 1980[9]. Selon lui, la pensée se forme deux fois : d'abord au niveau social ou « interpsychologique », dans l'interaction entre personnes, puis au niveau de la cognition individuelle, « intrapsychologique ». Les « instruments psychologiques », comme les mots et les concepts, assurent la transition entre ces deux niveaux. Vygotsky a notamment théorisé ce qui a été traduit en anglais comme « *zone of proximal development* » et en français « zone proximale de développement » (ZPD), selon laquelle une tâche pédagogique est considérée comme optimale lorsque l'élève peut la réussir en recevant une aide.

Une seconde théorie fondamentale pour penser les stratégies de guidage est le cognitivisme de Jerome Bruner. Celui qui avait pris part à la diffusion du constructivisme aux États-Unis s'intéresse ainsi aux stratégies mobilisées par l'apprenant pour comprendre la réalité. Dans ce cadre, l'enseignant assume « une interaction de tutelle » appelée aussi « étayage ».

Ces théories fondent des chantiers de recherche portés notamment sur l'autorégulation de l'apprentissage, ou la capacité d'une personne à comprendre et contrôler son environnement d'apprentissage. Selon le *Teaching Excellence in Adult Literacy Center,* celle-ci inclut trois composantes : les « stratégies cognitives » mises en place par l'apprenant pour apprendre un domaine, ses « capacités métacognitives », portant sur ses stratégies cognitives, et enfin sa « motivation », constituée notamment par la confiance dans ses propres compétences pour accomplir une tâche (TEAL Center, 2010). L'autorégulation doit émerger d'un apprentissage guidé par le *feedback* de formateurs (Paris et Paris, 2001). Ces chantiers réhabilitent la contribution des pairs et de l'enseignant, loin de la conception tutorale de l'apprentissage.

Selon une distinction opérée par le chercheur en sciences de l'éducation Sylvain Connac (2021), un objectif est dorénavant la « personnalisation », l'aménagement d'espaces de travail collectif visant à « lier individuation (conscientisation de soi) à socialisation (développement de l'importance de l'autre) » (Connac, 2021)[10].

Pour les chercheurs de ce courant, l'apprentissage individualisé face à un système de tutorat aurait perdu de son intérêt, car il ne considèrerait pas suffisamment les besoins de l'apprenant. Certains en viennent également à réfuter la pertinence de représenter informatiquement l'état de connaissances

[9] C'est à ces époques qu'a été traduit son ouvrage phare paru en 1934, intitulé en français *Pensée et Langage*.

[10] Sylvain Connac reprend la distinction entre individu et personne opérée par les philosophes personnalistes (Emmanuel Mounier, Martin Buber) des années 1920 à 1950. Elle se fonde sur le constat philosophique selon lequel les êtres humains ne sont pas que des individus, au sens d'êtres insécables, mais également des êtres en relation. Étymologiquement, le *persona*, en latin (en grec, πρόσωπον - prosôpon) est un masque de théâtre antique, qui permet d'endosser un rôle, de s'adresser aux spectateurs en tant que personnage (Mauss, 1970). Une personne est à ce titre un individu « augmenté » d'une identité sociale, qu'il endosse dans ses relations aux autres.

d'un apprenant. Sharon Derry et Lois Hawkes (1993) remarquent que les enseignants ne conduisent pas de diagnostics détaillés avant d'intervenir pour aider les apprenants. La simulation de l'activité enseignante, fondamentale dans les tentatives d'incorporation de l'intelligence artificielle aux « tuteurs intelligents », est délaissée au profit des possibilité d'interactivité offertes par les nouveaux types d'EIAH. L'adaptativité y apparaît, au mieux, comme un objectif secondaire, et, au pire, comme une fonctionnalité aliénante pour les publics éducatifs.

Du côté du technocentrisme mis en exergue par Marie Lefèvre figurent les fidèles des « tuteurs intelligents », pour qui les algorithmes doivent en premier lieu être perfectionnés pour simuler efficacement le tutorat, et *in fine*, soutenir l'élève en face à face. Ce courant priorise davantage l'adaptativité dans ses recherches. Pour ces chercheurs, l'article de Benjamin Bloom de 1984, montrant la supériorité du tutorat humain sur toute autre forme de pédagogie, est une preuve supplémentaire de la nécessité de perfectionner les systèmes de tutorat existants. L'auteur de la théorie du « *mastery learning* » concluait d'une expérimentation à l'université de Chicago que les étudiants travaillant avec un tuteur humain obtenaient des notes supérieures de deux écarts-types à celles d'étudiants suivant un enseignement conventionnel (Bloom, 1984). Autrement dit, les étudiants bénéficiant d'un tutorat individuel obtenaient des résultats aussi élevés que les 2 % des meilleurs étudiants suivant des cours collectifs et simultanés.

Bloom mettait néanmoins le doigt sur une question de recherche qu'il nomma le « problème à deux sigmas » : comment concevoir des environnements éducatifs aussi efficaces que des tuteurs humains individuels, mais suffisamment abordables pour être diffusés à grande échelle ? Question fondamentale qui inspirera nombre de développements scientifiques et industriels ultérieurs en matière de « tuteurs intelligents ».

Dans ce sillage se constituent, dans les années 1980, deux communautés scientifiques « sœurs », se réunissant dans des conférences internationales et publiant leurs propres journaux, sous les syntagmes « *Artificial Intelligence in Education* » (AIED) et « *Intelligent Tutoring Systems* » (ITS). En France, Nicolas Balacheff contribue à la création du groupe « Intelligence artificielle et Didactique » (Balacheff, 1994).

Souvent passé sous silence en raison de ses échecs passés, l'héritage behavioriste reste pourtant vivace, offrant des outils conceptuels pour répondre à ce problème. Fidèles à l'agenda constructiviste initial visant à appréhender la cognition de l'apprenant, les chercheurs y adjoignent des bases de psychologie comportementale pour programmer des systèmes adaptatifs censément efficaces. Parmi eux, John Anderson, chercheur à l'université Carnegie Mellon, est l'auteur de la théorie de l'« *Active Control of Thought – Rational »* (ACT–R) dans les années 1980. Celle-ci considère l'apprentissage comme l'accumulation de « règles de production » – du type « Si…alors… » – permettant à l'apprenant de transformer des connaissances déclaratives en connaissances procédurales (Anderson, 1987). Aux règles de stimulus-réponses fondant la connaissance skinnerienne, Anderson substitue les règles de production, moins nombreuses et plus facilement computables (Bruillard, 1997, p.184). C'est à partir de ce principe de conception que lui et ses collègues (Albert Corbett, Kenneth Koedinger, Steven Ritter) développent durant les années 1980 et 1990 des « tuteurs cognitifs » (« *cognitive tutors* »), voulus plus proches de la cognition de l'apprenant que les « tuteurs intelligents », comme on le verra plus bas. Cette dénomination est révélatrice d'une tentative paradoxale de marier les héritages constructivistes et behavioristes.

Les développements du *machine learning* ouvrent également des possibilités nouvelles aux partisans de ces recherches. Les recherches en IA connaissent en effet un *aggiornammento* au tournant des années 1980 et 1990. Les systèmes experts n'ont pas tenu leurs promesses, à cause d'une complexité extrême et de performances limitées. En parallèle, comme le rapportent Cardon, Cointet et Mazières en citant James Fleck, « la diminution du coût et l'augmentation des capacités de calcul dans les années 1980 ouvrent l'accès à des calculateurs puissants aux courants hétérodoxes et déviants qui avaient été mis à l'écart par le monopole du courant symbolique sur le financement des grands projets informatiques » (Cardon et al., 2016, p. 196). En découle l'avènement de l'IA dite « connexionniste », qui postule que « penser s'apparente à un calcul massivement parallèle de fonctions élémentaires [...] dont les comportements signifiants n'apparaissent au niveau collectif que comme un effet émergent des interactions produites par ces opérations élémentaires » (Cardon et al., 2018, p.178). Contribuant au retour sur le devant de la scène d'une focale comportementale, les dispositifs d'analyse de données, de plus en plus nombreuses, préfigurent des capacités nouvelles en termes d'adaptativité.

De nouvelles avancées en matière d'adaptativité

Les développements techniques autour de l'adaptativité portent les empreintes de ce paysage scientifique segmenté. Au sein des « tuteurs intelligents », elle reste un objectif premier tout en étant revisitée, voire augmentée, par les avancées issues de chantiers constructivistes comme l'interaction humain-machine ou l'aide à la décision.

L'interface est en effet de plus en plus prise en compte par les chercheurs en tant que médium de l'apprentissage. Elle connaît des développements importants dans les années 1990, à travers notamment la popularisation du modèle WIMP (« Windows, Icons, Menus, Pointers », soit « Fenêtres, icônes, menus, curseurs ») dans les systèmes d'exploitation grand public, et l'arrivée de la navigation hypertexte sur le web. Ces innovations permettent de diversifier les formats des contenus éducatifs.

Un aperçu des avancées sur ce chantier est offert par le système de tutorat en mathématiques *Cognitive Tutor Algebra I.* Celui-ci est lancé en 1998 par un collectif de chercheurs de Carnegie Mellon emmené par John Anderson[11]. L'écran ci-dessous (Figure 5) donne à voir la diversité des formats rencontrés par l'apprenant. L'interface est divisée en plusieurs espaces : « *scenario* » affiche la consigne d'un exercice d'arithmétique. Deux fenêtres proposent des manières variées d'y répondre : la « feuille de travail » (« *worksheet »)* invite l'élève à poser les termes du problème dans un tableau à double entrée ; le « graphe » (« *grapher »)* l'incite à représenter le problème sous forme graphique. Ces deux fonctionnalités intégrées à l'exercice donnent à voir l'influence des micromondes sur les « tuteurs intelligents ».

[11] Ces chercheurs le commercialisent en créant leur entreprise, Carnegie Learning. Nous reviendrons sur cette « spin-off » universitaire au cours de notre thèse.

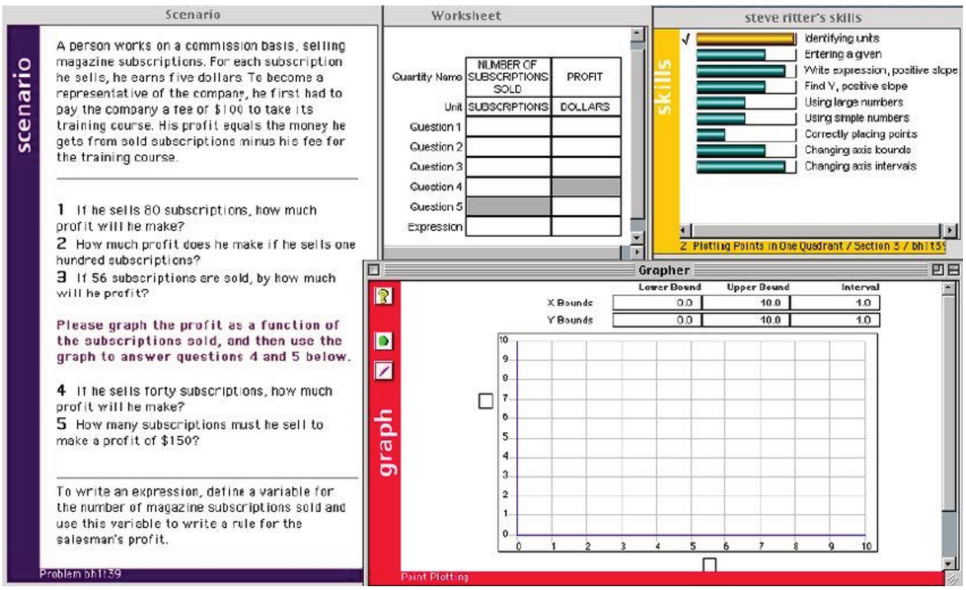


Figure 5. Capture d'écran du Cognitive Tutor Algebra I. *Source : Ritter et al., 2004.*

Ces avancées en termes d'interactivité étendent les possibilités de collecte de données. Pour prendre en compte la diversité des apprenants, émerge l'objectif d'interfaces *adaptables* selon les souhaits de l'utilisateur, voire *adaptatifs* – dans un sens propre au design et non pédagogique –, c'est-à-dire capables de découvrir automatiquement les préférences d'utilisation de celui-ci. Aux côtés de modèles de l'apprenant, les chercheurs programment des « modèles de l'utilisateur ». Ceux-ci sont nourris par des « traces », qu'Hugues Labarthe et Vanda Luengo définissent comme « [les] interactions d'un usager avec son environnement d'apprentissage » (Labarthe et Luengo, 2016). Des fichiers (« logs ») contiennent ainsi l'historique des actions d'un utilisateur sur une page, comme les clics ou la durée de consultation d'une vidéo.

Un autre domaine d'innovations qui distingue les EIAH de l'époque est l'aide automatique à la décision, auprès des apprenants comme des professionnels éducatifs (enseignants, administrateurs). Au centre de ces considérations se situe le *monitoring,* défini par Éric Bruillard comme « une procédure d'observation, de recueil, de traitement et de mise en forme d'informations en provenance de l'environnement d'apprentissage, sur les caractéristiques et les comportements des différents éléments constitutifs, utiles aux différents niveaux de régulation du système évaluatif » (Bruillard, 1997, p.185). Ce chantier de recherche se concrétise par des tableaux de bord (« *dashboards* »). Via des techniques d'intelligence artificielle dites d'« *infovizualisation »,* il s'agit de rendre intelligible des calculs statistiques complexes réalisés à partir des données des élèves, pour par exemple permettre à l'enseignant de juger leurs progressions par rapport aux projections du modèle. On retrouve là un chantier émergent de ce qui, aujourd'hui, constitue une norme dans les environnements numériques éducatifs.

Sur l'interface de *Cognitive Tutor* (Figure 5), l'icône figurant un point d'interrogation sur la gauche du graphique permet à l'apprenant de solliciter de l'aide, sous la forme d'un indice adapté à son niveau de maîtrise. Ce dernier est affiché dans une dernière fenêtre intitulée « *skillometer* », fonction de *monitoring* destinée à l'autorégulation de l'élève. Elle lui permet de voir sa maîtrise des compétences impliquées dans la résolution du problème, les « *knowledge components* » – par exemple « identifier des unités » ou « placer correctement des points » –, via la métaphore de jauges plus ou moins remplies.

Sur le plan de l'adaptativité pédagogique, ce système de tutorat témoigne également de traitements algorithmiques plus dynamiques. Parmi les méthodes de modélisation de l'élève ayant émergé dans les années 1990[12], John Anderson et son collègue Albert Corbett proposent le « *knowledge tracing »* en 1994*,* prolongeant la théorie ACT–R. À chaque fois que l'élève répond à un exercice lié à une compétence *(*« *knowledge components »)*, le système observe sa performance (réponse correcte ou incorrecte) pour mettre à jour sa croyance sur le fait que l'élève a ou n'a pas encore maîtrisé la compétence. Cette croyance consiste dans le calcul d'une probabilité de succès sur d'autres exercices liés à cette compétence. Il est opéré par un modèle bayésien programmé par les chercheurs, intégrant des paramètres indépendants comme la probabilité que l'apprenant maîtrise la compétence suite à une tentative, qu'il devine sans maîtriser la compétence, ou qu'il commette une erreur tout en maîtrisant la compétence. *In fine*, si l'élève réussit, la probabilité de succès augmente, et s'il échoue, elle diminue. Cette méthode représente une avancée significative dans le diagnostic des capacités de l'élève par rapport à l'*Item Response Theory* et le modèle de Rasch, qui n'offraient que des diagnostics ponctuels.

Lorsque toutes les compétences d'une section du programme sont jugées suffisamment maîtrisées, l'étudiant passe à la section suivante du programme, qui introduit de nouvelles compétences. Si celle-ci excède les 95%, la notion est considérée comme maîtrisée (Corbett et Anderson, 1995). C'est à la fois une mise en application de la théorie du « *mastery learning* » de Bloom et du concept de zone proximale de développement de Vygotsky : le système cherche à afficher des problèmes que l'élève ne maîtrise pas encore, mais peut résoudre avec une aide adéquate. Un syncrétisme est ainsi opéré entre le behaviorisme et le socioconstructivisme, l'un apportant sa focale verticale et comportementale, et l'autre ses outils conceptuels.

Cependant, comme le relatent Piech et al. (2015), le *knowledge tracing* souffre de plusieurs limites. La méthode ne peut d'une part être appliquée qu'à des domaines comme l'arithmétique où une représentation binaire de la compréhension des élèves (acquis/non-acquis) est valable. Deuxièmement, la signification des *knowledge components* et leurs correspondances dans les exercices peuvent être ambiguës et répondent rarement à la contrainte fonctionnelle d'un seul concept par exercice. Enfin, les paramètres du modèle bayésien demeurent statiques, découlant d'expérimentations menées lors de la phase de développement sur des cohortes limitées d'élèves. « Non apprenant », le modèle ne peut donc modifier ces variables en fonction des usages faits au jour le jour par les utilisateurs. Risque d'en découler, avec le temps, des écarts entre des paramètres comme les probabilités de deviner ou de se tromper, et les pratiques réelles des usagers.

L'arrivée des hypermédias adaptatifs dans la formation à distance

Parmi les nombreuses perspectives ouvertes par les micromondes (comme la réalité virtuelle ou augmentée), les hypertextes sont des dispositifs informatisés permettant de lier des documents de natures diverses, sur la base d'associations que l'utilisateur peut consulter librement. Le terme, inventé par Ted Nelson en 1965, est parfois complété par celui d'*hypermédia* pour y ajouter des documents audiovisuels (images, vidéos, sons). Les hypertextes apportent aux environnements

[12] Ohlsson (1994) propose une autre méthode de diagnostic, le « *Constraint-based modeling* ». Elle postule que l'apprentissage se fait par la confrontation de l'apprenant avec ses erreurs, en se basant sur la théorie « *performance errors »* (Ohlsson, 1996). Les « contraintes », ou principes fondamentaux d'un problème, sont programmées pour permettre de diagnostiquer des erreurs significatives chez un élève et lui fournir des retours spécifiques pour les corriger. En France, Nicolas Balacheff (1995) propose le modèle didactique cKc devant prendre en compte les conceptions de départ de l'élève dans le domaine de la démonstration mathématique (Labarthe et Luengo, 2016, p.15).

informatiques de nouvelles approches dans la manière de structurer les domaines, en prenant en compte la capacité des apprenants à naviguer et contrôler leurs parcours. Le web, passé dans le domaine public en 1993, va se structurer sur ce type d'architecture, permettant progressivement aux étudiants du monde entier de se former à distance au gré de « liens hypertextes ».

De voies nouvelles d'adaptativité sont alors investiguées. En 1996 est distribué le premier EIAH sur le web, *ELM-ART*, par Peter Brusilovsky, Elmar Schwarz, et Gerhard Weber. Leur ambition est de proposer un manuel « adaptatif » (« *adaptive textbook* ») pour l'apprentissage du langage Lisp (Brusilovsky et al., 1996). Cela s'est rapidement avéré être une gageure, en raison de deux principaux obstacles : l'ordre séquentiel de la rédaction des manuels, rendant leurs contenus difficilement réagençables dans le désordre, et l'impossibilité de modéliser l'état de connaissance d'un utilisateur à partir de la simple lecture d'une page (Weber et Brusilovsky, 2001, p.353).

Les deux chercheurs et, à leur suite, les adeptes des hypermédias se sont plutôt tournés vers l'industrie de l'enseignement à distance. Les usages permis par le web représentent des opportunités nouvelles en termes d'adaptativité selon Weber et Brusilovsky (2001) :

> « Premièrement, la plupart des applications éducatives basées sur le web sont utilisées par un plus grand nombre d'utilisateurs que n'importe quelle application autonome. […] Deuxièmement, les étudiants travaillent généralement seuls (souvent chez eux) avec des systèmes éducatifs basés sur le web et ne peuvent pas bénéficier de la même assistance intelligente et personnalisée que celle qu'un enseignant ou un de leurs pairs peut fournir dans une classe normale »[13] (Weber et Brusilovsky, 2001, p.351).

Brusilovsky et Weber cherchent ainsi à formaliser un modèle éducatif à distance et personnalisé à l'heure du web. Leur ambition est de concilier l'adaptativité avec la polyvalence des dispositifs à distance, notamment en matière de communication entre utilisateurs, dans ce qu'ils nomment des « hypermédias adaptatifs » (« *adaptive hypermedias* », Brusilovsky, 1996). Ils créent l'année suivante une seconde version, *ELM-ART II*, comprenant des exercices. Ces derniers permettent de mieux modéliser la maîtrise de l'élève et d'adapter sa navigation dans les contenus, lui ouvrant l'accès ou non à certains liens via la métaphore d'un feu tricolore. De façon originale et conforme aux thèses « anthropocentrées », le modèle de l'élève lui est directement accessible, pour qu'il puisse juger de ses progrès et besoins. Cela ouvre la voie de la « modélisation ouverte de l'apprenant », ou « *open student modeling* ».

Ces chantiers « anthropocentrés » et distanciels, connexes à l'adaptativité, trouvent un écho significatif en France. Les plateformes de formation à distance sont investiguées pour permettre l'interaction et la collaboration entre élèves et professeurs dans des formations. En 2002 le projet PIXED (« *Project Integrating eXperience in Distance Learning* ») issu de la thèse de Jean-Mathias Héraud, consiste ainsi à développer un hypermédia adaptatif sur le web, avec pour objectif la réutilisation des logs d'utilisateurs pour apporter une aide aux utilisateurs suivants (Héraud, 2002).

L'arrivée du « web social » et les défis de l'industrialisation

La diffusion massive du « web social » et la démocratisation des ordinateurs personnels dans les années 2000 galvanisent les recherches sur l'apprentissage en ligne et collaboratif, en particulier

[13] *« First, most Web-based educational applications are used by a much wider variety of users than any standalone application. A Web application designed with a particular class of users in mind may not suit other users. Second, students usually work with Web-based educational systems on their own (often from home) and can't get an intelligent and personalized assistance that a teacher or a peer student can provide in a normal classroom. »*

dans la formation à distance. Citant l'essayiste Tim O'Reilly, Fabrice Pirolli et Raphaëlle Crétin-Pirolli définissent le « web social » comme « un ensemble d'applications et d'outils dont les finalités peuvent être variées, mais qui ont en tous en commun les notions de partage et de collaboration », et mentionnent comme exemples les réseaux sociaux, les blogs, ou les wikis (Pirolli, et Crétin-Pirolli, 2011). Ce type de dispositifs a fait émerger de nouveaux outils d'enseignement asynchrones, tels que les « campus virtuels », les plateformes de formation à distance, et, à la fin des années 2000, les portails de « *Massive Open Online Course* » (MOOC).

À l'aube du *Big Data*, l'audience massive et mondialisée auxquels ces dispositifs commencent à donner accès ouvre des perspectives en termes d'adaptativité, en raison du volume croissant de données collectées. Hugues Labarthe et Vanda Luengo distinguent deux sortes de données récoltées par les informaticiens (2016, p.7) :

- Les « données qualitatives » correspondent aux réponses d'un usager à des formulaires, comme ses tentatives lors d'exercices, dont est inférée la progression et la quantité d'effort de l'apprenant, ou ses posts et messages, dont sont déduits sa participation ou son engagement dans l'environnement. Ce sont des données « épaisses », ou « *thick* », sur lesquelles travaillent les chercheurs pour saisir le sens d'un phénomène, en lien avec des théories des sciences humaines.
- Les « traces » regroupent quant à elles « l'ensemble des interactions d'un usager avec son environnement d'apprentissage ». Collectées depuis les débuts des EIAH, elles rassemblent des données d'usage dans des *logs*, comme les zones de clics, la durée de consultation ou les scores obtenus. Les traces sont censées renseigner la maîtrise, l'assiduité ou les préférences d'usage de l'apprenant.

Face à l'amoncellement de traces de natures et de formats variés, les méthodes de fouille et d'analyse sémantique basées sur le *machine learning* s'imposent dans le champ académique des EIAH. Elles font émerger, à la fin des années 2000, des modèles de l'apprenant basés sur des régressions logistiques. Il s'agit d'une méthode relativement basique d'apprentissage-machine, dite « supervisée », car entraînée à réaliser une tâche à partir d'exemples étiquetés en entrée et en sortie. Dans la lignée de l'*Item-Response Theory*, des régressions telles que l'*Additive Factor Model* (AFM, Cen et al., 2008) ou la *Performance Factor Analysis* (PFA, Pavlik et al., 2009) apprennent à opérer des prédictions de succès d'un apprenant sur une tâche donnée à partir de ses données d'apprentissage et des unités du domaine. Selon Labarthe et Luengo, l'on passe ainsi « d'une démarche centrée expert vers une démarche centrée données » (Labarthe et Luengo, 2016, p.15). Le chercheur en intelligence artificielle Alain Mille confirme que « la connexion avec l'activité est maintenant généralisée », et que « ce sont les traces d'activité d'apprentissage qui sont alors collectées pour alimenter les modèles servant à guider la personnalisation, l'adaptation » (Mille, 2013).

Pourtant, les plateformes à distance commercialisées sont loin d'être aussi adaptatives que les projets expérimentaux des chercheurs (George, 2001). Plusieurs facteurs expliquent les difficiles transferts de technologies des laboratoires vers les marchés éducatifs. Le premier obstacle est la faiblesse sémantique des traces issues du web par rapport à celles collectées en conditions expérimentales. Par exemple, les recherches d'obédience constructivistes menées à la fin des années 1990 tendent à diversifier les observations de l'apprentissage vers la prise en compte de facteurs émotionnels. Il s'agit notamment d'évaluer la concentration, les émotions et les capacités cognitives d'un apprenant via l'analyse de ses expressions faciales ou de mesures oculométriques (Gluck, 1999). Il apparaît *a priori* difficile d'industrialiser de tels dispositifs de collecte.

Un autre obstacle est la variabilité des contextes d'apprentissage et la multiplicité des facteurs contribuant à l'apprentissage, remettant en cause le transfert de résultats obtenus d'une classe à l'autre. Les facteurs hors ligne (psychologiques, biographiques), non saisis par les données, demeurent une limitation structurelle. Ces angles morts apparaissent d'autant plus dommageables pour l'apprentissage en ligne qu'ils font surgir un risque élevé d'erreurs dans l'adaptation des exercices. Face à ces limitations, la communauté scientifique s'organise, notamment pour partager des bibliothèques de données, où sont renseignés les contextes de collecte, à l'image du *Pittsburgh Science of Learning Center DataShop* en 2004.

Qui plus est, les médias éducatifs constructivistes souffrent, à l'image des micromondes avant eux, d'une difficulté à concilier assistance et autonomie de l'apprenant, ainsi que d'un manque de consensus sur un projet pédagogique dans lequel inclure ces outils. Les taux de décrochage rencontrés sur les plateformes de MOOC durant les années 2010 actualisent ce constat, affectant leur ambition de démocratisation de l'éducation dans le monde entier (Compagnon, 2014).

À ces obstacles s'ajoute la difficulté des chercheurs à prouver l'efficacité pédagogique de leurs dispositifs. Pourtant l'« *Evidence-Based Education »* devient une attente sociale et politique très forte au début des années 2000. Théorisé par le pédagogue anglais David Hargreaves en 1996, et traduit en français par « éducation fondée sur des preuves », ce principe suppose que les pratiques éducatives soient basées sur les meilleures preuves scientifiques disponibles – découlant le plus souvent d'essais randomisés – plutôt que sur la tradition ou le jugement personnel de l'enseignant. Il a vocation à guider les politiques éducatives vers davantage d'efficience, et tend à être puissamment implémenté aux États-Unis, puis en France à partir des années 2000. Or, sur ce type de protocoles, les outils constructivistes ne montrent aucun progrès significatif des élèves (Mœglin, 2010, p.50).

À l'aube des années 2010, l'idéal d'adaptativité se trouve donc confronté à des défis de taille alors qu'adviennent le *Big Data* et les méthodes de *data mining*. Revient un débat ancien pour le secteur : l'adaptativité peut-elle concerner l'enseignement entier de connaissances déclaratives et procédurales, ou doit-elle s'en tenir à des aspects limités et optionnels de la pédagogie (conseils, recommandation d'exercices), en laissant l'essentiel de l'activité entre les mains des publics éducatifs ?

Si, à la fin des années 2000, on se retourne sur l'histoire de cet idéal ingénierique, on mesure le chemin parcouru depuis les premières machines à enseigner. Alors que l'ambition initiale était de conditionner l'affichage d'une question à la validation de celle précédente (Pressey), les améliorations des « programmes » jusqu'aux années 1960 ont permis de « brancher » l'élève à des contenus différents selon ses réponses (Crowder), ou d'adapter les tâches demandées selon un historique des réponses passées (Pask). L'arrivée du micro-ordinateur et ses perspectives interactives ont fait apparaître les limites de l'enseignement programmé et émerger les « tuteurs intelligents ». Profitant des avancées de l'IA symbolique pour endosser une focale davantage centrée sur l'apprentissage, ils ont adapté la pédagogie aux « états cognitifs » modélisés de l'apprenant. Leur complexité technique et leurs angles morts ont fait apparaître les limites de l'enseignement transmissif, par opposition à la liberté de découverte laissée par les micromondes. À partir des années 1980, une opposition est alors apparue au sein du constructivisme dominant entre les défenseurs d'« environnements » interactifs conçus du point de vue des publics, et les fidèles du mode tutoral et transmissif, tentant une synthèse délicate entre constructivisme et behaviorisme. Parmi leurs avancées, les modélisations dynamiques du niveau de l'apprenant sont apparues, permettant d'adapter

l'attribution d'exercices selon la progression de l'apprenant. L'accroissement des données collectées avec les hypertextes et le « web social » a ouvert de nouveaux champs de recherche en provenance de la formation à distance, bien que les données restent encore simplistes d'un point de vue sémantique.

Si nombre de chercheurs en informatique éprouvent du scepticisme quant aux capacités tutorielles des machines, l'avènement du *Big Data* au tournant des années 2000 et 2010 va apporter des réponses sur le problème du volume. Au risque d'appauvrir en contrepartie la valeur sémantique des données, ce bouleversement peut-il relancer, sous des formes nouvelles, l'idéal de l'adaptativité dans les EIAH ? L'arrivée du « *deep learning »,* ou « apprentissage profond », dans les techniques d'apprentissage automatique entérine un nouveau paradigme scientifique inspiré par l'IA connexionniste : l'« épistémologie de la donnée ». Une question est de savoir si celui-ci est de nature à redistribuer les cartes et faire émerger l'adaptativité comme une caractéristique centrale des médias éducatifs contemporains.

Chapitre II. Les apports du *Big Data* et de l'IA connexionniste aux recherches sur l'adaptativité

L'accroissement des volumes de données dans les années 2000 a mis sur le devant de la scène scientifique des techniques de fouille ou « *data mining* », comme on a commencé à le voir. Avec l'affirmation de l'*Evidence-Based Education* dans les politiques éducatives, ces techniques sont investies d'une promesse : révéler des phénomènes naguère invisibles pour les observateurs et les théories des sciences humaines. Émerge un paradigme scientifique nouveau, que Bruno Bachimont qualifie d'« épistémologie de la donnée », autour de laquelle se constituent deux nouvelles communautés, l'*Educational Data Mining* (EDM) en 2007 et les *Learning Analytics* (LA) en 2011. Ces recherches poursuivent l'espoir que les machines assument des tâches essentielles dans l'enseignement et dépassent enfin les limitations des dispositifs informatiques antérieurs.

Parallèlement, une nouvelle théorie pédagogique est formulée, le connectionnisme, qui adapte le socio-constructivisme aux sites et plateformes du « web social ». Elle vise à optimiser la circulation des connaissances au sein de réseaux, dans lesquels les enseignants seraient des facilitateurs assistés par la technologie. La dichotomie entre le connectionnisme et l'épistémologie de la donnée revisite l'opposition entre technocentrisme et anthropocentrisme. Soumises à cette tension, les communautés des LA et de l'EDM spécifient leurs méthodologies et projettent des usages différents de l'adaptativité. Avant de s'intéresser aux avatars techniques et aux voies éducatives qu'ils portent, il convient d'abord de caractériser ces deux conceptions épistémologiques.

1. Une tension épistémologique nouvelle

Une épistémologie de la donnée au service de l'éducation

L'explosion des données numériques collectées sur le web provoque un bouleversement scientifique au tournant des années 2000 et 2010. L'anglicisme qui en rend compte, « *Big Data* », est défini par Hugues Labarthe et Vanda Luengo comme « un jeu de données massives, dont la taille excède celle de capture, de conservation, de gestion et d'analyse des outils logiciels de bases de données typiques » (Labarthe et Luengo, 2016, p.6). À ce critère du volume, une définition de référence fournie par un rapport de Gartner en 2001 ajoute deux autres notions : « *Variability* », soit la variété des formats et des sources dont proviennent ces informations (du web, de bases publiques…) et « *Velocity* », pour la vitesse à laquelle les données sont engendrées, collectées, partagées et mises à jour[14].

L'explosion quantitative des traces touche notamment les EIAH distribués sur le web, comme les plateformes de MOOC. Labarthe et Luengo décrivent les implications méthodologiques de ce basculement :

> « En l'espace de 2-3 années, environnements, problématiques et méthodes ont changé : on est passé de données *a priori* destinées à évaluer certains aspects émotionnels, comportementaux ou cognitifs, à des données tout venant mais massives à la fois par le nombre des utilisateurs, l'intensité des usages d'un utilisateur particulier et

[14] D'autres définitions ont ajouté des « V » supplémentaires comme « *Veracity* », « *Value* », « *Volability* ».

> l'entrée du monde réel dans le monde numérique via la géolocalisation et les indicateurs biométriques » (ibid., p.7).

Les données collectées sur le web ont ainsi perdu en épaisseur ce qu'elles ont gagné en volume. En découle un basculement scientifique consacrant les méthodes inductives de la « science des données » (« *data science* »). Si le syntagme est inventé en 1996 par l'informaticien japonais Chikio Hayashi, ces méthodes se sont développées à partir des années 1960, à la croisée des statistiques et de l'informatique. Le programme d'enseignement de la géométrie conçu par Richard Smallwood en 1962, modifiant ses embranchements selon les résultats des élèves, en était une forme précurseure. Ces méthodes ont surtout innervé le champ scientifique à travers la biologie et la médecine, dans la délivrance de diagnostics automatiques par exemple. Elles ont commencé à se diffuser dans la sphère économique à partir des années 1990, rencontrant un franc succès dans le marketing et le management – ce dont témoigne le courant de la « *Business Intelligence* ».

Avec le *Big Data* se répand dans l'ensemble du champ académique une épistémologie inspirée de cette discipline, selon laquelle la vérité résiderait non pas dans les théories mais dans les données. D'une méthode hypothético-déductive, consistant à mettre à l'épreuve, par un protocole d'enquête, des hypothèses formulées en amont dans un cadre théorique, on passerait à l'induction, processus qui permet de passer du constat particulier offert par les données à une règle générale.

Selon Bruno Bachimont, qui préfère à *Big Data* le terme de « mégadonnées », « l'enjeu et la révolution attendue des mégadonnées sont […] bien la compréhension nouvelle promise des activités humaines » (Bachimont, 2015, p.395). L'avènement de ce qu'il nomme l'« épistémologie de la donnée » pourrait ainsi constituer « une révolution des sciences de la culture à l'instar de la révolution scientifique qui a permis, à l'orée de notre modernité, de transformer notre relation à la nature d'une description fondée sur le langage à un rapport de mesure expérimentale et de formalisation calculée » (ibid.). De façon analogue, « [a]u lieu d'interpréter le fait culturel par la compréhension qu'on en a à travers l'expression qu'il reçoit dans le langage, il s'agit désormais de calculer des corrélations entre les données rassemblées » (ibid., p.397).

Cette promesse scientiste reviendrait donc à enfin « objectiver » l'éducation, au moyen des corrélations révélées par les traces numériques. Dans *Objectivité* (2012), les historiens des sciences Lorraine Daston et Peter Galison ont écrit à cet égard une historiographie de l'idéal d'objectivité dans les sciences. Dans la préface de l'édition française, Bruno Latour résumait les trois régimes principaux de ce qu'il nommait « *sub-objectivité* » – relation entre subjectivité et objectivité dans la construction des savoirs –, ayant présidé à la démarche scientifique moderne selon les deux auteurs :

- le premier est le temps des « pionniers » qui luttaient contre le chaos du réel et insistaient sur des types observables « d'après nature », en suivant leur subjectivité propre (XVIII^{e} siècle).
- Le second est le temps des « travailleurs de la preuve », qui au XIX^{e} siècle ont fait de l'objectivité un idéal scientifique et honni toute subjectivité, poursuivant le caractère mécanique de l'enregistrement.
- Le troisième est celui de la réhabilitation de l'acte subjectif de l'interprétation, devant la complexité croissante des instruments de mesure du réel.

La démarche scientifique guidant les recherches en fouille de données éducatives s'apparente à un renouvellement du second modèle. Une analogie est possible entre l'histoire technique de l'objectivité

scientifique et celle de l'individualisation pédagogique : de la même manière que la photographie fut utilisée en sciences par les « travailleurs de la preuve » du XIX^{e} siècle pour dépasser la méthode ancienne du dessin, les algorithmes sont programmés pour dépasser les démarches d'individualisation reposant sur la seule intuition du professeur. Cette épistémologie débouche sur la justification d'une technologisation substitutive, l'algorithme étant censé effectuer plus efficacement des diagnostics ou des prescriptions pédagogiques que l'humain.

Cette démarche marque un nouvel élan pour le behaviorisme, qui refuse d'ouvrir la boîte noire du sens « construit » par l'individu pour se focaliser sur ses comportements. La donnée numérique doit à la fois représenter la cognition individuelle et la formater pour les algorithmes. La résurgence de ce courant se nourrit des difficultés des médias éducatifs constructivistes évoquées plus haut, comme l'absence de consensus sur le degré d'assistance et d'autonomie de l'apprenant, le défaut de projet pédagogique dans lequel inclure ces dispositifs, ou la difficulté à prouver leur efficacité pédagogique. Les médias simplifiant l'acte éducatif sur la base des traces et se donnant pour objectifs des taux de maîtrise et d'autres indicateurs quantitatifs représentent un horizon souhaitable pour les pouvoirs publics adeptes de l'*Evidence-Based Education*.

Face à ce retour de flamme behavioriste, le socio-constructivisme trouve un nouvel élan à travers une théorie de l'apprentissage adaptée au « web social », le connectivisme, qui investit les données massives de visées pédagogiques et politiques différentes.

Le connectivisme, résurgence du socio-constructivisme

Dans un article de 2005, le chercheur canadien George Siemens expose les grands principes de ce courant. Celui-ci doit répondre aux prétendus manques des « trois grandes théories pédagogiques », à savoir le behaviorisme, le cognitivisme et le constructivisme, devant la généralisation des technologies numériques (Siemens, 2005). Cette nouvelle donne nécessiterait selon le chercheur une théorie n'appréhendant plus l'apprentissage sous un prisme individuel, mais comme un phénomène émergent au sein de réseaux, dont les individus et les organisations constitueraient les nœuds.

Cette théorie postule une dépréciation du savoir, voué à une obsolescence accélérée : « La connaissance croît de façon exponentielle. Dans de nombreux domaines, la vie de la connaissance est désormais mesurée en mois et en années » (ibid., p.1)[15]. Les individus et les organisations seraient alors contraints d'être « flexibles » et « ouverts » aux échanges. Siemens formule cette idée à travers un aphorisme rappelant ceux de Marshall McLuhan : « Le canal est plus important que le contenu. Notre capacité à apprendre ce dont nous avons besoin pour demain est plus important que ce que nous savons aujourd'hui » (ibid., p.7)[16]. Cette théorie se présente ainsi comme un outil pour le management des organisations, et notamment des établissements éducatifs : ceux-ci doivent considérer la connaissance comme une entité collective et en mouvement, pour faciliter sa circulation parmi les individus.

Dans ce cadre, Siemens rappelle que l'enseignant doit jouer un rôle crucial pour orienter les « expériences » éducatives, réactualisant les thèses de Dewey. Les données massives sont alors

[15] *« Knowledge is growing exponentially. In many fields the life of knowledge is now measured in months and years ».*
[16] *« The pipe is more important than the content within the pipe. Our ability to learn what we need for tomorrow is more important than what we know today ».*

interprétées comme des instruments au service des « expériences ». Un autre penseur du courant, John Goldie, explique en 2016 que face à l'autonomie accrue des apprenants, le rôle des enseignants est désormais celui de « facilitateurs », chargés d'évaluer « non seulement les gains d'apprentissage individuels mais aussi la qualité des processus d'apprentissage. Par exemple, les manières dont les apprenants se connectent les uns aux autres pour collaborer et partager la création de connaissances, ainsi que la production et la consommation de ressources d'apprentissage par les étudiants ». Or « la quantité de données générées par les réseaux en ligne offre des possibilités d'évaluation » (Goldie, 2016)[17].

Le connectivisme promeut donc une technologisation additive à rebours de celle substitutive supposée par l'épistémologie de la donnée. L'adaptativité y serait secondaire, la focale collective étant prioritaire, et les analyses algorithmiques devant surtout permettre à l'enseignant d'évaluer et de faciliter la circulation de la connaissance.

La tension entre l'épistémologie de la donnée, aux accents behavioristes, et le connectivisme, succédané du socio-constructivisme, nourrit la distinction entre l'*Educational Data Mining* et les *Learning Analytics*. Les représentants de l'EDM tendent à reprendre le plus fidèlement les préceptes du premier courant en poursuivant le perfectionnement des algorithmes, pour adapter au mieux et de façon plus étendue la pédagogie des EIAH. Les LA mettent plutôt l'accent sur la transmission des modélisations auprès des acteurs de l'apprentissage (apprenants, enseignants, administrateurs d'institutions éducatives, parents) pour les aider dans leurs pratiques respectives. Ils appliquent la *data science* comme un instrument parmi d'autres au service d'une vision collaborative et connectiviste de l'apprentissage en ligne.

En conséquence de cette tension entre technocentrisme et anthropocentrisme, ces deux communautés se positionnent différemment sur les méthodologies et les interdisciplinarités à mettre en place pour appréhender l'apprentissage. Présenter les chantiers propres à chaque courant peut apparaître comme une gageure, car les chercheurs de l'EDM et des LA mobilisent fréquemment des méthodologies communes. Tel est le cas de la prédiction de l'apprentissage : un algorithme prédictif pourra instrumenter un système de tutorat autonome, ou assister un enseignant en quête d'informations sur ses élèves pour différencier ses cours. Nous restituons néanmoins les chantiers préférentiels de ces deux communautés en nous focalisant sur leurs différences en termes d'intentions de recherche.

2. L'*Educational Data Mining,* ou la poursuite étendue de l'adaptativité

Perfectionner les traitements algorithmiques pour dépasser les limites des enseignants

L'*Educational Data Mining* (EDM) s'est institutionnalisé en tant que discipline avec la création, en 2007, du réseau scientifique *International Educational Data Mining Society.* Ryan Baker et Kalina Yacef, des représentants de l'EDM, définissent ce courant comme « un domaine de recherche impliquant : 1. le développement de méthodes dédiées à l'exploration des spécificités des données d'apprentissage ; 2. le déploiement de ces méthodes pour une meilleure compréhension des

[17] *« Teachers have a pivotal role in the process although as students become increasingly autonomous their role would be more facilitative. In the connectivist approach learning would be assessed not only in terms of individual knowledge gain, but also through evaluation of the learning process. For example, the ways learners connect with others to collaborate and share the creation of knowledge and also the production as well as consumption of learning resources by students. The amount of data generated by the online networks provides opportunities for such evaluation ».*

apprenants et de l'environnement d'apprentissage » (Baker et Yacef, 2009, p.4, traduits par Labarthe et Luengo, 2016, p.8). Les représentants de l'EDM poursuivent le perfectionnement des algorithmes, pour adapter plus précisément la pédagogie d'un EIAH.

Ce courant de recherche applique l'épistémologie de la donnée de façon radicale, privilégiant la quête de corrélations statistiques et la méthode inductive à partir de traces massives. L'affirmation selon laquelle la machine dépasserait l'enseignant dans sa capacité à percevoir des facteurs d'apprentissage est un principe fondamental. Dès l'introduction de leur rapport sur l'« analytique des données d'apprentissage », Hugues Labarthe et Vanda Luengo y font référence :

> « Évaluer les capacités d'abstraction des apprenants, détecter leur perte d'attention, adopter une pédagogie différenciée, dresser un bilan personnalisé actualisé au fil de l'apprentissage : voici autant de tâches qui reposent sur la capacité d'un enseignant à observer, analyser et réinvestir les traces comportementales et cognitives d'un apprentissage. Bien malgré lui, ce professionnel ne capte qu'une partie infime de ces données ce qui limite ses possibilités d'interprétation d'un geste, d'un exercice inabouti, d'une erreur de réappropriation. Comme dans d'autres domaines, l'observation humaine non instrumentée est limitée et fragile » (ibid., p.4).

L'enseignant est perçu comme un agent capable d'erreurs de diagnostics, ce qui justifierait la substitution de la machine. En cela, l'EDM est marqué par la culture éducative états-unienne, rappelée plus haut par Pierre Mœglin : « chez les Anglo-saxons et sous l'influence du taylorisme, l'*Educational Technology* n'attribue à l'enseignant aucun rôle prééminent. Ce sont l'organisation et les organisateurs qui constituent la cheville ouvrière du dispositif » (Mœglin, 2005, pp.146-147). Selon l'agenda de l'EDM, les algorithmes doivent désormais intégrer le cœur de l'organisation éducative. Ainsi, par la place marginale donnée à l'enseignant et la masse des données nécessaire aux méthodes de *data mining*, la configuration pédagogique projetée est essentiellement une formation à distance sur un environnement informatique.

La prédiction de l'apprentissage

Un exemple d'application de ces principes est la prédiction de l'apprentissage, un domaine de recherche privilégié par l'EDM. Il consiste à affiner les modèles de l'apprenant hérités de décennies de recherche grâce aux méthodes de fouille de données. Labarthe et Luengo expliquent que « ces analyses sont menées à différents niveaux de granularité, de la simple interaction en temps réel d'un individu (microgenèse) aux apprentissages d'une cohorte sur une période donnée » (Labarthe et Luengo, 2016, p.14).

La prédiction vise deux focales différentes :

- Au niveau macroscopique, les institutions éducatives sont en quête de modèles pour prévenir l'échec et améliorer leurs pratiques. Témoin les travaux de Jared Knowles visant à perfectionner le système de prédiction du risque de décrochage des collégiens du Wisconsin (Knowles, 2015).
- Au niveau individuel, les chercheurs tentent d'inventer de nouveaux modèles pour prédire au plus près la progression de l'apprenant, en captant la moindre interaction lors d'une situation d'apprentissage. Une frappe, un clic, un coup d'œil seraient potentiellement des « signaux faibles » de l'apprentissage. L'influence du behaviorisme est patente, les données qualitatives et les traces étant mobilisées selon une approche comportementale. Le but ultime est l'adaptativité au sens d'une individualisation, c'est-à-dire d'une focalisation de la pédagogie sur les performances de l'individu, selon Sylvain Connac (2021).

Ce chantier capitalise sur les algorithmes prédictifs développés depuis des décennies, itérés pour être adaptés aux nouveaux environnements d'apprentissage et à leurs enjeux. Nous avons évoqué les versions de l'*Item-Response Theory* ayant émergé dans les années 2000, basées sur des régressions logistiques à partir des données de l'apprenant, comme AFM ou PFA. En 2011, une autre méthode a commencé à être appliquée pour dépasser les diagnostics statiques de l'IRT : le système de notation « Elo » (Klinkenberg et al, 2011 ; Wauters et al. 2011). Celui-ci fut développé dans les années 1960 par le scientifique du même nom pour évaluer le niveau des joueurs d'échecs et associer des adversaires lors des parties (Elo, 1978). Dans ce système, si à l'issue d'un match, le favori sort victorieux, le classement de chacun des joueurs change peu. En revanche, si l'*outsider* l'emporte, son classement augmente et celui du favori diminue à mesure que l'écart initial de classement était important.

Les applications éducatives de ce système tendent à évaluer la maîtrise des élèves et la difficulté des questions pour les associer, comme deux adversaires aux échecs. L'algorithme assigne à chaque élève un classement évoluant dynamiquement selon ses performances à chaque question, et à chaque question un indice de difficulté évoluant selon les réponses des apprenants. Ce système vise *in fine* à attribuer à l'élève des questions ni trop simples ni trop compliquées, conformément à la théorie de la zone proximale de développement. Celle-ci se trouve quantifiée via un intervalle de probabilité de succès optimal. Grâce aux données amassées sur l'élève et les contenus, des algorithmes de *machine learning* simulent donc la probabilité de réussite de l'apprenant sur une banque d'exercices, et retiennent un contenu dont le taux est compris dans cet intervalle. Partageant avec l'IRT la mise en perspective de la maîtrise des apprenants et de la difficulté des questions, Elo s'en distingue par sa capacité à opérer des diagnostics dynamiques et relatifs à partir de données massives, alors qu'IRT calcule de façon ponctuelle des probabilités de réussite individuelle, à partir de variables nécessitant des données plus sophistiquées.

Une nouvelle avancée en termes de modélisation de l'apprenant survient au mitan des années 2010 avec l'application de « réseaux de neurones ». Ceux-ci sont basés sur le « *deep learning »,* ou « apprentissage profond ». Selon le site de la CNIL, le *deep learning* est « un procédé d'apprentissage automatique utilisant des réseaux de neurones possédant plusieurs couches de neurones cachées. Ces algorithmes possédant de très nombreux paramètres, ils demandent un nombre très important de données afin d'être entraînés » (Commission Nationale de l'Informatique et des Libertés, 2024). Ces données, qu'elles soient qualitatives ou quantitatives, sont préalablement transformées en vecteurs, pour permettre au réseau de mettre en évidence des relations et des structures signifiantes.

Comme le relate l'informaticien Yann Le Cun dans son ouvrage *Quand la machine apprend* (2019), les « réseaux de neurones » ont été utilisés pour la reconnaissance visuelle dès les années 1980. L'arrivée du *Big Data* au tournant des années 2000 et 2010 a conduit à un réemploi de cette architecture dans de nombreux secteurs à des fins prédictives, de l'énergie (prédiction des consommations d'électricité d'un quartier) à la publicité en ligne (prévision des taux de clics), en passant par la médecine (diagnostic de tumeurs) ou le transport (conception de voitures autonomes)[18].

[18] Leur avènement dans la communauté de l'IA remonterait selon Cardon, Cointet et Mazières (2016) au concours de computer-vision ECCV de 2012. Les chercheurs en IA du monde entier mettaient en compétition leurs modèles de reconnaissance d'images sur une bibliothèque commune. Contre toute attente, les méthodes d'apprentissage profond obtinrent des performances largement au-dessus de leurs homologues à vecteurs de support. La prise de conscience de la supériorité de ces machines inductives, ostracisées par la communauté de l'intelligence artificielle comme par celle de l'apprentissage automatique pendant des décennies, enclencha leur diffusion dans tous les domaines de recherche où interviennent des données.

En éducation, la nécessité d'un entraînement sur des bases de données très vastes conduit les chercheurs à conjuguer les gisements laissés ouverts par des initiatives universitaires (comme le *Datashop* du *Pittsburgh Science of Learning Center*) et des sites de *e-learning*. Le premier « réseau de neurones » employé en éducation, le *deep knowledge tracing* (DKT), est entraîné à Stanford par un collectif de chercheurs emmenés par Chris Piech (Piech et al., 2015), à partir des données de trois sites de MOOC : Khan Academy, Coursera, et EdX. Comme son nom l'indique, ce réseau reprend le principe du *knowledge tracing* d'Anderson et Corbett : comportant en entrée les données d'apprentissage de l'apprenant, les *knowledge components* (KC) mobilisés dans chaque exercice ainsi que la réponse attendue, le réseau de neurones est entraîné à prédire l'exactitude de la réponse de l'élève à partir des réponses passées de millions d'autres apprenants.

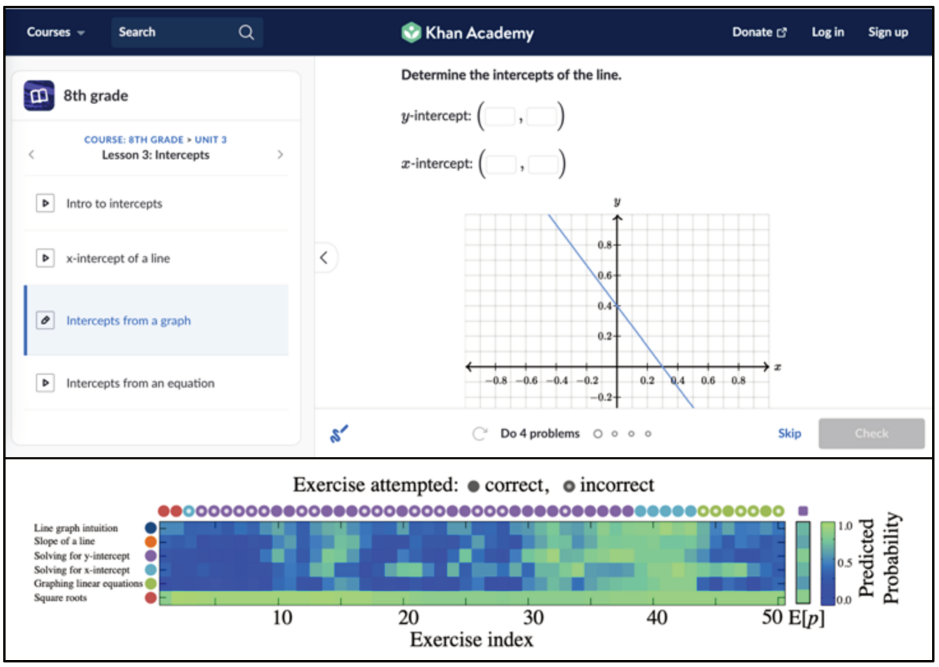


Figure 6. (haut) Capture d'écran d'un exercice de mathématiques de 8th grade sur Khan Academy *; (bas) Visualisation des probabilisations opérées par* DKT.
Source : Piech et al. 2015

La Figure 6. ci-dessus montre, du point de vue de l'utilisateur en haut, un exercice de mathématiques de « *8th grade* » (équivalent américain de la quatrième). La partie basse offre une visualisation des calculs de probabilités opérés par le réseau pour cet élève : sur six *knowledge components* (« *Line graph intuition* » ou « *Solving for y-intercepts* ») figurés par couleur, une séquence de cinquante exercices donne à voir ses réponses correctes et incorrectes dans chaque KC. Selon celles-ci, le modèle projette des probabilités de succès E[p] pour de nouveaux exercices mobilisant ces KC. Ces probabilités sont représentées par des tuiles de couleur, du bleu foncé, figurant une probabilité faible – par exemple quand l'élève a commis des erreurs récentes sur cette compétence – au vert foncé, représentant une probabilité forte de succès – quand l'élève a répondu correctement par le passé.

L'« état caché » du réseau, c'est-à-dire l'ensemble des couches de « neurones » intermédiaires capables d'ajuster leurs réglages, figure l'état des connaissances d'un étudiant. Gardant en mémoire ses résultats précédents, il doit « découvrir » les paramètres de sa progression au fil du temps. En fonction de ces facteurs et de critères pédagogiques, il doit permettre de sélectionner les exercices à afficher pour maximiser son apprentissage.

L'ensemble de ces calculs intermédiaires sont inaccessibles pour un observateur extérieur, ce qui rend d'autant plus brûlante la problématique de l'explicabilité des décisions pédagogiques dans les ITS, comme on le verra par la suite.

La fouille de données a donc intégré, sous des formes variées, les méthodes de l'*Educational Data Mining,* en vue de prédire l'apprentissage individuel. Dans ce type de travaux, la place laissée à l'interprétation humaine et aux protocoles hypothético-déductifs des sciences humaines est très maigre. Sur ce point, les *Learning Analytics* entretiennent leur spécificité.

3. Les *Learning Analytics,* ou les limites de la technologisation additive

Perfectionner les visualisations pour repousser les limites des enseignants

La seconde discipline, les *Learning Analytics* (LA), émerge en 2011 avec la création de la communauté *Society for Learning Analytics Research* (SoLAR). Selon Labarthe et Luengo, le terme « *analytics »* désigne des « techniques informatiques, mathématiques et statistiques pour révéler une information à partir de très larges ensembles de données. Par extension, les *analytics* permettent, sur la base d'actions réalisées, de comprendre, voire de prédire, le potentiel de futures actions, dans une quête de performance et d'efficience » (ibid., p.6). Par cette quête d'efficacité à partir des données, les LA reposent comme l'EDM sur l'épistémologie de la donnée.

Comme expliqué plus haut, une différence notable des LA vis-à-vis de l'EDM réside cependant dans l'accent mis sur la transmission de visualisations auprès des acteurs de l'apprentissage (apprenants, enseignants, administrateurs d'institutions éducatives, parents). L'objectif final est de les aider dans leurs pratiques respectives suivant une technologisation additive. La donnée numérique apparaît comme un instrument au service d'une vision connectiviste de l'apprentissage, devant notamment assister l'enseignant dans son rôle de « facilitateur ». Le théoricien du connectivisme George Siemens s'est ainsi intéressé aux *analytics* dans un article conjoint avec Phil Long, où il les définit comme « l'évaluation, l'analyse, la collecte et la communication des données relatives aux apprenants, leur contexte d'apprentissage, dans la perspective d'une compréhension et d'une optimisation de l'apprentissage et de son environnement » (Long et Siemens, 2011, traduits par Labarthe et Luengo, 2016, p.8).

Dans le cadre des LA, l'adaptativité n'est pas tant visée comme une fin en soi que comme un ingrédient de *personnalisation*, au sens donné par Sylvain Connac, d'aménagement d'espaces de travail collectif visant à « lier individuation (conscientisation de soi) à socialisation (développement de l'importance de l'autre) » (Connac, 2021). À la différence de l'EDM qui privilégie la formation à distance, les usages pédagogiques projetés par les LA sont divers : entre pairs et/ou avec l'enseignant, synchrone ou non, en classe, sur le web ou hybride.

Une place ambiguë réservée à l'enseignant

Les types d'analyse conduits par les LA sont cependant de natures différentes, participant d'une certaine ambiguïté sur la place laissée à l'intervention humaine. Le cabinet Gartner a tenté de synthétiser ces approches à travers quatre niveaux croissants de responsabilités laissées aux *analytics,*

et, ce faisant, de complexité (Figure 7.) : descriptif (« Que s'est-il produit ? »), diagnostique (Pourquoi cela s'est-il produit ?), prédictif (« Que va-t-il se produire ? ») et prescriptif (« Que devrais-je faire ? »).

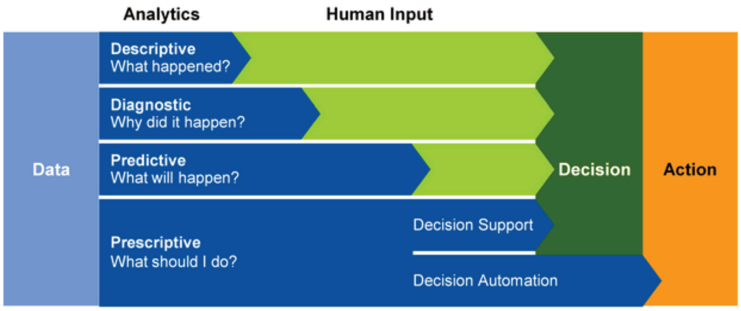


Figure 7. Les quatre approches des Learning Analytics *par Gartner. Source : Gartner, 2014.*

En matière de description et de diagnostic, les données collectées ne sont pas toujours massives, mais peuvent être issues de protocoles expérimentaux visant à capter des facteurs sociaux ou émotionnels dans l'apprentissage. Ces données plus « épaisses » peuvent être complétées par des observations qualitatives réalisées par des humains. En matière d'analyse, les résultats des fouilles de données et des méthodes statistiques sont régulièrement combinées avec des théories des sciences humaines, en vue de rendre les résultats intelligibles auprès des publics éducatifs.

Vanda Luengo, une des figures des *Learning Analytics* en France, expliquait en 2018 à Daniel Peraya que les méthodes déductives et inductives se complétaient souvent (Peraya et Luengo, 2018) :

> « Par exemple, dans le cadre d'analyse de l'apprentissage et dans une perspective prédictive, il est rare que la prédiction d'un certain type d'erreurs ne puisse être expliquée au regard d'une théorie. Si l'on fait du *clustering [partitionnement]*, on va distinguer des groupes et même si la démarche de distinction de ces groupes a été menée sans aucun *a priori* théorique, ce sont bien les théories qui vont permettre d'expliquer les distinctions entre ces groupes ».

Outre-Atlantique, ces méthodes croisées sont également fréquemment utilisées. Rencontrée en septembre 2023 à Pittsburgh, Carolyn Rosé, figure du « *Computer-supported collaborative learning* » et chercheuse à l'université Carnegie Mellon, nous a expliqué mêler des méthodes quantitatives et qualitatives pour appréhender le langage :

> « Dans le travail de modélisation que je fais sur les technologies linguistiques, c'est tout à fait ainsi, parce que je pense que le langage est quelque chose qui contourne souvent les règles. Les modèles discrétisent le monde et il y a donc quelque chose de fondamentalement faux dans ce que le modèle fait quand il s'agit du langage et de la façon dont nous l'utilisons. […] Je pense qu'il nous manque beaucoup d'aspects dans ce qui se produit en réalité, et c'est pourquoi je les mélange dans mon travail »[19] (Entretien 1).

Les théories et méthodes qualitatives sont cependant mobilisées dans un cadre expérimental, organisé autour d'essais randomisés, en vue de mesurer l'efficacité d'un dispositif par rapport à un autre.

[19] *« [In] the modelling work that I do in language technologies it's very much that way, because I think that language is something that is more often than not actually bending rules you know. Models discretize the world and so there's something fundamentally wrong about the nature of what a model does when it comes to something like language and how we actually use it. [...] I think that we're missing a lot of what's really there and that's why I blend those together in my work. »*

Exit des méthodes interprétatives voire critiques qui n'auraient pas de retombées pratiques immédiates. Les sciences cognitives et la linguistique sont ainsi mobilisées pour aider à diagnostiquer, dans les travaux écrits ou les interactions en ligne de l'individu, la construction de ses connaissances (Labarthe et Luengo, 2016, pp.18-19).

Les neurosciences fournissent par ailleurs des axes de recherche à la neurodidactique, interdiscipline liée aux sciences de l'éducation. Son objectif est d'améliorer la qualité de l'éducation en développant des stratégies didactiques plus efficaces. La neurodidactique cherche à saisir les processus neuronaux influant sur l'attention, l'effort, l'engagement, les émotions sociales et *in fine* les résultats d'apprentissage. Sont par exemple investigués des dispositifs visant à stimuler la production de la dopamine et d'endorphines – les deux neurotransmetteurs liés au plaisir – à des fins éducatives (Tang, 2017 ; Bonomo, 2017). L'adaptativité qui en découle porte notamment sur la recommandation d'activités ludiques censées augmenter la motivation de l'élève. Cette focalisation sur les *inputs* et les comportements désirés en sortie est de nature à relancer une vision behavioriste de l'apprentissage.

Selon le schéma de Gartner, sur des tâches prédictives, les *analytics* fournissent une aide à la décision. Des algorithmes d'*infovizualisation* sont chargés de rendre les descriptions et diagnostics automatiques facilement réutilisables pour l'enseignant. Celui-ci peut ainsi se voir conseiller, dans son tableau de bord, l'aménagement de sa classe en groupes de travail selon les besoins et les styles d'apprentissage.

En matière de prescriptions, les algorithmes peuvent soit présenter des arbitrages à l'enseignant, en lui laissant le soin de les appliquer ou non, soit automatiser des décisions pédagogiques de manière coercitive. Cette dernière éventualité se rapproche de l'EDM et d'une technologisation substitutive. Même dans le premier cas de figure où les prescriptions ne sont qu'additives, celles-ci restent soumises à une contradiction découlant de leur parenté avec l'EDM : devant conseiller les formateurs en leur offrant de nouvelles visions de l'apprentissage, elles reposent sur des techniques empruntées au *data mining* qui ne sont pas directement intelligibles. Prétendant ouvrir la boîte noire de la cognition de l'apprenant, elles enferment dans une autre les causes qui les poussent à conseiller telle décision aux enseignants et responsables de formations.

En définitive, les communautés de l'EDM et des LA actualisent les débats entre technocentrisme et anthropocentrisme. L'une applique les préceptes de la *data science* en laissant peu de place à l'interdisciplinarité, quand l'autre sollicite des méthodes exogènes compatibles avec une quête d'optimisation des pratiques pédagogiques. Ces positionnements ont des implications sur l'adaptativité visée. Les chercheurs en EDM approfondissent cet idéal en affinant les modélisations et les arbitrages pédagogiques des algorithmes. Pour les LA, le lien à l'adaptativité est plus ambigu, entre une assistance prioritaire des professionnels éducatifs, sans garantie d'intelligibilité, et une automatisation coercitive des décisions. Ces orientations ont des implications sur les usages pédagogiques projetés, entre un tutorat en « vase clos » et une collaboration synchrone ou non des publics éducatifs. Quelles qu'elles soient, ces intentions de recherche sont grevées par des limitations techniques et pédagogiques, qui hypothèquent l'impulsion d'un projet éducatif cohérent vers les organisations éducatives.

Chapitre III. Des limitations scientifiques en question

À l'aube du *Big Data*, l'idéal d'adaptativité se trouvait suspendu à un débat existentiel : l'adaptativité pouvait-elle concerner l'enseignement entier de connaissances déclaratives voire procédurales ? Ou devait-elle s'en tenir à des aspects limités et optionnels de la pédagogie (conseils, recommandation d'exercices), en laissant l'essentiel de l'activité entre les mains des professionnels éducatifs ? Avec l'avènement des méthodes de fouille de données massives, plusieurs barrières ont été levées. D'abord l'explosion quantitative des traces d'usage a permis d'appréhender davantage de facteurs d'apprentissage. L'avènement du *machine learning,* appliqué à des cohortes de milliers, voire de millions d'apprenants, a permis de dégager des régularités statistiques et des opportunités pédagogiques nouvelles en termes de parcours d'apprentissage ou d'interactions sociales. Toutefois certains problèmes demeurent entiers, voire plus urgents encore, comme la question du potentiel sémantique des données récoltées, la nécessité de contenus de remédiation nombreux et variés ou la difficulté des chercheurs à prouver l'efficacité pédagogique de leurs dispositifs. Ils confirment ensemble la gageure que représente la perspective d'une technologisation substitutive. D'autres obstacles sont en outre apparus avec la complexification des processus algorithmiques, comme le défi de l'explicabilité des décisions éducatives auprès des publics. Ces problèmes demeurent autant de barrières, sinon de menaces pour la viabilité des recherches sur l'adaptativité ainsi que leur réception sociale.

1. Le dilemme entre données massives et épaisses, ou l'illusion d'une technologisation substitutive

Un dilemme se pose pour les chercheurs concernant la qualité des données à collecter, entre deux idéaux divergents : le volume et l'épaisseur. Le premier revient à privilégier l'aspect quantitatif des données, et le second leur portée sémantique. Derrière leurs avantages spécifiques, les renoncements qu'ils supposent rappellent la gageure d'un recours à la technique en remplacement de l'enseignant.

La question du volume de données

Parmi les arguments en faveur du volume, l'analyse de jeux de données massives, telle que pratiquée dans la communauté de l'EDM, a diversifié les possibilités d'adaptativité dans les EIAH. Les capacités accrues des processeurs informatiques permettent aux modèles de *machine learning,* voire de *deep learning,* d'effectuer des prédictions plus précises et dynamiques sur la capacité des apprenants à maîtriser des contenus. En matière d'interventions pédagogiques, l'adaptation peut concerner simultanément les feedbacks en cas d'erreur, le contenu des exercices à réaliser et l'agencement des notions à étudier, sur un mode directif ou via des recommandations.

Cependant les données numériques, le « carburant » de ces infrastructures, gagnent en volume ce qu'elles perdent en valeur sémantique. Le risque consubstantiel de l'adaptativité, identifié dans plusieurs travaux critiques du *data mining* en sciences humaines, est de surestimer le caractère signifiant des données. Ainsi Alan Ouakrat et Julien Mésangeau soulignent-ils que les traces ne demeurent que des « traductions partielles et limitées des pratiques » (Ouakrat et Mésangeau, 2016). En éducation, une approche exclusivement « centrée données », qui ne reconnaîtrait pas

l'existence de facteurs hors ligne et psychologiques non-computables, conduirait à des diagnostics et des adaptations erronées, avec des conséquences néfastes pour l'apprentissage.

En outre, les capacités nouvelles de traitement informatique nécessitent des moyens techniques et financiers élevés. Le déploiement du *deep learning* repose sur des ordinateurs ayant plusieurs dizaines de milliers de processeurs – des millions pour les plus puissants, les supercalculateurs –, capables de réaliser des quantités astronomiques d'opérations par seconde, nécessaires à l'entraînement des réseaux de neurones. Ces équipements restent très onéreux et ne sont accessibles qu'aux universités suffisamment dotées. Si la France investit dans le développement de ses supercalculateurs, ceux-ci sont bien plus nombreux aux États-Unis, du fait de l'ampleur des subventions fédérales et notamment du sponsorat d'acteurs industriels. Selon l'étude Top500 de juin 2024, recensant les 500 supercalculateurs les plus puissants au monde, 171 seraient basés aux États-Unis contre 24 en France (Top500, 2024). Des recherches nombreuses tendent cependant à réduire le besoin en énergie et en mémoire des modèles de *deep learning*, pour pouvoir les entraîner sur des terminaux personnels (Schwartz et al., 2019).

L'enjeu de l'épaisseur des données

D'autres chercheurs privilégient pour leur part l'« épaisseur » des données récoltées. Pour dépasser leurs limitations sémantiques, des chercheurs, souvent issus de la communauté des LA et soucieux de mesurer des aspects sociaux ou émotionnels de l'apprentissage, en appellent à des données plus « épaisses ». En France, Vanda Luengo confie ainsi ses vœux à Daniel Peraya en 2019 :

> « Ce que j'appelle une donnée épaisse est une donnée qui décrit l'activité de l'apprenant de façon plus fine, comme on peut le faire dans un EIAH : je peux savoir que l'apprenant a construit un point, un segment, qu'il a déplacé la figure. Je peux donc reconstituer la sémantique de l'action qui est liée au processus d'apprentissage » (Luengo et Peraya, 2019).

Des données épaisses mais en nombre insuffisant risquent cependant de pénaliser la reconstitution de cette « sémantique ». C'est pourquoi la chercheuse y adjoint l'objectif d'« hétérogénéité » pour multiplier les protocoles de collecte, afin de renseigner différentes dimensions de l'action d'apprentissage :

> « J'ai aussi avancé la notion de donnée hétérogène qui est liée à celle d'épaisseur. Il y a bien sûr plusieurs formes d'hétérogénéité de la donnée, mais j'utilise le terme dans l'acception que voici. Il s'agit de multiplier les traces qui décrivent une action de l'apprenant afin d'enrichir le plus possible la description de celle-ci, autrement dit de lui donner cette épaisseur. Il convient donc de croiser plusieurs traces de nature différente relative à la même action de l'apprenant : je peux par exemple savoir que celui-ci a déplacé un objet (trace dans l'EIAH), quels sont les gestes qu'il a effectués, quels sont les événements (déplacement de la souris, par exemple) qui se sont déroulés et ce qu'il a regardé au même moment (donnée oculométrique) » (ibid.).

La quête d'épaisseur et d'hétérogénéité se heurte cependant à des contraintes pratiques et éthiques. Sur le plan pratique, ces traces doivent être étiquetées, manuellement ou automatiquement sous la supervision de chercheurs, pour reconstituer leur lien sémantique à une action. Cette contrainte rend les jeux de données utilisables plus difficiles d'accès. Les chercheurs doivent ainsi conjuguer les gisements laissés ouverts par des initiatives universitaires (comme le *Datashop* du *Pittsburgh Science of Learning Center*) et certains sites de *e-learning* ouverts au partage.

Sur le plan éthique, la quête de données épaisses dépend de l'identification personnelle de l'apprenant et potentiellement du partage de ses données au-delà de plateformes propriétaires – par exemple entre EIAH et systèmes d'information et de gestion –, dans une logique d'interopérabilité. Cela pose des questions de confidentialité importantes, eu égard aux obligations légales liées à la sécurisation des données personnelles, comme celles du règlement général sur la protection des données (RGPD) en Europe. L'anonymisation, processus de chiffrage de données consistant à éviter l'identification de la personne à laquelle elles appartiennent, est une voie couramment privilégiée pour s'y conformer. Elle complique cependant les recherches sur la personnalisation sans garantir une confidentialité parfaite (Rocher, 2019).

Si la pseudonymisation, procédure de « chiffrage-déchiffrage » permettant de relier des données à un apprenant, s'accorde mieux aux protocoles de recherche, elle ne résout pas le problème de la confidentialité. Ce problème est d'autant plus brûlant que la quête de capacités diagnostiques plus grandes dans les EIAH semble conduire à collecter et partager des données au-delà de la sphère proprement éducative. Par exemple, pour rendre les ITS capables d'identifier des troubles comme la dyslexie, l'accès à des documents médicaux est une piste heuristique. Néanmoins le partage de ces données sensibles contient des risques importants en termes de respect de la vie privée.

Ainsi le dilemme entre le volume et l'épaisseur des données suscite des débats et des positionnements variés au sein des communautés de l'EDM et des LA. Ces deux objectifs révèlent, à la fois sur des plans épistémologiques, pratiques et juridiques, la gageure d'une substitution de l'algorithme au jugement critique de l'enseignant.

2. La question des contenus

Depuis les premiers « tuteurs intelligents », l'adaptativité requiert des contenus très nombreux et divers, en termes de niveaux de difficulté, de granularité et de formats. Ces matériaux, qu'ils prennent la forme de feedbacks, de questions, de problèmes à résoudre ou de leçons, doivent en outre être conformes aux attentes pédagogiques des organisations éducatives. Ces contraintes nombreuses sont difficiles à concilier pour des équipes de recherche généralement dominées par les sciences informatiques – *a fortiori* en EDM –, au détriment des sciences de l'éducation, aux budgets de recherche souvent insuffisants pour répondre à ces défis.

Les opportunités des grands modèles de langage

Des perspectives sont néanmoins ouvertes par la génération automatique de contenus, relancée à travers les développements récents de ce qui est communément appelé l'« IA générative ». Dans le champ des technologies éducatives, le recours à la génération de contenus est relativement ancien. Les premiers programmes génératifs des années 1960 produisaient et résolvaient automatiquement des problèmes. Ils furent ensuite améliorés grâce aux progrès du traitement automatique du langage naturel pour instrumenter des agents conversationnels, à l'image du pionnier *Scholar*.

Des avancées notables sont arrivées dès les années 2000 par le recours au *machine learning* dans le domaine de la traduction automatique. Comme le relate Yann Le Cun (2021, p.236), un collectif de chercheurs emmené par Yoshua Bengio (Bengio et al., 2003) a introduit le principe de réseaux

de neurones transformant chaque mot en vecteur dans un texte d'entrée, afin de prédire, en sortie, la probabilité qu'un vocable suive une séquence de mots. Le principe a été peu à peu amélioré, avec l'adoption de nouveaux types de réseaux et de l'accroissement des bases de données d'entraînement. En 2017, des ingénieurs de Google inventent notamment une architecture, dite « Transformeur », dont l'innovation principale est le principe de la « *self-attention* ». Celui-ci permet alors d'encoder le contexte dans lequel chaque mot est utilisé dans sa représentation vectorielle (Vaswani et al., 2017).

Arrivent ensuite les modèles « auto-supervisés », capables de s'entraîner seuls sur des ensembles de données, sans annotation humaine. Ainsi Google propose-t-il en 2018 le modèle *BERT* (*Bidirectional Encoder Representations from Transformers*). Celui-ci cache 10 à 20% d'un texte et le réseau doit prédire les mots manquants en utilisant le contexte des autres mots dans la phrase. « Pour cela, le réseau doit apprendre le sens des mots et la structure des phrases » selon Le Cun (2021, p.244). Les transformeurs ont surtout l'avantage décisif de nécessiter moins de temps de formation que les architectures neuronales antérieures.

OpenAI reprend ce principe en lançant en 2018 son premier « grand modèle de langage » ou « *large language model* » (LLM) intitulé *GPT-1* (*Generative Pre-Trained Transformer*). Son principe est similaire à celui de BERT, à ceci près que le système est doté d'un décodeur lui permettant de générer librement du texte en opérant des prédictions statistiques mot après mot. D'autres mastodontes du numérique se jettent rapidement dans la bataille de « l'IA générative », comme Meta, Microsoft ou Nvidia. Depuis lors, chacun de ces acteurs investit dans des améliorations de son modèle[20].

En novembre 2022, la distribution gratuite par OpenAI de son agent conversationnel *ChatGPT* (fonctionnant à partir de la version 3.5 de GPT) a créé une effervescence médiatique mondiale autour des performances techniques de ces modèles. Ceux-ci sont désormais capables d'entretenir des conversations variées avec une grande précision syntaxique. Surtout leur entraînement sur des bases de données issues du web donne une impression d'omniscience, là où leurs réponses ne sont que des calculs de probabilité – les chercheurs parlent de « perroquets stochastiques » (Bender et al., 2021). Les grands modèles de langage offrent ainsi des opportunités pour répondre au manque de contenus dans les EIAH, tout en soulevant des questionnements.

Notons qu'ils ne sont pas en soi « adaptatifs » au sens pédagogique que nous avons fixé au début de cette partie, car ils ne prennent pas en compte les besoins pédagogiques non-formulés explicitement par l'apprenant. Ils n'ont pas la vocation éducative des « tuteurs intelligents » qui modélisent les capacités des apprenants sur un temps plus ou moins long. En revanche, pour des besoins explicités par l'usager (une demande d'explication dans un style énonciatif spécifique, ou la décomposition d'une notion), les perspectives en termes de personnalisation sont nombreuses. Ainsi ces architectures, couplées à celles des systèmes de tutorat, peuvent répondre au manque de contenus de remédiation et contribuer aux recherches sur l'adaptativité.

Interrogé en septembre 2023 sur la réception de ces innovations chez les chercheurs, Peter Brusilovsky notait l'émoi qu'il a vu, en tant que membre actif de la communauté de l'*Educational Data Mining* :

[20] Les améliorations concernent le volume croissant des bases de données d'entraînement ou la multi-modalité, pour permettre à ces IA de « comprendre » et générer des contenus autres que textuels, comme des images, des sons ou des vidéos.

> « Dans presque toutes les conférences auxquelles j'ai participé au cours des six derniers mois, les discussions se sont concentrées sur les LLM d'une manière ou d'une autre. Par exemple, une conférence récente sur l'IA était censée comporter trois présentations sur trois sujets différents, les graphiques de connaissances, le traitement du langage naturel et un sujet complètement différent de ceux-là. À la fin, tous les trois portaient sur les grands modèles de langage, et *ChatGPT* en particulier. Les experts essayaient de comparer ce qu'ils faisaient auparavant dans leurs domaines et ce qui s'y passe actuellement, pour voir ce que pourraient apporter les LLM. Cela nous a donc tous influencés. C'est pourquoi je pense que ce sera également la prochaine expansion dans le domaine de l'éducation »[21] (Entretien 2).

Les LLM suscitent en particulier l'intérêt des chercheurs pour la génération de feedbacks. Bruce McLaren, chercheur en *Educational Data Mining* à Carnegie Mellon travaillant en particulier sur les jeux éducatifs, explore par exemple la manière dont ces systèmes peuvent générer des commentaires pour soutenir le développement de connaissances déclaratives chez l'apprenant :

> « Par exemple dans un jeu, nous invitons les étudiants à expliquer eux-mêmes les étapes qu'ils suivent pour résoudre des problèmes. S'ils le font de manière ouverte, en tapant simplement leur explication, on peut imaginer utiliser *ChatGPT* pour analyser ce qu'ils disent et ensuite fournir des commentaires en utilisant toutes les connaissances qu'il pourrait glaner sur internet […] »[22] (Entretien 3).

Une autre perspective d'adaptativité est la génération de problèmes adaptés aux centres d'intérêt de l'élève. Des bénéfices éducatifs seraient attendus *a fortiori* chez des enfants ayant des difficultés dans la matière en question. D'un point de vue pédagogique, ce mode de génération obéit en apparence à l'agenda constructiviste : il s'agit d'amener l'élève à « construire » lui-même du sens, en personnalisant les objets rencontrés selon ses domaines de prédilection. Cependant, l'éditorialisation des exercices conserverait une perspective behavioriste, le programmeur imposant des exercices dans l'attente de comportements en sortie.

Des doutes et des velléités de « *fine-tuning* »

Toutefois la tendance de ce type d'agent à répondre coûte que coûte aux questions posées, quitte à commettre des erreurs, de même que l'inégale pertinence pédagogique de leurs sources d'entraînement, sont autant de limites à la fiabilité des contenus générés. Peter Brusilovsky pointe du doigt le registre encore trop « encyclopédique » et peu pédagogique des explications fournies par les agents conversationnels, du fait d'un entraînement sur des documents écrits pour des publics savants. Ce reproche fait écho à celui qui était déjà adressé aux systèmes experts comme *GUIDON* dans les années 1980 :

> « Dans notre nouveau travail avec l'université de Memphis, nous [...] construisons un tuteur qui peut dialoguer avec les étudiants. Ainsi nous essayons d'utiliser les modèles de langage à la place d'experts. Toutefois ce que nous découvrons, c'est que les explications produites sont plus compliquées, plus verbeuses et plus éloignées des étudiants que celles produites même par des experts. Je pense que c'est le résultat de l'entraînement du LLM sur des bases de connaissances professionnelles sur le web. Cependant, pour réussir en éducation, il faudrait l'entraîner davantage

[21] *« [At] almost every conference attended in the last half a year, yeah, all the discussion has been focusing on LLMs on one way or another. Just an example of one of the recent AI conference, which was supposed to have three keynotes on three different topics, [...] knowledge graphs, NLP and some topic completely different from that. At the end, all three keynotes were about large language models and ChatGPT specifically and the experts tried to compare what they did before in their fields and what in the fields is happening now and how potentially we can leverage LLMs together with what they do. So it's influenced all of us. That's why I believe it will be the next expansion in education too... »*

[22] *« Like for instance in a game we prompt students to self-explain the steps that they take in solving problems, and you know if they do that in an open-ended way where they sort of just type in their explanation, you could imagine using ChatGPT to analyze what they say and then provide some feedback using all of the internet knowledge that it's able to scoop up […] »*

sur de la littérature pédagogique, spécialement écrite pour expliquer les choses clairement aux personnes qui n'ont pas les connaissances… »[23] (Entretien 2).

Carolyn Rosé, autre chercheuse à Carnegie Mellon et figure de proue du *Computer-Supported Collaborative Learning*, se montre moins optimiste quant aux apports des LLM à l'éducation. Elle défend une position additive de la technologie, soulignant la primeur de l'interaction humaine : « Je pense que la pédagogie est bien plus importante que la technologie »[24] (Entretien 1). Prônant des incursions spécialisées de la technologie, elle réfute de ce fait le mythe d'une « IA générale » derrière ces robots conversationnels, selon lequel ceux-ci seraient capables de simuler des activités intellectuelles, et notamment pédagogiques dans n'importe quel domaine. Elle y voit plutôt un vecteur d'aliénation :

> « Je préconise fortement de cibler des applications spéciales parce que je pense que, encore une fois, il y a toute cette rhétorique sur l'intelligence générale qui n'est pas vraie, n'existe pas, c'est un mythe. Je pense qu'il y a beaucoup de choses en général que vous pouvez faire avec quelque chose comme *ChatGPT* jusqu'à ce que vous en dépendiez vraiment, puis vous voyez où il tombe en panne »[25] (ibid.).

Parmi ces dysfonctionnements, elle cible les garde-fous que rajoutent les programmeurs, de nature à dissimuler l'état des connaissances réelles dans un domaine :

> « En fait, je dirai juste qu'un problème avec *ChatGPT* est […] qu'ils ont construit des couches de protection autour de lui qui affectent ce que vous voyez réellement […]. Je pense qu'il est dangereux qu'il y ait cette couche qui vous empêche en quelque sorte de voir quelles sont les sources […], car cela isole vraiment l'utilisateur final des choses que nous enseignions comme étant de la littéracie numérique. Donc je regarde ce qui est fait à ce sujet et je pense qu'en fait, ce qui va se passer, c'est que *ChatGPT* ne continuera pas à être la norme pour toujours, les gens finiront par se rendre compte que c'est une mauvaise idée »[26] (ibid.).

À rebours d'une « IA générale », la chercheuse évoque le besoin de « *fine-tuner* » les modèles, c'est-à-dire de les entraîner sur des ensembles de données plus spécialisés, pour mieux les adapter à des tâches spécifiques. Cette nécessité est partagée parmi la communauté des chercheurs en EDM et LA. Peter Brusilovsky l'évoque par exemple pour résoudre le problème du manque de pédagogie des explications générées, en entraînant ces modèles sur des corpus de manuels ouverts, ou par le rapprochement entre des acteurs de l'édition et des LLM. Ces voies lui font prédire un dépassement de cet obstacle dans un avenir proche : « C'est une petite barrière, il faut peut-être deux ou trois ans jusqu'à ce que les gens s'adaptent »[27] (Entretien 2).

En définitive, la réception de ces technologies par la communauté scientifique est un autre révélateur des positionnements pédagogiques différenciés des chercheurs. Favorisant jusqu'ici une approche

---

[23] *« [In] our new work with University of Memphis we're [...] building a tutor which can interact with the students in a dialogue way. And we try to use language models instead of experts. But what we're finding out is that explanations produced by large language models are more complicated, more verbose and farther away from students than an explanation produced even by experts. And I think that is the outcome of LLM being trained on a professional board of knowledge on the web, while to succeed on the educational domain, you kind of need to train it more in education literature which has been specifically produced to explain things clearly to people who don't have the knowledge ».*

[24] *« I think that the pedagogy is way more important than the technology »*

[25] *« I strongly advocate for targeting special applications because I think that again there's all this rhetoric about generalized intelligence that's not true, doesn't exist, it's a myth. So I mean I do think that there's a lot in general that you can do with something like ChatGPT until you really depend upon it to do it well and then you see where it breaks down…».*

[26] *« Actually I'll just say one problem with ChatGPT is that it isn't even just the large language model but they built some layers of safeguards around it that affects what you actually see […] I think that it's dangerous that there's this layer that sort of prevents you from seeing what are the sources […] because it really insulates the end user from the things that we used to teach as information literacy so we're looking at what do you do about that and I think actually what's going to happen is that ChatGPT is not going to continue to be the standard forever, people are eventually going to realize that that's a bad idea ».*

[27] *« [It]'s a small barrier, it just like take maybe two or three years until people will adapt on both sides ».*

essentiellement behavioriste de l'apprentissage, les modèles de langage suscitent l'intérêt des partisans des « tuteurs intelligents », pour qui l'adaptativité pourrait être étendue. Toutefois des résistances émergent chez les défenseurs d'une technologisation additive, qui craignent autant une illusion technique qu'une perte de littératie des usagers.

3. L'explicabilité des décisions algorithmiques

L'informaticien Michael Jordan soulignait, dans un article de 2019 intitulé « *Artificial Intelligence – The Revolution Hasn't Happened Yet* », les dangers du déploiement d'un système d'aide à la décision dans des secteurs stratégiques comme la santé ou l'éducation, où les humains ne sauraient décomposer la « *provenance »* des décisions algorithmiques :

> « Le problème ne concern[e] pas seulement l'analyse des données en soi, mais aussi ce que les chercheurs en bases de données appellent la *provenance* : de manière générale, d'où proviennent les données, quelles déductions ont été tirées de ces données et dans quelle mesure ces déductions sont-elles pertinentes par rapport à la situation actuelle ? »[28] (Jordan, 2019).

En découlent, depuis plusieurs années, des recherches très actives dans des domaines comme l'« *explainable AI* » (XAI), la « *human-AI collaboration* » ou la « *human-centered AI* », afin de permettre aux humains d'implémenter ou non des décisions algorithmiques de façon éclairée.

L'émergence des enjeux d'explicabilité dans l'éducation

Dès les années 1990, la collecte et le traitement croissants de données dans les EIAH ont fait émerger le défi de l'explicabilité des décisions algorithmiques auprès des publics. Celui-ci est en particulier mis à l'agenda scientifique par l'informaticienne Judy Kay au tournant des années 1990 et 2000, lorsqu'elle appelle à la prise en compte de la « *scrutability* » – que nous traduisons par « explicabilité » – dans la conception des environnements informatiques. En 2006, elle le justifie du point de vue de l'apprenant par plusieurs raisons (Kay, 2006). Trois d'entre elles sont :

- Le droit de l'individu à juger de l'utilisation faite de ses informations personnelles, en ayant connaissance du rôle de la machine, des méthodes de personnalisation et des responsabilités du programmeur ;
- Son droit à exercer un contrôle sur l'outil, en corrigeant si nécessaire les erreurs du modèle ;
- La pertinence pédagogique d'une visualisation de son propre comportement d'apprentissage, à des fins d'autorégulation.

Toutefois les décisions algorithmiques des EIAH, avec le passage à une IA de type connexionniste et le développement du *data mining*, ont perdu en explicabilité. Censés défendre la capacité d'action des publics éducatifs et notamment des enseignants, les *Learning Analytics* sont, comme on l'a vu plus haut, confrontés à une contradiction : les nouvelles interprétations de l'apprentissage qu'ils offrent reposent en partie sur les techniques de *data mining*. Charge alors aux usagers de faire confiance à ces dispositifs, et de fermer les yeux sur les risques d'erreurs ou de biais algorithmiques.

[28] « *The problem had to do not just with data analysis per se, but with what database researchers call provenance — broadly, where did data arise, what inferences were drawn from the data, and how relevant are those inferences to the present situation?* ».

Les biais des modèles de *machine learning*

De nombreux travaux critiques ont ainsi souligné la tendance des modèles de *machine learning* à reproduire, par leurs décisions, des biais contenus dans leurs données d'entraînement. Hu et Rangwala (2020) ont souligné que les inégalités de genre ou de race contenues dans les jeux de données éducatives étaient nombreuses et susceptibles de conduire à des décisions algorithmiques discriminantes. Parmi les nombreux risques pédagogiques, les prédictions d'abandon ou d'échec scolaire de certains groupes d'élèves peuvent influer sur les pratiques des professeurs en classe et conduire, selon une prophétie auto-réalisatrice, à leur stigmatisation et leur découragement. Doroudi et al. (2019) ont en outre montré, sur les modèles de *knowledge tracing*, que les inégalités se multipliaient à mesure que la quantité de contenus pédagogiques augmentait.

Du côté des grands modèles de langage, si l'instauration de garde-fous prévient la génération de déclarations racistes, des chercheurs indiquent également l'existence de biais latents. Hofmann et al. (2024) ont notamment insisté sur les stéréotypes que les LLM contiennent sur les interlocuteurs écrivant en *African American English* (AAE), à même de les conduire à pénaliser ces personnes dans des domaines de recommandation variés, dont l'éducation.

Les risques sont d'autant plus pressants que les ITS se complexifient. Depuis quelques années, les recherches croissent en EDM et LA pour neutraliser ces biais d'une part, et rendre les décisions algorithmiques plus intelligibles pour les publics éducatifs d'autre part.

Peter Brusilovsky, qui a poursuivi à partir des années 1990 la recherche sur les « *open learner models »* comme on l'a vu, se félicite d'un retour en grâce de ces questions au sein de la communauté des chercheurs, après un passage à vide :

> « En fait, l'IA dans l'éducation a été l'un des premiers endroits où l'IA « centrée sur l'humain » a commencé. Il s'agissait des modèles d'apprentissage ouverts. C'était le premier cas dans l'histoire de l'IA où le système essayait de montrer à l'utilisateur certaines connaissances sur lui tout au long du processus, et permettait à l'utilisateur de changer cela. C'était complètement révolutionnaire. Le reste de l'IA n'est arrivé là que dix ans plus tard. Cependant, après cela, l'IA dans l'éducation n'a plus vraiment contribué à la collaboration « humain-IA ». Toutefois, cela progresse aujourd'hui. J'ai vu un certain nombre de personnes qui ont commencé à travailler là-dessus. Dans notre équipe, cela a toujours été un sujet, mais il n'a pas vraiment trouvé d'échos, donc nous le présentions rarement. Quand nous le faisions, personne ne s'en souciait. Mais maintenant, quand nous présentons des choses sur le sujet, les gens s'enthousiasment, et nous commençons à travailler davantage sur la même chose »[29] (Entretien 2).

Le regain d'intérêt pour l'explicabilité vise ainsi à fournir aux publics éducatifs une compréhension des méthodes ayant présidé à une décision algorithmique. L'enjeu est d'une part de permettre à l'apprenant de comprendre les raisons de l'affichage d'un parcours adapté, dans une logique critique proche de celle décrite par Michael Jordan, et d'autre part de l'aider à prendre conscience de sa progression et de ses besoins cognitifs, dans une logique métacognitive décrite plus haut par Judy Kay. L'enseignant doit pour sa part pouvoir juger des facteurs ayant causé l'affectation de tel parcours ou tel contenu à un élève. Il doit ainsi revoir ses jugements et ses pratiques personnelles,

[29] *« In fact AI in education was one of the first places where human-centered AI started, and that was this open learning models. It was the first case in the whole AI where the AI system tries to show the user some knowledge about the user through the process, and engaged the user in potentially changing that. That was a completely groundbreaking. The rest of AI didn't get there until 10 years later. But kind of after that, AI in education hasn't been really contributing to this human-AI collaboration, but this is actually pushing now. So I've seen a number of people who started to work on that, and in our team it was always a topic, but it hasn't been really resonating, so we rarely presented to them. We presented, nobody cared. But now when we present on the topic, people get enthusiastic so we start doing more work about the same thing ».*

ou au contraire annuler une décision algorithmique, afin de prendre lui-même en charge la différenciation.

Des méthodes de *computer science* et interdisciplinaires réinterrogeant l'adaptativité

Pour expliquer des décisions tirées du *machine learning*, des méthodes algorithmiques sont premièrement développées au sein de l'« *explainable AI* », branche de la *computer science* en plein essor. Parmi les multiples techniques investiguées, l'une des plus courantes est le recours à une « analyse de sensibilité », qui consiste à chercher la plus petite perturbation des données d'entrée qui inverserait la décision, afin de localiser les variables critiques.

Des chercheurs en EDM et LA mènent également des recherches spécifiques au domaine de l'éducation, en investiguant par exemple l'adaptation d'explications selon le modèle de l'apprenant (Brusilovsky, 2024).

Cependant, des doutes subsistent sur la possibilité de parvenir, en toute circonstance, à une explication précise et intelligible. En cause, la multiplicité des variables présentes dans les calculs opérés par les algorithmes de *machine learning, a fortiori* dans les réseaux de neurones. De plus, la pertinence d'une explication pose question sur des décisions algorithmiques isolées, ciblant un niveau d'adaptation fin (comme l'affichage d'un mot dans un exercice de langues). Celle-ci peut être fastidieuse pour le développeur et peu intéressante pour le formateur.

Le chantier de l'explicabilité entretient en définitive un rapport plus mesuré et éthique à l'idéal d'adaptativité, loin de la technologisation substitutive et d'un tropisme selon lequel l'humain devrait s'adapter aux dernières innovations technologiques. Cette posture peut au contraire conduire les acteurs de l'éducation à conditionner l'intégration éducative de techniques à leur intelligibilité, et refuser la perte de libre-arbitre qu'entraînent les systèmes de *machine learning* dans des tâches éducatives critiques.

Ainsi ces recherches posent des questions fondamentales pour l'avenir de l'adaptativité. Renoncer à accroître la capacité prédictive (et dès lors la complexité) des calculs algorithmiques pour privilégier leur explicabilité est également susceptible d'affecter l'utilité pédagogique de ces médias, et leur capacité à prédire des aspects de l'apprentissage non-visibles voire contre-intuitifs pour les enseignants.

Conclusion provisoire

Il est temps de poser à nouveau frais notre problématique de départ, celle de l'établissement, avec l'*adaptive learning,* d'un paradigme éducatif singulier, cohérent et transversal aux cultures françaises et états-uniennes ainsi qu'aux secteurs éducatifs. Dans les laboratoires où l'idéal d'adaptativité s'est formé au début du XX^{e} siècle, le chemin parcouru sur le plan technique a été long – de l'adaptation sommaire permise par l'enseignement programmé, aux tuteurs informatiques reposant sur l'IA symbolique puis aux EIAH usant de techniques de fouille de données et de grands modèles de langage. Avec chacun des artefacts, les champs d'adaptativité se sont accrus : adaptation des tâches à réaliser dans les machines à enseigner, premières adaptations des feedbacks et de la difficulté des questions dans les tuteurs modélisant la maîtrise de l'apprenant, puis adaptation des remédiations (*adaptive content),* des contenus des questions (*adaptive assessment*) et de l'ordre des notions

(*adaptive sequencing*) dans les EIAH, grâce à la collecte et au traitement en continu des données des élèves. Les modes pédagogiques ont eux aussi varié, d'un « *drill and practice* » behavioriste à un tutorat d'abord très encadré puis assoupli par des formes de pédagogie actives et collaboratives, sous l'influence du constructivisme et de ses courants affiliés. Aussi divers soient-ils, la singularité des projets adaptatifs, à rebours d'une forme scolaire impersonnelle, est bel et bien avérée.

Toutefois les antagonismes théoriques se répercutent sur la question de la place de la machine par rapport à l'enseignant, reposée à chaque innovation technologique. Le débat oscille traditionnellement entre une technologisation substitutive, technocentrée, équivalente à un remplacement du formateur par la machine, et une technologisation additive, équivalente à une assistance automatisée du formateur. Si à première vue le behaviorisme et son corollaire technocentriste ont décliné dans le champ académique au profit du constructivisme et d'une meilleure prise en compte des publics éducatifs, nous avons instillé jusqu'ici des nuances importantes : les tuteurs intelligents et les EIAH ont démontré d'une part la survivance du behaviorisme sous des formes souterraines et nouvelles, notamment au contact du *Big Data*. Nous avons d'autre part observé la mutation de ses avatars au contact des thèses constructivistes, à mesure que les fonctionnalités sociales et interactives des EIAH se sont développées.

Les communautés actuelles des *Learning Analytics* et de l'*Educational Data Mining* réactualisent ces débats. L'une prône un usage de la technologie en appoint des pratiques pédagogiques, vers davantage de personnalisation et de collaboration entre pairs, malgré le défi de l'intelligibilité des décisions algorithmiques. L'autre favorise la substitution de l'algorithme à l'intervention humaine sur de nombreuses tâches, à partir des préceptes de l'épistémologie de la donnée, masquant la gageure d'une technologisation substitutive. Ces positions entraînent des réponses variées aux principaux défis techniques qui animent la recherche actuelle, comme le dilemme entre volume et épaisseur des données récoltées, la nécessité de contenus de remédiation nombreux et variés ou l'explicabilité des décisions algorithmiques auprès des publics. Ces dissensions, si elles sont heuristiques, pénalisent l'impulsion d'un paradigme éducatif cohérent depuis les laboratoires. Au mieux, les différences de vues représentent-t-elles un nuancier de modèles « en puissance » pour les acteurs de l'éducation et de la formation.

Il est temps de confronter ces projets aux paysages socio-économique de l'éducation et de la formation aux États-Unis et en France. Il s'agit d'analyser les stratégies des acteurs industriels concourant au déploiement de ces technologies sur les trois plans détaillés en introduction : les politiques publiques, les stratégies de firmes et les professionnalités. Par l'étude de leurs rapports de force, nous retraçons les conditions de formalisation d'un paradigme éducatif.

DEUXIÈME PARTIE. De l'expérimentation à l'industrialisation : morphologies des industries éducatives revisitées par l'*adaptive learning*

La diversité technique des prototypes adaptatifs a démontré que de nombreux secteurs et domaines éducatifs pouvaient être revisités. Il s'agit maintenant de jauger la cohérence et la complémentarité des stratégies d'acteurs publics et privés prenant part à l'intégration de ces algorithmes dans les industries éducatives françaises et états-uniennes. Appréhender leur cohésion au-delà des spécificités nationales et industrielles nécessite d'adopter un point de vue socio-économique. Le chapitre suivant introduit la morphologie actuelle de la socio-économie des médias éducatifs et les problématiques que l'adaptativité y soulève.

Chapitre IV. Une socio-économie des médias éducatifs en suspens

Une précaution lexicale mérite d'être posée d'emblée. La circulation du syntagme de l'*adaptive learning* dans l'espace public en fait une « expression obstacle » (Bachelard, 1938), au sens où il crée une confusion sur la définition de l'adaptativité. S'en réclament des services plus ou moins sophistiqués, d'une part en termes d'ambitions pédagogiques – du tutorat dans une matière à un simple aide à la mémorisation – et d'autre part en termes techniques – de médias reposant sur des algorithmes de *machine learning* à des systèmes experts écrits par des ingénieurs, en passant par des programmes arborescents n'évaluant les élèves que de façon ponctuelle. Ces derniers ne répondent pas au critère d'adaptativité qu'est la collecte de données en continu.

Nous tâchons donc de distinguer dans l'espace socio-économique les médias proprement adaptatifs, au regard des trois critères définitionnels que nous avons donnés en introduction, des dispositifs statiques profitant de l'« aura » technologique de l'*adaptive learning*. Ainsi, nous comptons cerner les stratégies d'acteurs industriels participant au déploiement de l'adaptativité *stricto sensu*, de même que les opportunités d'ordre communicationnelle et marketing que d'autres acteurs saisissent dans cet idéal.

1. Des questionnements *dans* et *entre* les industries éducatives

Pour appréhender la diversité des secteurs éducatifs revisités par les algorithmes d'apprentissage adaptatif, nous considérons une « socio-économie des médias éducatifs » (Moeglin, 2005). Empruntée à la théorie des industries culturelles, cette focale englobe les modes de régulation des pouvoirs publics, les conditions de conception, de production et de valorisation des médias, et les relations contradictoires que les filières industrielles entretiennent les unes avec les autres.

Comme nous l'avons expliqué précédemment, celles-ci sont au nombre de cinq : l'édition de manuels, l'informatique pédagogique, la formation à distance, les produits pour l'éducation non-formelle et les systèmes numériques d'information et de gestion.

Notre hypothèse exposée en introduction est que les algorithmes d'apprentissage adaptatif contribueraient à la centralité des acteurs des technologies dans cette socio-économie, ainsi qu'à la convergence entre les industries culturelles et éducatives. Pour l'éprouver, il est nécessaire de délimiter en amont, pour chacune de ces industries, les rapports de force et les questionnements spécifiques qu'y soulève le déploiement d'algorithmes d'apprentissage adaptatif.

Cinq filières revisitées

L'édition de manuels

Plus ancienne filière, l'édition de manuels entretient un lien ténu avec les systèmes éducatifs nationaux, pour lesquels elle fournit ce média représentatif de la forme scolaire. Cette industrie concentre des oligopoles puissants. En France, le marché de l'édition scolaire, dont le chiffre d'affaires a atteint 296,8 millions d'euros en 2023 selon les chiffres du Syndicat national de l'édition, est tenu par six éditeurs. Ce sont des filiales de grands groupes de l'édition généraliste (Syndicat

National de l'Édition, 2024, p.10). Nous en dressons ici la liste, avec leurs chiffres d'affaires respectifs lorsqu'ils sont disponibles :

- Hatier et Hachette Education, filiales de l'éditeur Hachette Livre. Hatier représentait près de 57,5 millions d'euros de chiffre d'affaires en 2023 en France selon Infogreffe. Hachette Livre totalisait 2,8 milliards d'euros de recettes dans le monde en 2023 (Hachette, 2024).
- Nathan et Bordas, deux filiales d'Editis. Nathan représentait près de 11,1 millions d'euros de chiffre d'affaires en 2023 en France selon Infogreffe. La maison-mère comptabilisait 789 millions d'euros de chiffre d'affaires en 2022.
- Magnard-Vuibert, totalisant 37,2 millions d'euros de recettes en 2023. Sa maison-mère, Albin Michel, rassemblait près de 86,5 millions d'euros de recettes en 2022 selon Infogreffe.
- Belin, dont les derniers comptes publiés datent de 2015 et font état d'un chiffre d'affaires de 17,7 millions d'euros selon Infogreffe. Humensis, la maison d'édition qui a résulté de sa fusion avec les Presses Universitaires de France en 2016, comptabilisait 41,4 millions d'euros de chiffre d'affaires en 2022 selon Le Monde (Vulser, 2023).

Aux États-Unis, le marché de l'édition scolaire, de *Pre-K* (équivalent de la grande section de maternelle) à *K-12* (équivalent de la terminale), aurait représenté 5,61 milliards de dollars de chiffre d'affaires en 2022 (Curcic, 2023). Il s'agit également d'un oligopole tenu par les divisions éducatives de cinq éditeurs. Nous en dressons ici la liste avec leurs chiffres d'affaires respectifs, ceux des maisons d'édition de manuels n'étant pas disponibles :

- Scholastic, ayant amassé 1,7 milliards de dollars en 2023 (Scholastic, 2024).
- McGraw Hill, qui a comptabilisé 1,54 milliards de dollars de chiffres d'affaires en 2021 (Watson, 2022) ;
- Cengage, qui a amassé 1,48 milliards de dollars pour son année fiscale 2023 (Cengage Group, 2023) ;
- Houghton Mifflin Harcourt totalisant 1,05 milliards de dollars en 2021 (Houghton Mifflin Harcourt, 2022) ;
- Savvas Learning Group, ancienne division éducative de l'éditeur Pearson, qui ne communique pas son chiffre d'affaires.

Leur transition vers des catalogues numériques est un enjeu ancien et ambivalent, différemment traité par les éditeurs français et états-uniens. En France, la rentabilité du manuel numérique est bien inférieure à celle de son homologue imprimé selon Xavier Levoin, pour des questions de coûts de production et de mode de rémunération (Levoin, 2017). Toutefois le chercheur rapporte la tendance des éditeurs à se positionner sur tout nouveau marché pouvant représenter à terme un avantage concurrentiel, pour ne laisser personne « rafler la mise ». Aux États-Unis, cette transition est engagée depuis plus longtemps et à plus grand renfort de capitaux. Les raisons en sont multiples : la taille du marché *K-12* des médias numériques éducatifs est bien plus importante qu'en France. Par ailleurs, le marché des manuels imprimés y est depuis longtemps en décroissance, en raison de l'augmentation des prix de vente, du développement de services de location et de l'essor des logiciels éducatifs et autres ressources éducatives libres sur le web. Les grands groupes états-uniens ont ainsi entamé depuis les années 2000 leur transition vers le numérique éducatif, rachetant notamment des structures de production informatique. Cette tendance s'est confirmée avec l'annonce, en 2019 par Pearson, de la vente de sa division *K-12* et d'une stratégie « *digital first* », consistant à louer ses titres par défaut sous la forme d'*e-book*s, et à ne plus éditer sous format papier qu'une centaine de titres (Rosenblatt, 2019).

L'adaptativité, incompatible avec l'ordre séquentiel de lecture propre au manuel numérique, est susceptible d'apporter un nouveau bouleversement. Face aux perspectives que cette technologie représente, notamment en termes de connaissances des publics scolaires via les traces d'usage, les éditeurs français et états-uniens n'adoptent-ils qu'une attitude de participation marginale, laissant des petites structures prendre la plupart des risques ? Ou anticipent-ils un changement de paradigme en y investissant des ressources significatives ?

L'informatique pédagogique

La deuxième industrie, l'informatique pédagogique, est directement concernée par l'intégration d'algorithmes d'apprentissage adaptatif. Elle a émergé dans les années 1970, dans le sillage des « machines à enseigner ». Elle se distingue par l'ambition de fournir un environnement informatique d'apprentissage à un apprenant, pour une utilisation personnelle lors d'un cours ou d'une formation en présentiel. Prévaut donc un emploi de l'outil en complément de l'enseignement du formateur, selon une organisation individualisée et modulaire de l'enseignement en présentiel, à rebours de la forme scolaire. La place des acteurs publics y est prégnante, en France comme aux États-Unis : les gouvernements tracent les stratégies en matière d'équipement et les établissements, financés par les collectivités territoriales, constituent les acheteurs, commandant le plus souvent des licences annuelles.

Il est délicat de trouver des données chiffrées sur la taille de cette industrie, la plupart des mesures mélangeant les usages présentiels, hybrides, distanciels. L'on peut affirmer en revanche que le secteur, encore peu mature, est constitué d'une multitude de petites structures de production plus ou moins fragiles. Les barrières à l'entrée sont historiquement fortes : difficulté d'accès aux institutions éducatives centralisées en France, flotte de commerciaux insuffisantes pour vendre leurs solutions dans les districts dispersés des États-Unis, coûts élevés de production de contenus pédagogiques…

Parmi ces petites structures, nombre de « start-ups » ont émergé à partir des années 2010 avec la promesse d'une croissance rapide, par levées de fonds auprès des investisseurs publics et privés. On compte cependant des acteurs anciens de l'informatique pédagogique, antérieurs à l'engouement pour l'intelligence artificielle dans les années 2010, qui se sont arrogé des parts de marché importantes.

Aux États-Unis, où le marché scolaire est plus développé du fait de ses circuits de distribution décentralisés, quelques groupes concentrés, ayant grandi par fusions-acquisitions et dotés d'équipes de commerciaux étendues, atteignent des centaines de millions de dollars de chiffres d'affaires, comme Renaissance Learning ou Cambium Learning Group. En France, les ordres de grandeur sont bien moindres, le marché étant administré et les investissements privés bien plus faibles. Le groupe Educlever, l'un des plus importants après sa fusion avec Orthodidacte en 2020, comptabilisait alors entre 7 et 8 millions d'euros de chiffre d'affaires (Seveyrat, 2020). De même des éditeurs de manuels français ou états-uniens se sont positionnés dans cette filière, développant leurs propres outils ou rachetant des structures de production, tel Editis acquérant Educlever en 2022.

L'arrivée d'une nouvelle valeur ajoutée avec les algorithmes d'apprentissage adaptatif a-t-elle offert des avantages concurrentiels aux petites structures qui ont bâti leurs offres autour de ces technologies ? La nécessité d'entraîner les algorithmes sur des bases d'utilisateurs suffisamment

grandes constitue-t-elle une nouvelle barrière à l'entrée pour de jeunes structures, et un avantage pour des acteurs anciens aux publics déjà constitués ?

La formation à distance et le e-learning

Les algorithmes d'apprentissage adaptatif revisitent également l'industrie de la formation à distance. À la différence des deux filières précédentes, la plupart des outils de formation à distance se vendent directement auprès des consommateurs, que ce soient des parents pour le travail « à la maison » de leurs enfants, ou des publics plus âgés et autonomes de l'enseignement supérieur et de la formation professionnelle. D'abord enseignement de substitution, la formation à distance s'est développée avec l'audiovisuel et internet comme une solution concurrente au présentiel.

En France et *a fortiori* aux États-Unis, des formes d'oligopoles se sont constituées, bien qu'immatures et sujets à des perturbations, entre des entreprises aux bases d'utilisateurs mondialisées et une multitude de petites structures fragiles. Dans les divers segments éducatifs cohabitent cependant des offres marchandes avec d'autres à but non-lucratif. Le marché mondial des MOOC est ainsi dominé par deux plateformes états-uniennes, Coursera, cotée en bourse et annonçant 142 millions d'apprenants en 2023 (Coursera, 2023) et son concurrent à but non-lucratif edX, comptant plus de 40 millions d'utilisateurs en 2021 (edX, 2021).

Quels changements apportent les algorithmes d'apprentissage adaptatif à ce paysage morcelé ? Avant eux, le tutorat s'est déjà installé au centre des modalités pédagogiques de nombreuses plateformes, qui ont recours à des tuteurs humains opérant à distance. Néanmoins leurs coûts augmentent avec le nombre d'étudiants. Automatisé avec l'adaptativité, ce mode pédagogique trouverait-il enfin une voie financière viable, à même de renforcer l'industrie ?

D'autres environnements adaptatifs s'apparentent en outre à un sous-secteur de la formation à distance, fruit de son rapprochement et de sa rivalité avec l'informatique pédagogique : le *e-learning*. Né avec internet, ce secteur offre des prestations à distance en complément de formations en présentiel. Celles-ci ont d'abord été proposées par des établissements d'enseignement supérieur capables d'hybrider leurs formations, puis se sont étendues à tous les secteurs éducatifs où un travail autonome « à la maison » est sollicité. Ces médias, souvent appelés *Learning Management Systems* (LMS), permettent à l'enseignant d'éditer des devoirs à assigner à ses étudiants, de communiquer avec eux et de superviser leurs progrès dans des interfaces de *monitoring*. Les étudiants, en plus de réaliser les travaux demandés par l'enseignant, peuvent communiquer entre eux dans des forums.

Le circuit de commercialisation est globalement identique à celui de l'informatique pédagogique, l'acheteur étant l'institution éducative. Au printemps 2021, les trois premiers LMS aux États-Unis en termes de nombre d'institutions d'enseignement supérieur clientes étaient Canvas (avec 38% de parts de marché), Blackboard (avec 25 %) et Moodle (15 %) selon Edutechnica (2021). En France et aux États-Unis, la structure de la filière demeure un oligopole à frange instable, séparant quelques offreurs solidement installés de petites structures innovantes, alors que nombre d'éditeurs scolaires ont lancé leurs propres produits pour l'école et le travail à la maison.

L'adaptativité semble apporter de nombreuses perspectives à ce secteur, tant sur l'aide à la décision des enseignants que l'affectation de parcours autonomes et « sur-mesure ». Ces technologies seraient-

elles à même d'installer cette sous-filière comme une industrie de premier plan, porteuse d'un paradigme éducatif mêlant présentiel et distanciel ?

Les « produits pour l'éducation non-formelle »

Des algorithmes d'apprentissage adaptatif sont également intégrés à une industrie très différente de celles vues jusqu'ici : les « produits pour l'éducation non-formelle ». Nous empruntons cette désignation à Pierre Mœglin, qui rappelle que cette filière est difficile à caractériser. En effet, elle « ne relève ni de l'enseignement formel ni de la culture et du divertissement, et les produits qu'[elle] regroupe sont disparates : livres, journaux, films, programmes audiovisuels et multimédias, émissions de radio et de télévision, sites Internet, etc » (Mœglin, 2010, p.77). D'autre part, « ils font acquérir des connaissances et des compétences, certes, mais en se démarquant de l'éducation instituée et leurs utilisateurs les recherchent d'abord pour la satisfaction qu'ils leur procurent […] » (ibid.). Les disciplines et les tranches d'âge visées sont variées, de l'apprentissage de la lecture chez les enfants à celui des langues chez les adultes. Comme pour la formation à distance, la distribution s'effectue directement auprès du consommateur, parent ou adulte-apprenant.

Le secteur est en croissance rapide du fait des usages toujours plus importants et mondialisés des smartphones, d'internet et des applications mobiles. Les acteurs significatifs en termes de chiffres d'affaires sont des firmes essentiellement états-uniennes, comme Duolingo. Spécialisé dans l'enseignement des langues, cet acteur comptabilise 531,1 millions de dollars en 2023, en hausse de 44% sur un an (Duolingo, 2024). Les emprunts de ce secteur aux industries culturelles sont cependant des obstacles traditionnels pour sa réception par les publics éducatifs, percevant un apprentissage « au rabais » et craignant une soumission de l'éducation, notamment infantile, à la marchandisation publicitaire.

De nombreux services adaptatifs revisitent cette industrie en personnalisant les parcours suivant des contenus diversifiés et censément divertissants. Un exemple représentatif des formats empruntés est la gamification, désignant « l'usage d'éléments de *game design* dans des contextes autres que de jeu » (Deterding et al., 2011). Le déploiement d'algorithmes d'apprentissage adaptatif mêlant divers formats de remédiation serait-il susceptible d'asseoir la légitimité pédagogique du secteur ?

Les systèmes numériques de gestion et d'information

Les dispositifs algorithmiques touchent enfin la filière récente des systèmes numériques de gestion et d'information. Celle-ci se décline en deux types de dispositifs. Le premier concerne les infrastructures de service, tels que les environnements numériques de travail (ENT), permettant d'administrer l'enseignement et de superviser les résultats des élèves. Le second type de dispositifs concerne des outils techno-pédagogiques assurant la gestion de ressources. Ce sont des plateformes d'intermédiation donnant aux élèves et enseignants un accès à un catalogue de ressources numériques.

Le paysage de l'intermédiation est morcelé, avec d'une part des dispositifs gouvernementaux comme le Gestionnaire d'Accès aux Ressources (GAR) dans l'enseignement secondaire français, qui justifie sa présence par un impératif de protection des données personnelles. Il y a d'autre part des plateformes privées telles que *Clever* ou *ClassLink* aux États-Unis, ou des multinationales des technologies comme Microsoft et son application *Teams* dans l'enseignement supérieur. Ceux-ci tentent de s'interposer tant pour des raisons pratiques, comme les protocoles d'identification unique, que pour

des raisons commerciales, en percevant des redevances sur les transactions opérées dans leurs « places de marché ».

Dans quelle mesure ces acteurs sont-ils en mesure d'imposer des normes éditoriales aux structures éditant des médias adaptatifs, notamment en termes de gestion des données personnelles, et de favoriser ainsi de nouveaux avantages concurrentiels ?

Des enjeux transversaux

Ces dispositifs soulèvent en définitive des enjeux socio-économiques *dans* et *entre* les industries éducatives. Outre les questionnements inhérents à chaque filière, opposant des firmes de tailles différentes et d'activités complémentaires ou concurrentes, transparaissent des enjeux transversaux, interrogeant les rapports de forces entre industries. Celles-ci sont traditionnellement animées par des logiques contradictoires, les poussant tantôt à se spécifier les unes des autres, et tantôt à s'allier.

Du côté de la spécification joue la tendance de chaque industrie à prétendre à l'hégémonie et à l'affaiblissement de ses concurrentes. Un exemple est leur concurrence lors de marchés publics pour l'équipement d'établissements en médias éducatifs. Si pendant longtemps le manuel scolaire imprimé est apparu comme l'avatar privilégié de la forme scolaire et de l'enseignement simultané et collectif, les autres industries œuvrent pour institutionnaliser leurs produits – EIAH, chaînes de télévision scolaire, ENT –, et remporter des appels d'offres. Les commandes récentes, *a fortiori* en France, de médias adaptatifs pour équiper les institutions éducatives publiques est susceptible de relancer ces rivalités.

Du côté des alliances jouent les collaborations entre éditeurs de manuels et structures de production informatique, pour la conception d'environnements d'apprentissage. Depuis l'échec retentissant de *Plato* et de Control Data à cause du manque de contenus éducatifs, les structures de l'informatique pédagogique ont pris conscience de la nécessité d'alliances avec les producteurs et ayant-droits de catalogues de contenus. L'adaptativité présente des défis pour chacun de ces acteurs pris seuls, devant encourager *a priori* leurs coopérations. Ainsi les éditeurs scolaires ont besoin de savoir-faire extérieurs pour repenser leurs médias traditionnellement linéaires, et les start-ups doivent se doter de contenus divers et accéder à des bases d'utilisateurs importantes pour entraîner leurs algorithmes.

Des acteurs exogènes semblent pousser également à une homogénéisation des filières du numérique éducatif sous leur autorité : les « *Big Tech* », multinationales états-uniennes ayant été fondées dans la Silicon Valley, dont le pouvoir tient à la fabrication de terminaux informatiques et/ou à la distribution des contenus numériques sur le web. Les plus identifiées médiatiquement sont les « GAFAM » (Google, Amazon, Facebook (aujourd'hui Meta), Apple et Microsoft). Désigner ces acteurs est en soi malaisé, étant donné la forte mythification qui les entoure dans l'espace public. Les expressions de « GAFAM », ou « *Big Five* » aux États-Unis tendent à les ranger dans un « club » élitiste de multinationales implicitement idéalisé. L'hyperbole « *Big Tech* » est certes également orientée. Néanmoins, ne cantonnant pas le signifié à cinq acteurs immuables, elle présente l'intérêt d'être adaptable aux évolutions du secteur.

La pertinence d'une expression pour désigner des acteurs aux cœurs de métiers divers reste questionnable (Bouquillion, 2023, p.38). Pourtant leurs stratégies de développement sont proches,

notamment dans le secteur du numérique éducatif. Ils mettent chacun en place depuis les années 2010 une stratégie de développement « par cercles concentriques » selon Thuillas et Wiart (2023) :

> « (…) Le centre de chaque cercle repose sur un type d'activité spécifique – la recherche d'informations pour Google, le réseautage social pour Facebook, le commerce électronique pour Amazon, les produits électroniques grand public pour Apple –, qui constitue leur point d'ancrage à la fois originel et principal (Bouquillion, Miège, Mœglin, 2013). À travers des lancements de nouveaux services, des partenariats et des acquisitions successives, les Gafa ont progressivement étendu leur périmètre jusqu'au point d'établir et d'imposer leurs propres écosystèmes numériques où l'organisation des activités chevauche les secteurs et se structure autour de plateformes (Pénard, Rallet, 2014). Aujourd'hui, les Gafa sont présents dans la plupart des filières culturelles avec des offres de plateformes spécifiques, mais se positionnent plus largement dans toutes des étapes qui sous-tendent l'accessibilité en ligne des contenus (cloud, système d'exploitation, appareils électroniques, site d'e-commerce, magasin d'application, navigateur, métavers, etc.) ». (Thuillas et Wiart, 2023, p.34)

On peut y ajouter Microsoft, qui a bâti son offre à partir des produits électroniques « grand public » comme Apple. Ces « géants » ont grandi par concentration horizontale, de plateformes d'intermédiation « *BtoC* » en aval jusqu'à la conception de services algorithmiques et de contenus en amont des filières (Smyrnaios, 2016). L'expression « *Big Tech* » traduit ainsi leurs tendances caractéristiques au gigantisme, par une diversification de leurs gammes de services technologiques et une expansion vers de nouveaux marchés géographiques. Cette stratégie leur a permis de diversifier leurs revenus et accroître leur influence sur les industries du web. Cela se vérifie dans le domaine de l'éducation, qui les intéresse historiquement pour trois raisons ciblées par Boissière et Bruillard (2021)[30] :

> « D'abord parce que c'est un marché mondial en expansion (les pays émergents étant souvent vus comme des relais de croissance aux marchés que constituent les pays développés) dans lequel la technologie a une part croissante. Ensuite, parce que c'est un marché récurrent (une fois les usages installés dans les établissements, les besoins de renouvellement sont immenses) et structurant (une fois la technologie reconnue d'utilité éducative, son acquisition est légitime voire vue comme nécessaire pour le bon développement de l'élève). Enfin, probablement pour certains d'entre eux parce qu'un formation initiale réalisée sur leur produit est un moyen de garantir l'acquisition et l'utilisation de leurs solutions à court et moyen terme par le futur adulte » (Boissière et Bruillard, 2021, pp.188-189).

Supposant des capacités de collecte et de stockage de données importantes, les algorithmes d'apprentissage adaptatif sont susceptibles de constituer une porte d'entrée supplémentaire pour ces acteurs centraux des infrastructures technologiques.

Ainsi les industries éducatives sont-elles unies par une socio-économie faite de relations contradictoires entre les filières. Pierre Mœglin relève quatre niveaux interdépendants où sont prises les décisions dans les industries éducatives, de la perspective la plus macroscopique à la plus microscopique : les initiatives publiques, les stratégies des firmes, les professionnalités et les usages (Mœglin, 2010, p.105).

Au premier échelon, les politiques des États et des collectivités sur l'adaptativité et l'intelligence artificielle sont conditionnées, au second échelon, par les stratégies des entreprises et organisations éducatives. Les réalisations de celles-ci dépendent troisièmement de la cristallisation de communautés et de pratiques professionnelles autour du numérique éducatif, et en particulier

[30] Enfin, par le recours à une expression anglaise, nous nous référons implicitement à des acteurs de cet espace linguistique, et écartons les géants chinois (Baidu, Alibaba, Tencent, Xiaomi, Huawei). La raison en est que ces derniers sont très peu implantés dans les marchés éducatifs français et états-uniens, pour des causes stratégiques leur appartenant, mais également en raison de réglementations instaurées dans ces pays, à visées protectionnistes et/ou de souveraineté des données.

de l'intelligence artificielle dans l'éducation. Enfin la légitimité de ces trois niveaux tient au dynamisme des usages, que feraient à la fois les enseignants, les élèves et les publics tiers (administrateurs, parents) des médias adaptatifs.

Ces quatre niveaux de décision conditionnent, outre le déploiement de médias adaptatifs, l'établissement d'un éventuel paradigme éducatif dans l'espace socio-économique. Chacun nécessite des méthodologies spécifiques pour être appréhendé. Notre perspective macroscopique nous limite à l'observation des deux premiers niveaux, les initiatives publiques et les stratégies de firmes, tandis que nos entretiens avec des professionnels et nos visites de salons *EdTech* vont nous permettre d'appréhender le spectre des professionnalités. Nous reprenons ci-après les questionnements liés à chacun de ces échelons.

2. Des politiques du numérique éducatif questionnées par l'adaptativité

En matière de numérique éducatif, les initiatives publiques sont, à de rares exceptions près, volontaristes, encourageant le déploiement des technologies dans les appareils d'éducation et de formation nationaux. Appréhender les politiques nationales nécessite cependant de circonscrire d'emblée leurs différences ainsi que leurs rayons d'action par échelon éducatif.

Des spécificités organisationnelles

En règle générale, le gouvernement français tend à encadrer étroitement son système d'éducation et de formation par rapport à l'État fédéral états-unien, qui laisse davantage de latitude aux établissements. Aux niveaux primaire et secondaire, le système français suit une organisation centralisée et descendante, du ministère aux établissements via les académies et les collectivités. Son homologue états-unien s'organise quant à lui sur un mode décentralisé, laissant une grande latitude aux États fédérés et aux districts regroupant les établissements.

Si les universités sont plus autonomes que les établissements scolaires, elles sont également plus dépendantes du gouvernement en France en termes de budget de fonctionnement – la part de l'État français dans le budget moyen des universités montait à 81% en 2017 (Calviac, 2019), alors que celle des *State University Systems,* premiers pourvoyeurs des universités américaines, avoisinait les 40% en 2021 (National Center for Education Statistics, 2023a).

Dans la formation professionnelle, l'Etat français encadre l'organisation du marché et la certification des organismes de formation, quand les employeurs états-uniens conservent une autonomie plus large pour choisir la formation de leurs employés, parmi un éventail d'acteurs divers, des universités aux organismes de formation. Aussi divers soient-ils, les échelons décisionnels des pouvoirs publics français et états-uniens restent cependant volontaristes en matière de numérique éducatif, et obéissent à des principes communs.

Des velléités et des incohérences à actualiser

Leur volontarisme face à l'intégration d'artefacts technologiques procède historiquement de deux types d'objectifs. Le premier est endogène à l'éducation, lié à la transformation des pratiques d'apprentissage, d'enseignement et d'organisation de l'éducation qu'induirait la technologie.

Y répondent notamment les volontés gouvernementales de soutenir l'individualisation de l'apprentissage dans les systèmes éducatifs, via le soutien fédéral aux États-Unis aux machines à enseigner et aux laboratoires de langue dans les années 1950, ou le soutien du Ministère français au développement de la micro-informatique dans les salles de classe dans les années 1980.

Le second est exogène et lié à des considérations d'ordres économiques et politiques. Y répondent la volonté d'assurer la souveraineté du pays face à des risques géopolitiques – les États-Unis soutenant l'automatisation des machines à enseigner face à la crainte d'un déclassement après le lancement du satellite soviétique Spoutnik –, ou l'inscription de l'éducation dans une politique industrielle nationale – les commandes du Ministère français en ordinateurs dans les années 1980 bénéficiant au fleuron Thomson.

Les soutiens aux dispositifs algorithmiques dans les politiques françaises et états-uniennes méritent d'être questionnés selon cette grille d'analyse. En deçà de leurs modalités, quels sont leurs objectifs sur ces deux plans éducatifs et non-éducatifs ? L'identification fréquente, dans les discours politiques, de l'*adaptive learning* à l'intelligence artificielle – courant technologique aux enjeux géopolitiques importants – ne motive-t-elle pas une focalisation sur la focale économique et politique au détriment de celle éducative ?

En outre, les politiques publiques connaissent traditionnellement des incohérences. Elles manquent, pêle-mêle, de données fiables sur l'efficacité de telle initiative menée par le passé ou de tel média éducatif, ainsi que de ligne directrice lorsqu'elles soutiennent alternativement des filières aux intérêts divergents. En découlent l'impression qu'elles éparpillent les financements entre politiques de l'offre, visant à soutenir des firmes éducatives et à diffuser leurs innovations de façon verticale dans les systèmes éducatifs, et politiques de la demande, consistant à faire émerger des usages volontaires chez les publics éducatifs sur un mode horizontal (Fourgous, 2010). En France, deux tendances récentes nuancent néanmoins ce tableau : d'abord l'équilibrage des dépenses ministérielles entre équipements et usages, puis le transfert de nombreuses aides de l'amont (la production de contenus éducatifs) à l'aval (la distribution des médias), ce dont témoigne la généralisation du GAR (Boissière et Bruillard, 2021).

Les soutiens publics aux dispositifs adaptatifs interrogent l'actualité de ces tendances. Les acteurs publics poursuivent-ils l'équilibrage des dépenses en finançant significativement l'émergence d'usages volontaires chez les enseignants et les élèves ? S'ils entérinent l'affaiblissement de politiques de l'amont au profit de l'aval, ne risquent-ils pas de pénaliser la conception de contenus de remédiation variés, indispensables à l'édition de médias adaptatifs ?

Par ailleurs, l'investissement ministériel dans le GAR conditionne la distribution de firmes éditrices dans les écoles, collèges et lycées à l'exigence de protection des données personnelles. Les dispositifs adaptatifs en profiteraient-ils, alors que la collecte et le traitement de données variées constituent leur principe fondamental ? Une politique des données est-elle une menace pour leurs activités, ou au contraire une chance pour leur permettre de rallier les publics et pénétrer davantage d'institutions éducatives ? Ces divers questionnements afférents aux initiatives publiques des deux pays doivent apporter des réponses à la question de la constitution d'une socio-économie singulière et cohérente. Celle-ci est en outre conditionnée par les stratégies des firmes œuvrant dans les industries éducatives.

3. Les stratégies des firmes

Pour observer les stratégies des firmes, la théorie des industries culturelles nous offre un outil conceptuel pertinent : le modèle socio-économique. Comme on l'a vu en introduction, celui-ci dépeint des modes d'organisation idéal-typiques formalisés *a priori*, régissant les conditions de production des produits, la structuration de leurs processus de valorisation ainsi que les modes de rémunération des diverses parties prenantes. Atténuant l'incertitude sur la rentabilité des produits, ils véhiculent également des représentations sociales susceptibles d'amener des acteurs différents à sceller des collaborations. Notre hypothèse est que ces outils conceptuels peuvent éclairer les stratégies des firmes prenant part au développement de l'adaptativité, et contribuer à révéler *in fine* la formulation d'un projet paradigmatique.

Cinq modèles en question

Nous entendons le modèle en tant que « règle du jeu », conformément à la position détaillée par Pierre Mœglin (2007). Il s'agit ainsi d'un « idéal-type » au sens weberien, « tableau de pensée homogène » sommant de caractéristiques poussées à l'extrême. Celles-ci ne se retrouvent pas telles quelles dans les stratégies des firmes, mais c'est au contraire par rapport à ces données *a priori* que s'apprécient les mutations des stratégies, « par exemple lorsqu'une règle de départ est remplacée par une autre ou lorsque deux ou plusieurs règles sont pratiquées à la fois » (ibid.)[31].

Ces modèles « canoniques » sont au nombre de cinq. Les deux premiers sont le modèle « éditorial » et celui du « flux » (ou « flot »). Le premier a été formalisé en 1978 par Armel Huet et al. dans l'ouvrage *Capitalisme et industries culturelles* (Huet et al., 1978). Il désigne le mode d'organisation le plus ancien, ayant accompagné l'industrialisation de l'édition en Europe. Le paiement s'y fait en aval, individuellement, directement et à l'unité, pour un média reproduit de manière durable à l'image du livre, ou semi-durable comme un fichier audiovisuel. Il consacre la centralité de l'éditeur, collectant le fruit de la vente du produit auprès du consommateur, pour le redistribuer entre les différentes parties prenantes de la filière. En éducation, ce modèle est traditionnellement dominant, régissant notamment le secteur du manuel imprimé ou numérique.

Du point de vue des valeurs sociétales, il promeut un droit de propriété de l'acheteur sur le média. Ce droit fait écho à l'autonomie pratique des individus, notamment au plan économique, que Michel Freitag tient pour la caractéristique principale de la société bourgeoise moderne (Freitag, 1995). L'universalité de l'usage demeure cependant un objectif corollaire, lié à l'idéal kantien de l'espace public, le livre ayant prétention à être lu et approprié par tous.

Le modèle du « flux », décrit par Patrice Flichy (1980), régit quant à lui le fonctionnement de la télévision et de la radio. Les produits y sont de nature éphémère sans appropriation individuelle. L'objet de la transaction est l'audience, « vendue » en amont par les diffuseurs à des annonceurs. Les premiers sont donc centraux, agençant la grille des programmes et redistribuant les revenus aux autres acteurs de la filière. Il est plus récent en éducation, avec les incursions de la radio éducative dans les années 1950 puis de chaînes de télévision éducative dans les années 1960, mais aussi

[31] Cette acception diffère du modèle comme « règle de fonctionnement », ou « logique » chez des chercheurs comme Gaëtan Tremblay, considérant par là des « conditions structurantes qui définissent les paramètres du jeu des acteurs dans un secteur donné, à une époque historique donnée » (Tremblay, 1997, p.17). Ces modélisations *a posteriori* visent à caractériser au plus près l'évolution d'un média.

les fournitures et équipements vecteurs de publicité, des agendas et kits pédagogiques aux messageries avec bandeaux commerciaux.

À rebours du droit de propriété individuel, ce modèle qu'on nomme également du « flot » consacre un droit d'accès collectif et démocratique, garanti notamment par l'État et ses infrastructures de communication. Ses contempteurs en font un corollaire des *mass media* et de la standardisation de la culture, quand ses défenseurs louent à travers lui un modèle opposé aux valeurs normatives de l'institution scolaire et basé sur l'autodidaxie.

Trois autres modèles ont été proposés ces dernières décennies, alors que de nouveaux médias et des opportunités marchandes afférentes sont apparus. Le modèle du « club » régit le fonctionnement de filières dont l'accès aux produits se fait au forfait : l'objet de la transaction est un abonnement, payé en amont et donnant accès à un catalogue. L'acteur central est cette fois le gestionnaire du club qui effectue la liaison entre les abonnés et le catalogue. Outre la câblodistribution pour laquelle ce modèle a été formalisé (Tremblay et Lacroix, 1991), il concerne, en éducation, les revues par abonnement, les livres vendus en club, mais aussi les logiciels et environnements informatiques vendus via des licences annuelles par élève ou enseignant.

Il apparaît, *a priori*, comme un modèle privilégié pour la distribution des médias adaptatifs à destination des organisations éducatives. Selon un effet de réseau, plus nombreux sont les membres du club, plus l'offre est diversifiée et attrayante – principe de nature à concorder avec le « web social », et dans notre cas, les environnements adaptatifs de formation à distance. Au droit de propriété du modèle éditorial et au droit d'accès du modèle du flot, cette voie oppose le droit « de tirage », accordé à chaque usager pour la période prévue par l'adhésion. Elle promeut la construction d'une communauté éducative plus ou moins solidaire à travers le club.

Le modèle du « compteur » régit les services dont le financement est assuré par le consommateur à la durée et au volume. Théorisé par Bernard Miège et Patrick Pajon en 1990 pour la télématique (Miège et Pajon, 1990), il concerne la consultation de données par minitel ou la télévision à la demande. L'opérateur de télécommunication est l'acteur central qui contrôle l'accès, organise les interactions, procède à la facturation à la durée et redistribue les recettes aux fournisseurs de données. Ce modèle promeut un droit « d'usage » pour le consommateur, après le paiement, au *pro rata* de sa communication.

Rare aujourd'hui en éducation, il n'en demeure pas moins un référentiel puissant en termes de valeurs sociétales : il représente l'idéal d'un préceptorat à distance selon les désirs d'utilisation de l'apprenant. D'aucuns, surtout en France, perçoivent dans ce mode d'organisation une menace contre l'idéal d'égalité des chances, notamment en ce qu'il différencie la pédagogie et compense des handicaps individuels. S'y ajoute le soupçon d'un modèle de société « à péage » qui, inversement, étendrait des inégalités sociales. Aux États-Unis, la prévalence de la quête d'équité sur celle d'égalité, ainsi que l'institutionnalisation ancienne du tutorat rendent ces débats moins vifs.

Un dernier modèle est celui du « courtage informationnel ». Théorisé par Pierre Mœglin (1998), il a fait florès sur internet, en particulier sous l'impulsion des *Big Tech*, comme Amazon pour le commerce en ligne ou Google pour la recherche d'informations. Dans ce modèle, le courtier monnaye sa fonction d'intermédiaire entre usagers et producteurs via un financement au contact – à la commission, par référencement payant et vente de mots-clés à des annonceurs, ou par

la commercialisation d'informations acquises durant la transaction. Pour ajouter de la valeur à son service, le courtier enregistre les préférences des usagers et leur recommande des productions de façon personnalisée. En éducation, il est visible dans les modèles d'affaires des gestionnaires de centrales d'accès privé, qui mettent en lien producteurs de ressources informatiques et usagers, tels que les magasins d'applications pour la distribution d'applications éducatives directement au consommateur. Il apparaît *a priori* comme un modèle pertinent pour penser la commercialisation de médias adaptatifs directement auprès des consommateurs. Ce type de modèle est cependant susceptible de gagner également le circuit institutionnel devant l'essor des plateformes d'intermédiation évoquées plus haut, émanant d'acteurs privés et notamment des *Big Tech* comme Microsoft avec *Teams*.

Sur le plan des représentations sociales, le modèle conduit un idéal de l'assistance personnalisée. De même il véhicule une conception réticulaire et connexionniste de la société ainsi que l'a mis en exergue Pierre Mœglin (2005, pp. 233-234), y voyant un avatar socio-économique du « nouvel esprit du capitalisme » théorisé par Luc Boltanski et Ève Chiapello (1999) : il fait du réseau et de l'assistance personnalisée des valeurs universelles applicables au travail, aux relations privées voire aux politiques publiques. Ce modèle à l'actualité brûlante depuis l'avènement du *Big Data* fait écho à la personnalisation permise par les algorithmes d'apprentissage adaptatif, susceptibles d'orienter l'usager dans une abondance de contenus.

Les divers modèles ont des tendances hégémoniques, visant à s'imposer dans le paysage économique et réaliser leurs projets de société au détriment des autres. Néanmoins, en tant que « règles du jeu » abstraites et *a priori*, ils se retrouvent dans les stratégies des firmes éducatives non pas exclusivement mais de façon mêlée. Certains acteurs, notamment les grands groupes de l'édition scolaire ayant des filiales dans plusieurs industries, pratiquent plusieurs modèles selon leurs activités. Certains médias peuvent en outre obéir à plusieurs modèles socio-économiques selon les filières. À l'échelle d'une filière, de nouveaux modèles peuvent même émerger, lorsque le paysage technique, économique et social change, et qu'un nouvel acteur se place au centre du jeu et en définit les règles. Notre étude des stratégies des firmes devra déterminer le ou les modèles socio-économiques existants ou nouveaux qui accompagneraient l'industrialisation des médias adaptatifs.

Nous faisons donc l'hypothèse que les modèles socio-économiques issus de la théorie des industries culturelles, *a fortiori* ceux du club et du courtage, peuvent éclairer les stratégies des firmes prenant part à l'industrialisation des médias adaptatifs. Leur déploiement prolongerait la marchandisation croissante de l'éducation, notamment privative, que nous avons évoquée en introduction, symptôme de la convergence entre industries culturelles et éducatives. L'observation des modalités de distribution de l'adaptativité en France et aux États-Unis, entre commandes publiques et débouchés privés doit permettre d'éprouver cette hypothèse.

La thèse de la convergence

L'inscription des médias adaptatifs dans la convergence entre les industries éducatives et culturelles pose question. Il convient de détailler la manière dont ils réactualisent ou remettent en cause les arguments évoqués en introduction en faveur de cette thèse.

Outre la marchandisation, un facteur de convergence est la multiplication des innovations technologiques. Les acteurs des technologies, qu'il s'agisse d'équipements ou de logiciels,

diversifient et renouvellent en permanence leurs offres dans des marchés très concurrentiels et soumis à l'obsolescence. Ce renouvellement traverse puissamment les industries culturelles. Ainsi, dans le budget des ménages français, la part des dépenses liées aux équipements donnant accès aux contenus culturels excède depuis de nombreuses années celle des biens et services culturels (Direction des Études, de la Prospective, des Statistiques et de la Documentation, 2024, p.40).

Les industries éducatives sont touchées par cette même multiplication des innovations. Ainsi la marchandisation-privatisation se développe-t-elle à la faveur de l'augmentation des dépenses éducatives. Louis Porcher parle d'« économisation éducative » pour dépeindre le recours croissant à des outils et médias de plus en plus coûteux dans les salles de classe (Porcher, 1998, p.2), à l'image des outils et médias numériques (calculatrices connectées, tablettes, tableaux blancs interactifs).

Dans cette multiplication générale des innovations s'inscrivent les algorithmes d'apprentissage adaptatif, promus par les acteurs marchands comme une énième « disruption » technologique. Les modèles de *machine learning* profitent aussi de la diversification des supports et terminaux, les données d'entraînement s'en trouvant accrues. Dans une lecture plus pessimiste, l'avènement de nouvelles innovations, telles que les grands modèles de langage, pourrait aussi reléguer au second plan l'idéal d'adaptativité dans les stratégies des firmes.

La multiplication technologique sert depuis plusieurs décennies l'individualisation des médias – autre facteur de rapprochement entre industries culturelles et éducatives. Comme on l'a vu en introduction, les consommations individuelles culminent aujourd'hui sur les plateformes de streaming en raison de leurs algorithmes de recommandation. L'analogie entre ces derniers et les algorithmes d'apprentissage adaptatif, et plus largement entre individualisation des consommations culturelles et adaptativité pédagogique doit être questionnée. Concernant l'apprentissage, une activité aussi empreinte par des traditions artisanales et des enjeux politiques, affleurent rapidement des limites. La dichotomie issue des sciences de l'éducation entre individualisation, au sens d'une focalisation exclusive sur l'apprenant, et personnalisation, équivalent d'une adaptation plus souple appelant des pratiques collaboratives, peut apporter des nuances (Connac, 2021).

Une quatrième tendance rapprochant l'éducation et la culture est la financiarisation, ou le « développement de l'actionnariat financier, notamment transnational » selon Philippe Bouquillion (2012). Les politiques néolibérales des années 1980 et 1990 menées dans les pays occidentaux ont notamment déréglementé les marchés financiers de même que libéralisé des secteurs anciennement tenus par les États. Ces mouvements ont permis un déversement de capitaux financiers dans des marchés culturels et éducatifs. Des mouvements de concentration et de fusion-acquisitions, permis par la financiarisation, ont ainsi conduit à la constitution de groupes multimédias réunissant des filiales dans les industries de la culture, des télécommunications, de l'informatique voire de l'éducation.

Le développement de l'actionnariat est devenu un moyen de financement central jusque dans les industries éducatives, où les pressions concurrentielles se sont accrues. Par exemple, avec les coûts de recherche et développement induits par la multiplication des artefacts technologiques, les petites entreprises et les plus grands groupes doivent solliciter des financements externes accrus. L'émergence des médias adaptatifs est susceptible de favoriser un peu plus la financiarisation de l'éducation, selon des modalités restant à mettre en regard avec celles ayant cours dans les industries culturelles.

Une dernière tendance symptomatique de la convergence est l'augmentation des coûts marketing. Dans des marchés éducatifs très concurrentiels, les enjeux de *branding* sont cruciaux pour les industries de la culture comme de l'éducation. En éducation, pour attirer des masses critiques d'usager (conformément à un effet de réseau), les opérations de marketing reposant sur la segmentation de « publics-cibles » sont devenues monnaie courante, *a fortiori* chez des groupes concentrés. Les algorithmes d'apprentissage adaptatif, reposant en particulier sur le *machine learning,* supposent en outre la quête d'une vaste audience pour pouvoir être entraînés sur des jeux de données assez nombreuses. L'ampleur des budgets alloués par les firmes ainsi que les rhétoriques marchandes qu'elles mobilisent (autour de l'*adaptive learning* notamment) doivent être questionnées.

Les industries éducatives conservent toutefois des spécificités fortes, notamment sur le plan des enjeux politiques et des modalités d'intervention des pouvoirs publics. Elles devront être circonscrites pour jauger correctement l'éventuelle convergence industrielle qui sous-tend le déploiement de l'adaptativité.

Penser une filière intégrée

Pour appréhender une socio-économie transversale aux cinq industries éducatives, nous considérons une filière intégrée des médias adaptatifs. Malgré leur diversité technique (de systèmes de tutorat utilisables en classe et à la maison à des applications d'exercices gamifiés en ligne), on peut en effet considérer une chaîne de valeur générique. Celle-ci est concentrée sur trois étapes, dont deux en amont : la conception d'algorithmes d'apprentissage adaptatif et la production de contenus pédagogiques ; et une en aval : la distribution du média auprès d'un public éducatif en « *Software as a Service* » (SaaS). Ce mode de distribution est employé par les éditeurs mettant leurs logiciels à disposition des utilisateurs sur internet. Des hébergeurs tiers collectent et stockent les données des usagers via des serveurs de « *cloud computing* » à distance et sur internet, plutôt que sur un serveur local. La distribution en SaaS est aujourd'hui employée pour l'ensemble des ITS présents sur le marché.

Dans ces trois étapes interviennent cinq types de firmes, de statuts et d'activités très divers, impliquées sur une ou plusieurs industries éducatives :

- Les structures de production informatique de type « start-ups » ;
- Les acteurs plus anciens et concentrés du numérique éducatif ;
- Les éditeurs scolaires proposant des offres numériques ;
- Les gestionnaires de plateformes d'accès aux médias numériques ;
- Et des *Big Tech*, présents sur différents domaines, de l'hébergement *cloud* à la distribution dans des magasins d'applications, en passant par la conception de services algorithmiques voire de LMS.

Cette partie va donc s'appliquer à observer des tendances à la fédération ou à l'éclatement de ces acteurs, concourant à des stratégies rénovées à la lumière des modèles socio-économiques.

4. Des professionnalités dynamisées par l'*EdTech*

Une clé de voûte des industries éducatives est la constitution de professionnalités transverses, c'est-à-dire de compétences et d'intérêts partagés par des professionnels variés (haut-fonctionnaires, entrepreneurs, pédagogues, enseignants, *designers*, administrateurs d'établissements, investisseurs) et notamment issus de diverses filières industrielles, de l'édition scolaire à l'informatique pédagogique. Fréquentant les mêmes salons et s'informant dans les mêmes médias spécialisés, ils constituent une communauté technophile à même de soutenir et d'influencer en amont l'industrialisation des médias, tant dans les débats politiques que dans les stratégies des firmes.

Observer la morphologie et les orientations des réseaux

L'existence de ces professionnalités est d'abord visible dans la création de réseaux d'acteurs formulant des préconisations auprès des pouvoirs publics, mettant en relation leurs membres et soutenant de manière générale le développement des industries des technologies éducatives. Leurs orientations s'observent dans le choix de leurs dénominations. En France, une ombrelle ancienne est celle du « numérique éducatif », ou « numérique pour l'éducation », reprise notamment par l'Afinef (Association Française des Industriels du Numérique dans l'Education et la Formation). Une autre désignation plus récente est celle de l'*EdTech*. Contraction d'« *Educational Technology* », le terme a émergé aux États-Unis dans le courant des années 2010 puis s'est répandu dans le monde et notamment en France. Au réseau américain 1EdTech répondent dans l'Hexagone la fédération EdTech France ou le média EdTechActu. Cet anglicisme traduit l'imaginaire néo-libéral et innovant des start-ups, ainsi que le puissant tropisme innovateur qui accompagnent les technologies informatiques et en particulier l'intelligence artificielle. Il peut être interprété d'une part comme une référence aux autres secteurs industriels revisités par les start-ups et l'IA, comme la médecine (« *MedTech* ») ou la finance (« *FinTech* »), et d'autre part comme un écho lointain – volontaire ? – à la communauté antérieure de *l'Educational technology* aux États-Unis, ou « Technologie éducative » en France. Cet héritage est susceptible de jeter un éclairage sur les visées et le poids des communautés actuelles.

La communauté de la Technologie éducative s'est formée, comme on l'a vu en première partie, dès les années 1960 autour de pédagogues, chercheurs, enseignants et entrepreneurs de divers pays, partageant un attrait commun pour les médias éducatifs de l'époque. L'alliance était hétéroclite, tant leurs idéaux pouvaient varier, entre des éditeurs tenants du behaviorisme et des pédagogues progressistes, des militants de la gratuité et des chefs d'entreprise soucieux de rentabilité. Guy Berger interprétait l'origine de leur réunion dans la volonté de bousculer des systèmes éducatifs jugés sclérosés (Berger, 1982).

La même motivation réunit-elle les acteurs pluriels des communautés du numérique éducatif et de l'*EdTech* ? À première vue, l'intelligence artificielle relance puissamment ce tropisme innovateur, à travers l'idéal d'une « disruption » de l'École. Pour éprouver cette hypothèse, il convient néanmoins de questionner la diversité sociologique et professionnelle des membres des communautés actuelles, pour saisir leurs intérêts convergents.

Avec l'*EdTech* et l'IA, de nouveaux acteurs sont en outre arrivés : dirigeants de start-ups, cadres de multinationales des technologies… Ceux-ci, obéissant à des agendas propres, sont-ils de nature à agir

comme des intermédiaires et des agents du rapprochement entre des industries différentes, ou au contraire comme des éléments perturbateurs ?

La Technologie éducative a décliné à partir des années 1970, critiquée pour sa prétention à rendre compte de l'ensemble des phénomènes éducatifs à partir de principes simplistes. La disqualification du behaviorisme dans la sphère académique et l'émergence des théories constructivistes avec les développements des ordinateurs ont eu raison de cette communauté. Ces clivages épistémologiques sont-ils aujourd'hui révolus parmi les acteurs technophiles de l'éducation ? On a observé en première partie la survivance de dissensions entre chercheurs sur la place à donner à la technique, relancées par le *Big Data* et ses penchants behavioristes. Ces débats imprègnent-ils les communautés actuelles ?

Des compétences communes à questionner

La communauté du « numérique éducatif » a ensuite profité d'un facteur d'homogénéisation d'ordre technique que n'avait pas la Technologie éducative : l'avènement hégémonique de l'informatique et du « numérique ». Là où le courant précédent s'emparait de dispositifs hétéroclites (machines à enseigner, techniques audiovisuelles, ordinateurs), les communautés actuelles sont tout entières tournées vers les dispositifs informatiques. Ce parti-pris a normalisé, parmi les professionnels enseignants, administratifs du privé ou du public, dirigeants ou cadres d'entreprises, des compétences intersectorielles et communicationnelles basées sur des outils comme le Web. Toutefois, ces technologies ont aussi entraîné la constitution de nouvelles filières comme celle des systèmes numériques de gestion et d'information, et dès lors la spécialisation de nombreux acteurs dans des secteurs aux visées divergentes.

Entre spécialisation et polyvalence, des intérêts ambivalents sont susceptibles de s'être créés entre des professions (notamment entre professionnels de l'édition et de l'informatique pédagogique). La popularité récente de l'intelligence artificielle a-t-elle contribué à resserrer les liens entre ces professionnels ? L'émergence de nouvelles compétences avec l'intelligence artificielle et le *machine learning*, telles que celles liées à la *data science* et à la protection des données, pourraient créer des espaces de dialogue d'un nouveau genre, comme des lieux de débats et de dissensions. Il convient ainsi de questionner les motivations de chaque acteur industriel à rejoindre ces communautés, ainsi que l'homogénéité de celles-ci, à travers les concordes et les clivages émergeant collectivement.

Dans les chapitres suivants, nous étudions ces trois niveaux interdépendants que sont les politiques publiques, les stratégies des firmes et les professionnalités, à l'aune des cas états-uniens puis français. Suivant la méthode contrastive expliquée en introduction, il s'agit de délimiter l'influence des particularismes nationaux et de ceux liés aux niveaux d'enseignement, pour mettre finalement en exergue des permanences socio-économiques dans l'industrialisation des médias adaptatifs.

Chapitre V. Des industries éducatives revisitées aux États-Unis

En préambule de ce chapitre, il est d'abord nécessaire de cerner les modes d'organisation de l'appareil éducatif et de formation états-unien, pour comprendre le rayon d'action des trois principaux échelons publics : État fédéral, État fédéré, districts. Ainsi, nous comptons interpréter les discours à portée paradigmatique des décideurs étatiques, selon leurs capacités à encadrer et réguler le déploiement des dispositifs algorithmiques. Concernant les firmes, nous devons observer les tendances à la fédération ou à l'éclatement des principaux acteurs des médias adaptatifs, ainsi que les modèles socio-économiques que suivent leurs stratégies. Au niveau des professionnalités, nous étudions la manière dont des réseaux d'acteurs transversaux se forment sous la bannière *EdTech*, et analysons leurs agendas pédagogiques et stratégiques.

1. L'appel des acteurs publics au *personalized learning*

Les modes d'organisation de l'appareil éducatif et de formation

Le cas états-unien nécessite de renoncer au préalable à la métaphore de « système » éducatif et de formation, trompeuse comme le remarquait en 2004 Laurence Émile-Besse, alors attachée culturelle à l'ambassade de France à Washington : « parler du système éducatif américain est une sorte de contradiction […] tant le terme « système » suppose une cohérence et une unité qui s'opposent précisément à ce qu'est l'éducation américaine » (Émile-Besse, 2004). Les institutions éducatives sont en effet plurielles : écoles publiques, *charters schools* privées ou foyers adeptes du *home schooling* sont autant de lieux d'enseignement primaire et secondaire. Les établissements du supérieur peuvent quant à eux être des universités de statuts très différents, des *college* privés, ou des *community colleges* publics.

Par ailleurs, les lieux de décision sont divers et les circuits de décision décentralisés. L'État fédéral agit d'abord au niveau législatif. Les lois votées au Congrès donnent des orientations stratégiques à tous les échelons éducatifs du pays. L'action fédérale est ensuite exécutive, portée par le *Department of Education,* et consiste à établir, à partir de son budget annuel voté au Congrès, « des politiques sur l'aide financière fédérale à l'éducation, et distribuer et surveiller ces fonds »[32] (U.S. Department of Education, 2024a). Cette distribution s'opère sous la forme de subventions, visant principalement à pallier les disparités entre établissements et districts scolaires, et à amortir la baisse des financements des États dans les périodes de récession[33]. Le total des subventions fédérales reste minoritaire : celles-ci représentaient 11 % du budget total des écoles publiques en 2021 (Peter G. Peterson Foundation, 2023). L'État fédéral participe également au marché de la formation pour les individus en reconversion ou au chômage, selon les besoins économiques locaux (« *workforce development* »), en allouant des fonds aux États fédérés. Le marché de la formation continue (« *training and development* ») appartient quant à lui aux employeurs et aux individus.

[32] *« ED's 4,400 employees and $68 billion budget are dedicated to: Establishing policies on federal financial aid for education, and distributing as well as monitoring those funds ».*

[33] Le dernier exemple en date est la crise économique ayant suivi la pandémie de Covid-19. En 2020 et 2021, le Congrès a adopté successivement trois lois de subventions d'urgence à destination des services publics : le *Coronavirus Aid, Relief, and Economic Security Act (CARES Act),* le *Coronavirus Response and Relief Supplemental Appropriations Act (CRRSA Act),* et l'*American Rescue Plan Act (ARP Act)*.

Au niveau de chacun des cinquante États fédérés, la *State Education Agency* (SEA) est en charge de la politique locale en termes d'éducation primaire et secondaire. Dirigée par une assemblée, le *board of education*, cette juridiction doit d'une part mettre en œuvre les lois décidées au niveau fédéral, et d'autre part définir la politique scolaire du territoire, notamment en termes de définition des programmes et standards pédagogiques. Sur la prescription d'outils pédagogiques se distinguent les *adoption states* des *open territory states*. Les premiers chargent leurs SEA de définir des listes d'éditeurs agréés, parmi lesquels les districts en charge de l'équipement d'un réseau d'établissements devront passer commande pour toucher des subventions de la part de l'État. Les seconds laissent ces derniers libres d'acheter les médias de leur choix (Koehler Zeringue et Allen Fuller, 2007). La carte ci-dessous (Figure 8.) donne à voir les vingt *adoption states* et les trente *open territory states* actuels.



Figure 8. Répartition des vingt adoption states (en vert) et des trente open territory states (en rouge). Source : Koehler Zeringue et Allen Fuller, 2007

La SEA est en outre le premier financeur des dépenses générales de fonctionnement des établissements publics, incluant l'achat d'outils et de médias éducatifs. Dans l'enseignement primaire et secondaire, les SEA contribuaient à hauteur de 45 % du budget des écoles publiques en 2021 (Peter G. Peterson Foundation, 2023). Dans le supérieur, les États soutiennent les établissements publics du territoire via des allocations directes, plus importantes pour les « universités d'État » appartenant au *State University System* (SUS), ou des appels à subventions. Ces subsides représentent la majeure partie des financements des établissements – à hauteur de 40% en moyenne de leur budget en 2021 (National Center for Education Statistics, 2023a). Les États fédérés financent les programmes de « *workforce development* » et notamment la *Career and Technical Education* (CTE), filières techniques proposées dans les *high schools* et les *community colleges* aux lycéens, étudiants et adultes en reconversion.

Viennent enfin les institutions éducatives publiques (*grade, middle, high schools*, universités, *community* et *vocational colleges*) qui s'équipent en outils et médias éducatifs. Dans l'enseignement primaire et secondaire, les écoles d'un même territoire sont rattachées à un district scolaire. Il en existait plus de 13 000 en 2021 (National Center for Education Statistics, 2022). Cet échelon est en charge de sélectionner des éditeurs et de négocier, au nom des établissements sous son autorité, des prix de gros pour des licences. Dans l'enseignement supérieur et les programmes de *Career and Technical Education* (CTE), ce sont les universités et les *vocational colleges* qui gèrent leurs budgets et passent commande en outils éducatifs.

Malgré la forte décentralisation du système d'éducation et de formation, les pouvoirs publics états-uniens jouent un rôle structurant dans l'ensemble des industries éducatives. Cette influence est la plus forte sur les filières où ils constituent les principaux pourvoyeurs de fonds des établissements acheteurs (l'édition de manuels, l'informatique pédagogique, le *e-learning* ou les systèmes numériques de gestion et d'information). Elle est plus indirecte sur les industries dont les produits se vendent directement aux consommateurs ou aux entreprises (la formation à distance et les produits de l'éducation non-formelle), mais tout de même déterminante en raison de leur pouvoir législatif. Par exemple, les réglementations en termes de protection des données personnelles pèsent sur les activités des éditeurs d'applications. Témoin, aux États-Unis, le *Children's Online Privacy Protection Act* (COPPA), encadrant depuis 1998 la collecte en ligne de renseignements personnels sur des enfants de moins de treize ans.

Les politiques publiques sont ainsi de nature à structurer ces industries, tant dans les stratégies des firmes que dans les usages des apprenants. Au niveau fédéral comme au niveau des gouvernements fédérés, l'implémentation de technologies éducatives fait l'objet d'un consensus entre Républicains et Démocrates. Les deux principaux partis soutiennent leur déploiement dans une quête d'efficience héritée de l'*Evidence-Based Education*. Les clivages partisans tournent principalement autour de deux types de questionnements. Les premiers sont idéologiques, les Républicains appelant à bannir des programmes scolaires les enseignements ayant trait aux inégalités ethniques ou de genre, comme en Floride. Les seconds sont organisationnels, les Démocrates soutenant davantage de redistribution par l'État fédéral vers les districts et les publics défavorisés, quand le camp républicain appelle à accroître l'autonomie financière des États en supprimant le *Department of Education* – comme le révèlent les débats préalables à la campagne présidentielle de 2024 (Alfonseca, 2024).

Le déploiement de technologies éducatives étant un aspect consensuel dans le champ politique états-unien, quel projet éducatif les pouvoirs publics tendent-ils à soutenir à travers elles ? Des préconisations du *Department of Education* aux réformes votées au Congrès, en passant par les directives des SEA, les institutions tendent depuis plusieurs décennies à organiser les institutions suivant une quête industrialiste d'efficience, grâce à une individualisation de l'éducation. Au tournant des années 2010, un idéal plaçant la personnalisation au cœur des pratiques a pris l'ascendant dans les discours et textes de lois : le *personalized learning*. De Washington aux gouvernements fédérés, nous retraçons ci-dessous les facteurs ayant conduit à l'émergence de ce paradigme.

La construction d'un projet officiel

Le prisme politique est d'abord essentiel pour appréhender les réformes éducatives décidées à l'échelon fédéral, comme l'explique Pierre Mœglin : « Flagrante est […] l'influence du politique dans la persévérance du projet industrialiste à travers la succession des lois et réformes qui, depuis celles inspirées par le *Social Efficiency Movement* dans les premières décennies du XX^{e} siècle, mènent à NCLB [*No Child Left Behind*, loi fédérale votée en 2001] […] » (Mœglin, 2016, p.264). Le *Social Efficiency Movement* constitue en effet la matrice idéologique du pouvoir législatif au XX^{e} siècle, promouvant une approche tayloriste des services publics. Appliquée à l'éducation, celle-ci considère l'imputabilité de l'échec scolaire non seulement aux individus et aux familles, mais aussi à l'institution scolaire (conformément au principe de l'*accountability*). Une convergence s'opère alors avec le behaviorisme dominant dans la pédagogie d'après-guerre, visant à transformer les capacités initiales de l'enfant lors de son entrée à l'école (« *inputs* ») en comportements souhaités à sa sortie (« *outputs* »). Cette doctrine suppose la définition de comportements désirés sous la forme

d'un « curriculum », centré sur des compétences formalisées selon les attentes de la société, et non par le biais de programmes composés de savoirs. Le behaviorisme offre ainsi un cadre utilitaire aux pouvoirs publics soucieux de répondre à des problèmes de politique intérieure ou extérieure. Les technologies à usage individuel conçues suivant les préceptes behavioristes de l'enseignement programmé constituent alors un instrument important de la politique fédérale. Les principales lois éducatives du XX^{e} siècle s'inscrivent dans une veine industrialiste favorisant une politique de l'offre, en attribuant des subventions aux éditeurs de ces technologies éducatives.

L'émergence de la personnalisation comme idéal politique d'efficacité

En 1958, la ratification du *National Defense of Education Act* (NDEA) doit ainsi apporter des réponses politiques au lancement par l'Union Soviétique du satellite Spoutnik et à la crainte d'un déclassement des États-Unis dans la Guerre Froide. Il vise notamment à accroître la montée en compétences des élèves américains, en finançant des tests annuels dans les écoles et des expérimentations autour de « machines à enseigner ». Le *Title VII* prévoit des subventions aux éditeurs de ces outils d'enseignement programmé, dont B.F. Skinner profite pour ses machines, comme on l'a vu en première partie (Casas, 2002).

En 1965, l'argument de politique intérieure prévaut lorsqu'en pleine lutte pour les droits civiques est voté l'*Elementary and Secondary Education Act* (ESEA). Pierre angulaire du projet de « *Great Society »* du président L.B. Johnson, cette loi fixe encore aujourd'hui le cadre d'une intervention du pouvoir fédéral dans l'éducation primaire et secondaire, consistant, par des subventions ciblées, à lutter contre les inégalités sociales et communautaires à l'école. Cette loi s'inscrit dans le sillage de l'*affirmative action*, ou « discrimination positive » émergeant alors dans les politiques fédérales pour aider les minorités défavorisées, en vue d'égaliser les chances d'accès aux emplois notamment.

Renouvelant cette loi, le *No Child Left Behind Act* (NCLB) est voté en 2001 avec l'ambition qu'en 2014, l'ensemble des élèves américains soient compétents (« *proficient* ») en lecture, écriture, et calcul. Pour cela, les États se trouvent obligés de réaliser des tests de compétences chaque année, suivant des modalités et standards à leur discrétion, pour tous les élèves de huit à quatorze ans. Les établissements échouant à démontrer une progression annuelle (*Adequate Yearly Progress*, ou AYP) sont passibles de sanctions, pouvant aller jusqu'à leur fermeture. Ainsi le principe d'*accountability*, pivot du *Social Efficiency Movement*, est renforcé.

Pour atteindre cet objectif très ambitieux de 100% des élèves américains maîtrisant les compétences « fondamentales » en 2014 – alors qu'ils sont à peu près 30% en 2002 –, les technologies sont à nouveau appréhendées comme des solutions. Premièrement, l'article *Enhancing Education Through Technology* (EETT) vise à « fournir une assistance aux États et aux localités pour la mise en œuvre et le soutien d'un système complet qui utilise efficacement la technologie dans les écoles primaires et secondaires, afin d'améliorer la réussite scolaire des élèves » (Lemke et al., 2004, p.4)[34].

[34] C'est le premier objectif de la *Part D* du *NLCB Act : «* Les objectifs de cette partie sont (…) (1) de fournir une assistance aux États et aux localités pour la mise en œuvre et le soutien d'un système complet qui utilise efficacement la technologie dans les écoles primaires et secondaires pour améliorer la réussite scolaire des élèves »/*« The purposes of this part are (…) (1) To provide assistance to States and localities for the implementation and support of a comprehensive system that effectively uses technology in elementary schools and secondary schools to improve student academic achievement »* (Lemke, C. et al., 2004, p.4).

Deuxièmement, l'État fédéral s'efforce de structurer les recherches en sciences éducatives, liées en particulier aux technologies numériques. Il s'agit de faire émerger des « bonnes pratiques » bénéfiques à l'échelle nationale. Il reprend ainsi les principes de l'« *Evidence-Based Education* » (EBE), prisme politique évoqué en première partie, selon lequel les pratiques éducatives doivent être basées sur des preuves scientifiques obtenues par essais randomisés. Les technologies éducatives et les sciences de l'ingénierie sont en particulier visées par les investissements de la *National Science Foundation,* l'agence fédérale finançant la recherche.

Les financements se concentrent ainsi sur la création de pôles de recherche mêlant universités et industries : les *Science of Learning Centers* (National Science Foundation, 2004)[35]. Les universités de Carnegie Mellon et de Pittsburgh sont notamment dotées de 25 millions de dollars en 2004 pour créer ensemble le *Pittsburgh Science of Learning Center* (PSLC). Celui-ci vise à conduire des recherches sur les environnements informatiques pour l'apprentissage humain. Le site de Carnegie Mellon, relatant la création du PSLC et de son laboratoire LearnLab, la justifie par les supposées limites de la recherche en sciences de l'éducation de l'époque (Potts et al., 2004) :

> « LearnLab abordera ce qui constitue depuis longtemps un dilemme pour les chercheurs en éducation : les expériences menées dans les conditions artificielles du laboratoire produisent des résultats et des innovations qui ne sont pas largement transférables aux écoles, tandis que les études menées en classe ont tendance à être moins rigoureusement contrôlées, et ne fournissent donc pas de résultats suffisamment fiables pour être utilisés par d'autres. Ainsi, la recherche en éducation n'a pas fait preuve du type de rigueur que l'on retrouve, par exemple, dans la recherche médicale – un problème que les décideurs fédéraux s'empressent de corriger. En vertu de la loi fédérale *No Child Left Behind*, les écoles doivent démontrer leurs progrès annuels au moyen de tests standardisés, obligeant les enseignants et les administrateurs scolaires à prendre des décisions davantage fondées sur les données »[36].

Le PSLC s'inscrit dans la volonté fédérale de filtrer et structurer les recherches en sciences de l'éducation pour répondre à l'objectif politique de l'époque : la hausse des résultats éducatifs des écoles. La référence à la recherche médicale est révélatrice d'une stratégie de scientifisation de l'éducation par les données. Au LearnLab émergera ainsi la discipline de l'*Educational Data Mining*, sous l'impulsion de Ryan Baker, alors post-doctorant dans ce laboratoire. Les centres lauréats contribuent plus largement à la promotion de technologies d'apprentissage « personnalisées », dont les politiques fédérales vont progressivement devenir prescriptrices.

Le *Department* soutient également la recherche menée par les firmes privées via le *Small Business Innovation Research program*. Administré par l'*Institute of Education Sciences*, il s'articule en deux phases : la première alloue 250 000 dollars sur huit mois, avec l'objectif que le projet, à cette échéance, ait « (1) développé un prototype fonctionnel d'un nouveau produit technologique éducatif et (2) mené des recherches pour éprouver la facilité d'utilisation et la promesse du prototype

[35] Trois centres furent créés en 2004 : le *Center of Excellence for Learning in Education, Science and Technology* (CELEST) à l'université de Boston, le *Learning in Informal and Formal Environments* (LIFE) à l'université de Washington, et le PSLC. Trois autres furent lancés en 2006 : le *Spa3al Intelligence and Learning Center* (SILC) à l'université Temple, le *Temporal Dynamics of Learning Center* (TDLC) à l'université de Californie-San Diego, et le *Visual Language and Learning Center* (VL2) à l'université Gallaudet.

[36] *« LearnLab will address what has long been a dilemma for education researchers: Experiments conducted in the artificial confines of the laboratory produce results and innovations that are not broadly transferable to schools, while studies conducted in classrooms have tended to be less rigorously controlled and thus do not provide sufficiently trustworthy results that can be used by others. Thus, education research has lacked the kind of rigor that can be found, for example, in medical research, a problem federal policymakers are eager to correct. Under the federal No Child Left Behind Act, schools must demonstrate yearly progress through standardized tests, requiring teachers and school administrators to make more data-driven decisions ».*

de conduire aux résultats escomptés »[37] (Institute of Education Sciences, 2024). La seconde phase poursuit la première en dotant les entreprises sélectionnées d'un million de dollars supplémentaire pour deux ans, et en ajoutant trois autres visées : « (3) évaluer la promesse d'atteindre les résultats escomptés lorsque le produit est utilisé dans des salles de classe ou des écoles ; (4) élaborer un plan de commercialisation pour la distribution et la durabilité du produit ; (5) planifier des études d'efficacité ultérieures »[38] (Institute of Education Sciences, 2024).

En parallèle, les mesures de NCLB suscitent de nombreuses critiques dans la société états-unienne, révélant une attente sociale en termes de personnalisation pédagogique. Principale pomme de discorde : la prétendue standardisation de l'enseignement et de l'évaluation qui en découlerait. Derrière ce grief, deux dangers *a priori* antagonistes sont dénoncés : l'exclusion des enfants en difficulté scolaire, et le nivellement par le bas des élèves les plus en réussite.

La première crainte, exprimée par le camp progressiste, porte sur la place centrale du test, brocardé en tant que moteur bureaucratique d'exclusion. Des chercheurs du *National Center for Fair and Open Testing* pointent du doigt, comme cause d'un supposé appauvrissement de l'enseignement, le fait que, suivant NCLB, « renforcer les scores individuels aux tests standardisés devrait être le but principal de l'école. Cette supposition conduit à un enseignement standardisé, visant principalement la préparation des tests, ce qui contrevient aux efforts pour fournir à tous les élèves une éducation de qualité »[39] (Neill et al, 2004, ES-1).

À l'inverse, le second reproche, émanant des tenants de la *gifted education*, dénonce la focalisation exclusive du NCLB sur les élèves en difficulté au détriment des plus « talentueux ». À défaut de récompenser graduellement les États selon leurs performances, le NLCB pêche selon eux à créer un nivellement par le haut des établissements (Beisser, 2008) – un danger dans un contexte géopolitique de compétition technologique entre les États.

Ces deux reproches, en apparence antagonistes, attestent tous deux d'une attente sociale de personnalisation de l'enseignement. La standardisation qu'ils dénoncent tient en bonne partie aux méthodes d'évaluation statique des élèves. En effet, les tests administrés via le *National Assessment of Educational Progress* (NAEP) sont des questionnaires à choix multiples portant sur une batterie de questions communes à tous les élèves d'un même État. Nulle adaptation des questions selon le niveau de l'enfant n'est prévue, qui permettrait de concilier aspirations progressistes et élitistes. Les systèmes adaptatifs développés dans les *Science of Learning Centers* ouvrent au contraire des pistes pour y parvenir, dans le sillage du *Computer Adaptive Testing* observé dans le premier chapitre.

À ces critiques s'ajoutent les résultats décevants obtenus par les écoles du pays aux tests annuels, inférieurs aux projections attendues pour atteindre l'objectif de 2014[40]. Dès lors, un nouvel idéal pédagogique, le *personalized learning,* émerge comme une réponse dans les politiques fédérales du

[37] *« A funded Phase I ED/IES SBIR project would be successful if at the end of a 8-month period the team: (1) had developed a functioning prototype of a new education technology product, and (2) conducted research to determine the usability, initial feasibility, and promise of the prototype to lead to the intended outcomes."*

[38] *« (3) evaluated the promise of the product for achieving the intended outcomes when used in classrooms or schools; (4) developed a commercialization plan for the distribution and sustainability of the product, and (5) developed plans to conduct research to evaluate the efficacy of the product to lead to the intended outcomes ».*

[39] « *Two false assumptions undergird NCLB*: (1) Boosting standardized test scores should be the primary goal of schools. *This assumption leads to one-size-fits-all teaching aimed primarily at test preparation, and it works againstefforts to give all children a high-quality education (…) »*

[40] En 2015, Denis Meuret, comparant les résultats moyens entre 2002 et 2012, dresse un bilan nuancé : « les résultats de NCLB en termes d'efficacité sont décevants, sans être négatifs. Depuis 2003, au NAEP […], les résultats des élèves se sont améliorés, mais sans que cette augmentation soit plus rapide qu'avant la loi. D'après PISA, la performance moyenne des élèves de 15 ans n'a pas évolué de façon significative » (Meuret, 2015).

début des années 2010. Saisissante est par exemple l'évolution du programme de subvention *Race to The Top*, qui accorde dès 2009 une enveloppe de 4,35 milliards de dollars aux États poursuivant la refonte de leurs systèmes éducatifs suivant les préceptes de NCLB. À sa publication, l'appel ne comporte aucune attente en termes de *personalization.* Trois ans plus tard, sa déclinaison à destination des districts vise explicitement cet objectif :

> « Le programme *Race to the Top-District* soutient les améliorations audacieuses et dirigées localement vers l'apprentissage et l'enseignement, qui amélioreront directement la réussite des étudiants. Les lauréats du programme serviront de laboratoires d'innovations, avançant de nouvelles manières d'éduquer nos élèves via une approche personnalisée »[41] (U.S. Department of Education, 2024b).

En 2015, au programme d'investissement de NCLB, l'*Enhancing Education Through Technology* (EETT), est ajouté un objectif d'équipement des écoles en technologies éducatives « personnalisées », en vue d'« assurer que tous les élèves aient accès à des expériences d'apprentissage digitales personnalisées, rigoureuses et engageantes »[42] (Enhancing Education Through Technology Act, 2015).

L'*Every Student Succeeds Act* (ESSA), « *reauthorization* » de l'ESEA ayant succédé à NCLB, entérine ce tournant dans le financement fédéral des acteurs territoriaux. Point d'orgue de la mandature de Barack Obama, cette loi encourage explicitement la mise en place de pratiques de *personalized learning* dans tous les États et districts subventionnés sur la base de « preuves d'efficacité ». Pour permettre aux districts de mieux juger l'efficacité des logiciels qu'ils achètent, l'ESSA fournit une grille d'évaluation comportant quatre niveaux de preuve (« *four tiers of evidence* »)[43]. Elle comprend en outre un programme d'investissement, le *Student Support and Academic Enrichment Program* (SSAE), dont le volet « *Activities to Support the Effective Use of Technology »* finance les districts évaluant en amont les besoins des publics scolaires (élèves, enseignants, *school leaders*) en termes de *personalized learning :*

> *«* La section 4106(d) de l'ESEA requière qu'une agence recevant une allocation du programme SSAE d'au moins 30 000$ réalise une évaluation globale des besoins avant la réception de l'allocation, […] pour examiner ses besoins d'amélioration en termes : […] d'accès à des expériences d'apprentissage personnalisées soutenues par la technologie, ainsi que de formation professionnelle pour l'usage efficace des données et de la technologie *»*[44] (U.S. Department of Education, Office of Elementary and Secondary Education*,* 2016, p.11).

Le *Department* décerne également des *Education and Research Innovation (EIR) grants*, subventions sur cinq ans visant à développer et industrialiser des innovations entrepreneuriales, testées sur le terrain et fondées sur des « preuves d'efficacité ». Le montant prévu par projet en 2023 est fixé à 4 millions de dollars (U.S. Department of Education, 2023).

Ainsi l'État fédéral concilie politique de l'offre, finançant les entreprises selon l'agenda du *personalized learning*, et une politique de la demande, pour inciter les acteurs territoriaux

---

[41] *« The Race to the Top-District program supports bold, locally directed improvements in learning and teaching that will directly improve student achievement. Program grantees serve as innovation laboratories, advancing new ways to educate our students through a personalized approach »*

[42]« *The purposes of this part are (…) (2) to ensure all students have access to personalized, rigorous, and engaging digital learning experiences* » (*Enhancing Education Through Technology Act*, 2015).

[43] Les quatre types de preuves sont, par ordre décroissant de fiabilité : « Forte » (« *Strong* »), « Modérée » (« *Moderate* »), « Prometteuse » (« *Promising* »), et « Démontre une logique » (« *Demonstrates a Rationale* »).

[44] « *ESEA section 4106(d) requires that an LEA receiving an SSAE program allocation of at least $30,000 must conduct a comprehensive needs assessment prior to receiving its allocation, and subsequent needs assessments at least once every three years, to examine its needs for improvement of : (…) access to personalized learning experiences supported by technology and professional development for the effective use of data and technology ».*

à développer leurs usages de ces technologies. Cependant la quête industrialiste d'efficience, héritée du *Social Efficiency Movement* et revisitée par l'*Evidence-Based Education*, reste intacte. Seuls les moyens changent. Ainsi, lors de la cérémonie de signature de l'ESSA, Barack Obama (2015) réaffirme les objectifs bureaucratiques du NCLB, dont l'importance des évaluations annuelles, et pointe surtout du doigt les erreurs commises dans leur mise en application :

> « Les objectifs de *No Child Left Behind*, le prédécesseur de cette loi, étaient les bons : des standards élevés, l'imputation directe, le comblement des inégalités de réussite, l'assurance que tous les élèves apprennent, et non une partie. Mais en pratique, cela s'est souvent avéré limité. Étaient ignorés les besoins spécifiques de chaque communauté. Cela conduisait à trop d'évaluations durant le temps de classe. Cela conduisait les écoles et les districts à des réformes à l'emporte-pièce, qui ne produisaient pas toujours le type de résultats que nous attendions »[45] (Obama, 2015).

Derrière la réaffirmation de l'objectif du NCLB de « ne laisser aucun enfant de côté », la critique du Président Obama porte sur les moyens standardisés de le mettre en œuvre. L'ESSA doit combler cette lacune : les évaluations doivent être plus efficientes, moins nombreuses et adaptées aux communautés. La synthèse entre l'industrialisme traditionnel des politiques éducatives et le *personalized learning* est désormais officielle.

Le contexte politique est devenu propice au déploiement de techniques pédagogiques et évaluatives personnalisées. Parmi celles-ci, les algorithmes de fouille de données et l'intelligence artificielle de type connexionniste, ouvrant dans les années 2010 des perspectives techniques importantes, sont promus en réponse à deux nouveaux problèmes politiques. Premièrement, dans un contexte géopolitique de course à l'IA, l'administration prend conscience d'un risque de déclassement technologique du pays, en particulier face à la Chine (National Security Commission on Artificial Intelligence, 2021). Deuxièmement, les résultats des élèves américains aux évaluations nationales et internationales régressent. Dans les années 2010, et *a fortiori* suite à la pandémie de Covid-19 et la fermeture des écoles américaines, les États-Unis ont récolté des résultats jugés alarmants. En 2022, les notes moyennes obtenues par les élèves au NAEP sont au plus bas depuis 1994. L'enquête PISA menée la même année mesure en outre une forte baisse du niveau moyen des élèves américains en mathématiques (National Center for Education Statistics, 2023b).

L'État fédéral réagit alors de deux façons. La première est une production de documents normatifs, définissant officiellement le *personalized learning* et préconisant le déploiement de ce projet techno-éducatif dans les écoles et universités du pays. La seconde est une stratégie de structuration des recherches en matière d'IA éducative selon cet agenda.

La production de normes autour d'un projet techno-éducatif

Avant d'aborder cette première stratégie fédérale, il est nécessaire de noter que le *personalized learning,* thématique ancienne des sciences éducatives américaines, est devenu au fil du temps une notion aux contours flous, supposant des moyens parfois contradictoires selon les interprétations. Ses thuriféraires revendiquent souvent l'héritage de John Dewey, qui prônait comme on l'a vu un apprentissage individuel par l'« expérience » (Shemshack et Spector, 2020). Dans cette lignée s'inscrivent les experts appelant au développement des interactions entre pairs (travaux en groupes

[45] « *The goals of No Child Left Behind, the predecessor of this law, were the right ones: High standards. Accountability. Closing the achievement gap. Making sure that every child was learning, not just some. But in practice, it often fell short. It didn't always consider the specific needs of each community. It led to too much testing during classroom time. It often forced schools and school districts into cookie-cutter reforms that didn't always produce the kinds of results that we wanted to see* ».

de niveaux, tutorat) ou avec le professeur (classe inversée). Toutefois certains professionnels, plus proches du behaviorisme, prônent en guise de *personalized learning* l'usage d'outils informatiques « en vase clos ».

Le *Department of Education* et son bureau spécialisé dans les technologies éducatives, l'*Office of Educational Technology* (OET), ont expliqué leur position sur le sujet en 2017 dans le *National Educational Technology Plan,* ou NETP (U.S. Department of Education, Office of Educational Technology, 2017a). Ce dernier a été agrémenté d'un supplément pour l'enseignement supérieur la même année (U.S. Department of Education, Office of Educational Technology, 2017b). Tous deux intitulés « *Reimagining the Role of Technology in Education* », ils prétendent « établir une vision nationale et un plan pour l'apprentissage assisté par la technologie, à partir du travail de chercheurs phares en éducation, de dirigeants de districts, écoles et structures d'enseignement supérieur, de professeurs des écoles, de développeurs, d'entrepreneurs et d'ONG »[46] (U.S. Department of Education, Office of Educational Technology, 2017a, p.3).

Ils formulent des objectifs à destination des dirigeants et professionnels éducatifs pour réaliser une « vision nationale ». Nous restituons ceux-ci en reprenant les cinq sections des NETP : « Apprentissage », « Enseignement », « Évaluation », « Gouvernance » et « Infrastructure »[47]. Il s'agit en filigrane de cerner la cohérence et la nouveauté du projet éducatif, pour comprendre dans quelle mesure il représente une alternative au paradigme existant de la forme scolaire.

Section	Enseignements primaire et secondaire	Enseignement supérieur
Apprentissage	***National Educational Technology Plan :*** « Objectif : Tous les apprenants auront des expériences d'apprentissage engageantes et encapacitantes, à la fois dans des configurations formelles et informelles, qui les prépareront à être des participants actifs, créatifs, instruits et éthiques dans notre société connectée et mondialisée »[48] (p.9).	***National Educational Technology Plan (Higher Education) :*** « Objectif : Tous les apprenants auront des expériences d'apprentissage engageantes et encapacitantes, à la fois dans des configurations formelles et informelles, dans de multiples contextes, et à des étapes variées de leurs vies. Les apprenants seront assistés par des technologies qui étayent leur apprentissage, leur permettent de documenter leurs compétences, et les aident à se connecter utilement à des instructeurs, mentors et pairs, afin de rencontrer le succès dans des parcours professionnels et pédagogiques variés »[49] (p.17).

Dans ces deux rapports, l'apprentissage est invariablement tourné vers l'« expérience » de l'individu, rompant avec la clôture spatiale traditionnelle de la classe pour s'ouvrir à « configurations informelles » du quotidien. Sa vocation est de motiver l'individu, afin de le rendre actif et créatif dans la société, dont la projection (« connectée et mondialisée ») fait écho à la thématique de la « société de l'information » (Labelle, 2001). Dans l'enseignement supérieur, la technologie contribue à l'insertion professionnelle et sociale de l'apprenant, valorisant ses compétences et dans des contextes

[46] « *The National Education Technology Plan (NETP) sets a national vision and plan for learning enabled by technology through building on the work of leading education researchers; district, school, and higher education leaders; classroom teachers; developers; entrepreneurs; and nonprofit organizations* ».

[47] *« Learning », « Teaching », « Assessment », « Leadership » et « Infrastructure ».*

[48] *« Goal : All learners will have engaging and empowering learning experiences in both formal and informal settings that prepare them to be active, creative, knowledgeable, and ethical participants in our globally connected society »* (p.9).

[49] « *Goal : All learners will have engaging and empowering learning experiences in both formal and informal settings, in multiple contexts, and at various stages throughout their lifetimes. Learners will be supported by technology that scaffolds their learning, allows them to document their competencies, and helps them form meaningful connections to instructors, mentors, and peers to ensure their success along diverse career and educational pathways* ».

variés. L'« expérience » suit une acception proche de celle de John Dewey, remodelée suivant une vision néolibérale de la société, en tant qu'agrégat d'intérêts individuels.

Enseignement	***National Educational Technology Plan :*** « Objectif : Les éducateurs seront assistés par des technologies qui les relient aux individus, aux données, aux contenus, aux ressources, aux expertises, par des expériences d'apprentissage qui les encapacitent et les inspirent pour fournir un enseignement plus efficace à tous les apprenants »[50] (p.28).	***National Educational Technology Plan (Higher Education) :*** « Objectif : Les éducateurs de l'enseignement supérieur et post-secondaire utiliseront la technologie pour concevoir des expériences d'apprentissage qui soutiennent mieux et rendent davantage possible l'apprentissage de l'étudiant, tout en construisant et en utilisant des preuves d'efficacité pour améliorer et faire évoluer leurs approches éducatives dans le temps. En complément, les instructeurs vont : s'engager dans des expériences d'apprentissage continues les préparant à utiliser adroitement des approches basées sur la recherche pour enseigner avec les technologies ; utiliser des boucles de *feedback* alimentées par les données pour fournir des soutiens ciblés, académiques ou non, aux étudiants ; et utiliser des ressources et outils puissants pour créer des expériences collaboratives engageantes et adaptées aux besoins des étudiants »[51] (p.24).

Aux niveaux primaire et secondaire, l'éducateur voit sa marge de manœuvre élargie par les technologies, qui doivent lui donner accès à un riche réseau d'acteurs et d'informations, et par des méthodes pédagogiques reconnues comme « efficaces » par la recherche – signe de l'aspiration fédérale à une « *Evidence-Based Education* ». Concernant le professeur dans l'enseignement supérieur, l'accent est mis sur son utilisation de la technologie en vue de « concevoir » *(« design »)* des « expériences » pour ses étudiants, propices à leurs apprentissages suivant la conception constructiviste. Il doit ainsi leur fournir un soutien académique et motivationnel plus précis, ainsi que contribuer à la recherche de « preuves d'efficacité ». Il doit lui-même faire évoluer ses pratiques en accédant à une « expérience » de formation. Le paradigme fédéral prône donc une technologisation additive – le développement de la technologie en vue d'« assister » l'enseignant et d'accroître ses possibilités d'action – et non substitutive – un développement technologique qui viserait à le remplacer dans certaines tâches.

Évaluation	***National Educational Technology Plan :*** « Objectif : À tous les niveaux, notre système éducatif utilisera le pouvoir de la technologie pour mesurer ce qui compte et utiliser des données d'évaluation pour améliorer l'apprentissage »[52] (p.55).	***National Educational Technology Plan (Higher Education) :*** « Objectif : La communauté de l'enseignement supérieur collaborera pour développer des évaluations authentiques permettant de mesurer l'apprentissage et l'acquisition de compétences par les étudiants. Ces évaluations amélioreront l'apprentissage des étudiants en apportant des commentaires réguliers et en permettant la personnalisation ; aideront les enseignants à comprendre l'apprentissage des étudiants et à améliorer leur enseignement ; et aideront les institutions à suivre l'acquisition de compétences et la progression des

[50] « *GOAL: Educators will be supported by technology that connects them to people, data, content, resources, expertise, and learning experiences that can empower and inspire them to provide more effective teaching for all learners*».

[51] « *GOAL: Higher education and postsecondary instructors will use technology to design learning experiences that better support and enable student learning, while building and using evidence to improve and evolve their instructional approach over time. Additionally, instructors will engage in ongoing professional learning experiences that prepare them to adeptly apply research-based approaches to teaching with technology; use data-driven feedback loops to provide targeted academic and non-academic support to students; and use powerful tools and resources to create collaborative learning experiences that are engaging and responsive to student needs* ».

[52] *« GOAL: At all levels, our education system will leverage the power of technology to measure what matters and use assessment data to improve learning ».*

		étudiants. Les établissements éducatifs exploiteront la technologie pour permettre une évaluation précise et complète de l'apprentissage des étudiants à plus grande échelle, en s'alignant, le cas échéant, sur les normes validées en externe et élaborées par des professeurs, employeurs et autres »[53] (p.38).

Du côté des évaluations, la technologie doit permettre de mesurer « ce qui compte » pour bénéficier aux divers publics, en écho à l'*accountability* plus efficiente espérée par le président Obama. Dans l'enseignement supérieur, les perspectives sont nombreuses. Les évaluations dynamiques de l'activité des apprenants, accompagnées de « commentaires réguliers » prodigués par des humains ou des médias adaptatifs, doivent, pêle-mêle, améliorer leur progression, permettre l'intervention adéquate des enseignants, et offrir aux institutions publiques une compréhension plus fine des résultats des étudiants, dans une perspective technocratique. Les évaluations sont aussi censées permettre aux établissements de veiller à l'efficacité de leurs formations, suivant des standards formulés par la société et notamment les employeurs. La technologie évaluative et analytique doit donc impulser un cercle vertueux dans un cadre néolibéral : par l'accroissement des capacités d'action des divers publics, la « performance » individuelle se verrait améliorée et les besoins de l'appareil productif assouvis.

Gouvernance	***National Educational Technology Plan :*** « Objectif : Intégrer aux rôles et responsabilités des dirigeants à tous les niveaux une compréhension des enjeux de l'éducation basée sur la technologie ; définir des visions étatiques, régionales et locales pour la technologie dans l'apprentissage »[54] (p.42).	***National Educational Technology Plan (Higher Education) :*** « Objectif : Les dirigeants de l'écosystème de l'enseignement supérieur pourront mettre en œuvre des pratiques assistées par la technologie qui optimisent la réussite de tous les étudiants et créent une culture qui favorise la collaboration, l'innovation et le changement »[55] (p.60).

L'« éducation basée par la technologie » (« *technology-enabled education* ») doit être appropriée par les dirigeants pilotant l'enseignement primaire et secondaire, qui en seraient dorénavant les dépositaires. Sa mise en place doit aussi être « territorialisée », en se déclinant différemment selon les échelons territoriaux. De façon moins coercitive, elle pénétrerait les pratiques de gouvernance des établissements d'enseignement supérieur, pour *in fine* « optimiser » les performances de « tous les étudiants ». À l'idéal néolibéral esquissé plus haut s'ajoutent une dimension démocratique ainsi que trois valeurs cardinales faisant écho au « nouvel esprit du capitalisme » théorisé par Luc Boltanski et Ève Chiapello : la collaboration, l'innovation et le changement (Boltanski et Chiapello, 1997).

Infrastructure	***National Educational Technology Plan :*** « Objectif : Tous les étudiants et enseignants auront accès à une infrastructure solide et complète, n'importe où et quand ils en auront besoin	***National Educational Technology Plan (Higher Education) :*** « Objectif : Tous les étudiants et enseignants seront soutenus par une infrastructure robuste qui relie les environnements d'apprentissage formels, informels,

[53] *« GOAL: The higher education community will collaborate to develop authentic assessments that enable measurement of student learning and competency attainment. These assessments will improve student learning by providing frequent feedback and enabling personalization, helping faculty understand student learning and improve teaching, and assisting institutions in tracking student attainment of competencies and progress. Education providers will leverage technology to allow for the precise and comprehensive assessment of student learning at greater scale, aligning, where appropriate, with externally validated standards developed by faculty, employers, and others ».*

[54] *« GOAL: Embed an understanding of technology-enabled education within the roles and responsibilities of education leaders at all levels and set state, regional, and local visions for technology in learning ».*

[55] *« GOAL: Leaders across the higher education ecosystem will be empowered to implement technology-enabled practices that optimize student success for all of today's students and create a culture that promotes collaboration, innovation, and change ».*

	pour apprendre. Préparer les étudiants à réussir nécessite une infrastructure d'apprentissage robuste et flexible, capable de soutenir de nouveaux types d'engagement et de fournir un accès ubiquitaire à des outils technologiques qui permettent aux étudiants de créer, de concevoir et d'explorer »[56] (p.69).	en présentiel et en distanciel, pour connecter une diversité d'expériences d'apprentissage et documenter et partager les résultats d'apprentissage dans l'ensemble de l'écosystème »[57] (p.46). « Nous devons réfléchir aux moyens de garantir que l'apprentissage soutenu par la technologie soit abordable pour tous les étudiants. Certains établissements sont en mesure de rendre les opportunités de formation disponibles pour un grand nombre d'étudiants à un faible coût, via des ressources éducatives libres ; pourtant, ces économies ne reviennent pas toujours directement aux étudiants eux-mêmes. Les établissements et les professeurs devraient continuer à rechercher des expériences d'apprentissage engageantes qui tirent parti de la technologie pour réduire les coûts d'enseignement pour l'établissement ainsi que les frais de scolarité et d'autres coûts pour les étudiants »[58] (p.70).

L'ubiquité de l'infrastructure technologique, effaçant les frontières entre espaces de travail et de vie privée, ainsi que tout balisage *a priori* du temps d'apprentissage, doit permettre à l'élève des niveaux primaire et secondaire d'apprendre en expérimentant au sein de l'outil numérique, conformément à l'agenda constructiviste. Il s'agit ainsi d'« engager » de différentes manières les élèves en vue de les motiver. Dans l'enseignement supérieur, l'ubiquité doit donner lieu à des usages collaboratifs et efficients, du fait du partage croissant d'informations entre les publics. Le département fédéral encourage également l'usage de ressources éducatives libres (REL, ou *open educational ressources*, OER) disponibles sur le web. L'UNESCO (2024) propose la définition suivante des REL sur son site web :

> « Les ressources éducatives libres (REL) sont des matériels d'apprentissage, d'enseignement, et de recherche sur tout format et support, relevant du domaine public ou bien protégés par le droit d'auteur et publiés sous licence ouverte, qui autorisent leur consultation, leur réutilisation, leur utilisation à d'autres fins, leur adaptation et leur redistribution gratuites par d'autres » (UNESCO, 2024).

Les ressources éducatives libres sont encouragées par les instances internationales en vue de démocratiser l'enseignement et encourager des usages numériques dans le monde. Cela passe par des répertoires d'exercices et de leçons gratuits et certifiés comme *OER Commons*, que des acteurs individuels ou marchands peuvent consulter et réutiliser selon les fins qu'ils souhaitent. Pour l'État fédéral américain, il s'agit ainsi de réduire les coûts des études pour les élèves – une problématique ancienne et brûlante dans le pays[59].

L'idéal d'une société connexionniste est toutefois grevé par des précautions éthiques prises par l'État fédéral et réaffirmées par le *Department*. Les deux NETP mentionnent à cet égard les lois fédérales

[56] « *GOAL: All students and educators will have access to a robust and comprehensive infrastructure when and where they need it for learning. Preparing students to be successful for the future requires a robust and flexible learning infrastructure capable of supporting new types of engagement and providing ubiquitous access to the technology tools that allow students to create, design, and explore ».*

[57] *« GOAL: All students and educators will be supported by a robust infrastructure that bridges formal, informal, workplace, and mobile learning environments to connect a diversity of learning experiences and document and share learning outcomes across the ecosystem ».*

[58] *« We must consider ways to ensure technology-enabled learning is affordable for students. In some cases, institutions are able to make educational opportunities more widely available at a low cost through digital or open educational resources; yet those savings are not always passed directly to students themselves. Institutions and faculty should continue to explore engaging learning experiences that leverage technology to reduce instructional costs for the institution and cost of tuition and fees for students ».*

[59] La dette étudiante aux États-Unis s'est élevée en 2023 à 1 760 milliards de dollars. C'était plus que le PIB de l'Australie la même année (1 670 milliards de dollars) selon le cabinet FinMasters (Ciochia, 2024).

qui encadrent le développement et le déploiement des technologies éducatives (U.S. Department of Education, Office of Educational Technology, 2017a, pp.80-81) :

- le *Family Educational Rights and Privacy Act* (FERPA), voté en 1974, accorde aux parents le droit de consulter les dossiers scolaires de leurs enfants et le droit d'exercer un contrôle sur la divulgation d'informations personnelles identifiables. Le FERPA encadre également l'usage des données dans l'enseignement supérieur, ce droit étant transféré aux enfants à leur majorité (U.S. Department of Education, Office of Educational Technology, 2017b, pp.54-55).
- Le *Protection of Pupil Rights Amendment* (PPRA) de 1978 définit les informations pouvant être collectées par les districts auprès des élèves de l'éducation primaire et secondaire dans le cadre de programmes fédéraux.
- L'*Individuals with Disabilities Education Act* (IDEA) approfondit depuis 1990 les dispositions de *FERPA* en matière de protection des informations personnelles dans le cas des enfants en situation de handicap.
- Le *Children's Online Privacy Protection Act* (COPPA), voté en 1998, encadre d'une part la collecte en ligne de renseignements personnels sur des enfants de moins de treize ans, en requérant le consentement exprès d'un responsable légal, et restreint d'autre part le suivi des étudiants par les éditeurs de logiciels à des fins éducatives exclusives.

Ces textes sont complétés par les lois votées au niveau des États. Prises ensemble, ces directives obligent les éditeurs de logiciels à s'assurer du consentement des élèves ou de leurs tuteurs quant à l'utilisation des données d'apprenants. En outre, elles contraignent les districts à établir des chartes (*Acceptable* et *Responsible Use Policy*) encadrant les utilisations de logiciels éducatifs par les écoles et la manière dont les données sont transmises à des acteurs tiers.

Ainsi le projet fédéral de *personalized learning* remet en cause des traits saillants de la forme scolaire, tels que le balisage spatio-temporel de l'apprentissage, ou les revisite, comme le contrat didactique entre l'apprenant et le formateur. Bien qu'amendé par des précautions éthiques, sa cohérence d'ensemble permet d'en faire un idéal-type commun aux échelons éducatifs primaire, secondaire et supérieur. La technologie numérique y joue un rôle central. Elle doit d'abord offrir à l'élève une « expérience d'apprentissage » formelle et informelle, dans et en dehors des institutions éducatives, suscitant sa motivation et favorisant sa créativité et sa collaboration avec ses pairs. Il doit ainsi être préparé à évoluer dans une société « connectée » et à répondre aux attentes flexibles du monde du travail. La technologie doit permettre à l'enseignant de mieux l'accompagner dans la fourniture de ces expériences, dans le sens d'une technologisation additive, en se conformant à des « preuves d'efficacité ». L'administrateur doit aussi contrôler, par les données, l'efficience de l'échelon éducatif sous son autorité. Cette projection d'un « écosystème » rompt en définitive avec le paradigme standardisé du « *one size fits all* », en conciliant une pédagogie socio-constructiviste voire connectiviste avec un industrialisme empreint de l'« *Evidence-Based Education* ».

En 2023, le nouveau rapport de l'OET intitulé « *Artificial Intelligence and Future of Teaching and Learning: Insights and Recommendations* » confronte cette vision fédérale aux développements de l'intelligence artificielle, dont les avancées en matière d'adaptativité (U.S. Department of Education, Office of Educational Technology, 2023, p.18).

Comment les technologies adaptatives soutenues par l'IA sont-elles intégrées dans le projet fédéral de *personalized learning* ? Elles sont envisagées à la fois comme une opportunité et un danger. Dans le premier cas, le bureau fédéral explique que l'adaptativité est un moyen efficace d'améliorer

l'apprentissage (U.S. Department of Education, Office of Educational Technology, 2023, p.18). Il fonde cette assertion sur un article collectif dont deux des auteurs, Vincent Aleven et Ken Koedinger, sont par ailleurs les cofondateurs de Carnegie Learning, éditeur de l'ITS *MATHia*. Ce biais met en cause la neutralité du *Department*, participant, volontairement ou non, d'une prophétie industrielle « auto-réalisatrice » qui légitime l'introduction de l'IA dans l'éducation.

La seconde posture, critique, cible les limites techniques de ces technologies, telles que les angles morts du modèle de l'apprenant.

> « Les limites fondamentales découlent de la nature du modèle au cœur de tout système spécifique basé sur l'IA. Les modèles sont des approximations de la réalité. Lorsque des pans importants de l'apprentissage humain sont laissés de côté ou moins développés, l'adaptativité qui en découle est également limitée, et les supports d'apprentissage qui en résultent risquent d'être moins fiables »[60]. (ibid., p.18)

L'OET demeure donc précautionneux vis-à-vis des bénéfices pédagogiques à attendre pour les apprenants. Il soulève des questions rhétoriques telles que « l'IA mène-t-elle à des activités d'élèves plus restreintes (par exemple, des problèmes de mathématiques procédurales) ou à la gamme d'activités mises en avant dans le *National Educational Technology Plan* (NETP), qui met l'accent sur des fonctionnalités telles que le *personalized learning* […] ? »[61] (ibid., p.23). Ou encore « l'IA soutient-elle l'apprenant dans son ensemble, y compris les dimensions sociales de son apprentissage telles que permettre aux étudiants de participer activement à l'apprentissage en petits groupes et en collaboration ? »[62] (ibid., p.23). Selon le bureau fédéral, ces technologies seraient à un tournant : soit elles épousent le paradigme fédéral du *personalized learning* et son agenda socio-constructiviste, soit elles basculent du côté d'une pédagogie restrictive et aliénante pour les apprenants.

Pour soutenir la première voie, l'OET recommande une approche complémentaire entre l'algorithme et l'humain : « travailler vers des modèles d'IA qui correspondent à l'intégralité des visions de l'apprentissage et évitent de limiter l'apprentissage à ce que l'IA peut actuellement modéliser »[63] (ibid., p.18). Le *Department* apporte son soutien explicite à une technologisation additive et à des démarches d'« IA centrée sur l'humain », ou d'« humain dans la boucle ».

Le *Department* est tout aussi précautionneux sur des questions éthiques telles que la confidentialité des données personnelles et l'information des usagers quant au traitement de leurs données : « Lorsque l'IA est utilisée, la vie privée et les données des élèves sont-elles protégées ? Les étudiants et leurs tuteurs sont-ils informés de ce qui arrive à leurs données ? »[64] (ibid., p.23).

Des considérations politiques comme la lutte contre les inégalités sociales sont aussi soulignées : « Alors que nous façonnons les technologies éducatives améliorées par l'IA autour de principes fondés sur la recherche, un objectif clé doit être de renforcer et de soutenir l'apprentissage de ceux

---

[60] « *Core limits arise from the nature of the model at the heart of any specific AI-enabled system. Models are approximations of reality. When important parts of human learning are left out of the model or less fully developed, the resulting adaptivity will also be limited, and the resulting supports for learning may be brittle or narrow* ».

[61] *« Is AI leading to narrower student activities (e.g., procedural math problems), or the fuller range of activities highlighted in the National Educational Technology Plan (NETP), which emphasizes features such as personalized learning […] ? »*

[62] *« Is AI supporting the whole learner, including social dimensions of learning such as enabling students to be active participants in small group and collaborative learning? […] »*

[63] *« Consequently, this section on Learning focuses on one key concept: Work toward AI models that fit the fullness of visions for learning—and avoid limiting learning to what AI can currently model well ».*

[64] *« When AI is used, are students' privacy and data protected? Are students and their guardians informed about what happens with their data? »*

qui ont connu des circonstances défavorables, telles que celles causées par la pandémie de Covid-19 ou par des inégalités plus larges »[65] (ibid., p.18).

La dimension éthique continue d'être détaillée dans la dernière version du *National Education Technology Plan* publiée en janvier 2024. Celle-ci est construite autour de trois fractures à combler : celle des usages actifs des technologies (*Digital Use Divide*), celle de l'accès à des terminaux, des médias et une connexion de qualité (*Digital Access Divide)*, et celle des visées pédagogiques *(Digital Design Divide*), supposant une formation adéquate des enseignants à l'usage des technologies. Encourageant à nouveau le déploiement de médias d'obédience socio-constructiviste, l'OET en revient à recommander aux administrateurs des processus de sélection industrialistes et empreints par l'*Evidence-Based Education*, basés sur les « *four tiers of evidence »* prévus par l'ESSA (U.S. Department of Education, Office of Educational Technology, 2024, p.49).

La structuration des recherches en IA éducative : les AI Institutes

La seconde stratégie fédérale vise à structurer les recherches en matière d'IA éducative selon l'agenda du *personalized learning*. Le *National Artificial Intelligence Research and Development Strategic Plan*, signé en 2016, cible sept stratégies pour développer la recherche en IA. La première, *« Make Long-Term Investments in AI Research* », recommande des financements de la *National Science Foundation* dans la recherche fondamentale en IA. Son actualisation en 2019 y ajoute la nécessité de renforcer les partenariats public-privé. En 2020 sont créés, sous l'égide de la NSF, les « *AI institutes »* en réponse à ces deux recommandations. Parmi les six axes de recherche attendus dans ces instituts figure l'« *AI-Augmented Learning* ». Au début de l'année 2024, 25 *AI institutes* existent dans le pays, dont cinq axés sur l'IA en éducation[66]. Leurs caractéristiques principales sont restituées en annexe 1. Chacun est financé à hauteur de 20 millions de dollars sur cinq ans. Ils visent à développer des dispositifs d'IA et à prouver leur efficacité en vue d'une industrialisation future.

Ils reprennent ainsi le modèle du « démonstrateur technologique » (« *demonstrator model* »), usuel dans la recherche militaire, visant à développer et éprouver un prototype technologique en situation d'usage, pour fournir une « preuve de concept » auprès des acteurs socio-économiques. Ce mode de recherche fut impulsé par l'État fédéral dès la Seconde Guerre mondiale puis tout au long de la Guerre froide, pour soutenir la recherche militaire, nucléaire et spatiale. Les historiens des systèmes universitaires Christophe Charle et Jacques Verger soulignent alors « l'investissement énorme du gouvernement fédéral dans la recherche […] en fonction des besoins de ce que le président Eisenhower lui-même a baptisé le "complexe militaro-industriel" » (Charle et Verger, 2012, p.153).

Avec les « *AI institutes »*, cette organisation réticulaire de la recherche autour de prototypes est reprise par le pouvoir fédéral. Témoin les travaux du *National AI Institute for Student-AI Teaming,* qui visent à développer des « *AI partners* », soit des systèmes d'assistance aux enseignants devant favoriser le travail collaboratif en petits groupes d'élèves. Ceux-ci mobilisent les chercheurs de huit universités, une agence de développement web, un pôle de compétitivité public-privé, une association et deux districts scolaires, dans lesquels sont menés des ateliers de co-construction avec les publics scolaires (National AI Institute for Student-AI Teaming, 2024).

[65] « *As we shape AI-enhanced edtech around research-based principles, a key goal must be to strengthen and support learning for those who have experienced unfavorable circumstances for learning, such as caused by the COVID-19 pandemic or by broader inequities* ».
[66] Il s'agit du *National AI Institute for Student-AI Teaming* (Université du Colorado), du *National AI Institute for Adult Learning and Online Education* (Institut de Technologie de Géorgie), le *AI Institute for Engaged Learning* (Université de Caroline du Nord), du *National AI Institute for Exceptional Education* (Université de Buffalo) et du *National AI Institute for Inclusive Intelligent Technologies for Education* (Université de l'Illinois).

Les recherches menées dans les *AI institutes* épousent les recommandations techniques et éthiques formulées par l'OET. Les axes de recherche ciblent principalement des outils d'intelligence artificielle « additifs », devant accroître les capacités d'analyse et d'intervention des professeurs et formateurs. Il en va par exemple du *National AI Institute for Student-AI Teaming*, qui développe des systèmes d'assistance aux enseignants pour favoriser le travail collaboratif en petits groupes d'élèves.

Ces recherches obéissent en outre à des agendas éthiques. Le *National AI Institute for Exceptional Education* ancre par exemple ses travaux d'identification automatique des difficultés d'élocution chez les jeunes enfants dans le courant de l'« IA centrée sur l'humain ».

La circulation du personalized learning *en question*

Étant donnée la décentralisation forte de l'appareil éducatif états-unien, l'on est amené à interroger la circulation du projet de *personalized learning* au-delà de Washington, notamment à l'échelon des États fédérés, primordiaux dans l'organisation de l'éducation et de la formation. Reprennent-ils ce projet dans leurs propres textes normatifs, au-delà de leurs différences politiques et organisationnelles ? Une veille documentaire nous a révélé que de nombreux États fédérés, toutes étiquettes politiques confondues, avaient publié des feuilles de route à destination de leurs districts promouvant des approches similaires au *personalized learning*. Du côté des *adoption states,* encadrant plus étroitement les achats d'outils des établissements, l'État républicain de l'Utah a par exemple publié en 2023 un livret portant sur le « *Personalized, Competency-Based Learning (PCBL)* » (Utah State Board of Education, 2023). Dans l'État républicain de Géorgie, le *Governor's Office of Student Achievment* a mis en place un programme de subventions, le *Personalized Learning grants*, à destination de quatorze écoles implémentant ce projet technopédagogique (The Governor's Office of Student Achievment, 2024). Du côté des *open territory states*, laissant davantage de liberté aux districts et établissements dans le choix de leurs outils, le *Department of Public Instruction* du Wisconsin a collaboré dans les années 2010, sous des gouvernements démocrates puis républicains, avec l'*Institute for Personalized Learning*, afin de documenter et d'accompagner des districts dans l'implémentation de cette approche (Midwest Comprehensive Center, 2018). Cela confirme le consensus bipartisan portant sur l'implémentation de technologies éducatives aux États-Unis.

Une volonté politique d'implémentation d'un projet éducatif cohérent, le *personalized learning,* s'est ainsi construite à l'échelle fédérale pour gagner les États fédérés. Ce projet est affilié à l'*Evidence-Based Education,* courant industrialiste revisitant les préceptes du *Social Efficiency Movement.* Il opère néanmoins une rupture sur le plan pédagogique avec le behaviorisme de l'enseignement programmé, devenu peu à peu synonyme de standardisation, pour embrasser la cause socio-constructiviste. Eu égard à la distinction opérée en introduction, ce volontarisme procède à la fois d'objectifs endogènes à l'éducation, liés à la transformation des pratiques d'apprentissage, d'enseignement et d'organisation de l'éducation qu'induirait la technologie ; et exogènes, liés à des considérations d'ordres économiques et politiques, en l'occurrence la volonté de répondre au décrochage des élèves américains dans les études de niveaux, et de mener la course internationale à l'intelligence artificielle, notamment face à la Chine.

Le projet étant formalisé, des questions apparaissent rapidement sur un autre pan des initiatives publiques : la poursuite du *personalized learning* « sur le terrain », par les établissements éducatifs

publics. L'appareil scolaire et universitaire états-unien étant fortement décentralisé, la volonté politique et la capacité financière des établissements apparaissent comme des facteurs cruciaux dans le déploiement de médias adaptatifs conformes au *personalized learning*.

Les contraintes logistiques des districts et des établissements d'enseignement public

Le financement contraint de l'équipement

Pour inciter les établissements scolaires et universitaires à implémenter ce projet, les échelons fédéraux et fédérés accordent, à eux ou à leurs districts, des subventions ciblées. Cependant celles-ci s'avèrent loin d'être suffisantes pour impulser une politique *top-down.* Elles représentent en effet une part faible des ressources des établissements. Pour les États ayant communiqué les sommes reçues en 2021 dans le cadre du *Student Support and Academic Enrichment* (SSAE) – le programme de l'ESSA comprenant un volet consacré au *personalized learning* – celles-ci ne représentent que 1,08% de l'ensemble des subsides fédéraux (National Center for Education Statistics, 2023c, pp.18-19). Subsides qui constituent déjà une part minoritaire des budgets des établissements – 11% des revenus des *grade, middle* et *high schools* en 2021 (Peter G. Peterson Foundation, 2023).

L'influence des États fédérés sur l'équipement des établissements est complexe, variant selon leurs contributions financières d'une part, et selon leurs degrés d'intervention (*adoption states* ou *open territory states*) d'autre part. On sait que les États apportent l'essentiel des revenus des écoles primaires et secondaires – à hauteur de 45% en 2021 (Peter G. Peterson Foundation, 2023) –, et sont l'échelon participant le plus aux dépenses de fonctionnement des établissements d'enseignement supérieur – à hauteur de 17% en 2020 (Urban Institute, 2022). L'influence budgétaire des États est donc notable, malgré la volatilité de leurs dépenses en cas de récession et les inégalités de rattrapage qui les distinguent. Des disparités importantes existent ainsi entre les 17 États, qui, selon The Pew Charitable Trusts, dépensaient moins en 2019 qu'avant la crise financière de 2008, et les autres qui avaient accru leurs subventions (The Pew Charitable Trusts, 2019).

Les agences éducatives des vingt *adoption states* tendent d'autre part à structurer l'offre disponible pour les établissements publics sous leur juridiction, en formulant des listes de produits éducatifs agréés, parmi lesquels les districts doivent passer commande pour toucher des fonds publics. La *Texas Education Agency* propose par exemple un programme de subventions et d'accompagnement, les « *Learning Acceleration Support Opportunities* », aux districts qui commandent des médias pédagogiques parmi une liste d'éditeurs approuvés (Texas Education Agency, 2022).

Dans l'enseignement supérieur, certains *State University Systems* proposent des dispositifs d'achat centralisés et mutualisés à leurs universités membres. Le système d'achat *GeorgiaBEST* du *University System of Georgia* en est un exemple (University System of Georgia, 2024).

Cependant, étant données la fin des aides fédérales liées à la pandémie en 2024 et la réduction des subsides de certains États fédérés depuis les années 2010, les districts et établissements publics doivent de plus en plus composer avec des ressources limitées pour implémenter ce projet. Or les districts optent pour un investissement marginal dans le numérique éducatif. S'il n'est pas possible de connaître l'affectation exacte de leurs revenus, on sait du moins que leurs dépenses vont

très majoritairement aux salaires et avantages sociaux des personnels, et minoritairement aux achats de services comprenant les technologies éducatives[67].

Selon le réseau *Edtech Evidence Exchange*, « les estimations suggèrent que les écoles américaines, de la maternelle à la douzième année, ont dépensé entre 26 et 41 milliards de dollars par an en technologies éducatives au cours de l'année précédant la pandémie » (Edtech Evidence Exchange, 2021, p.5). Ces chiffres, rapportés aux dépenses totales déclarées par les établissements pour l'année 2019, soit 769 milliards de dollars (National Center for Education Statistics, 2021, p.4), représenteraient une part comprise entre 3,4% et 5,3%. L'investissement, non-négligeable, reste semble-t-il insuffisant pour instrumenter un projet aussi fondamental que le *personalized learning*.

Les doutes pesant sur l'efficience des médias numériques

Chargées par les pouvoirs publics de gérer leurs établissements selon le principe d'*accountability*, avec en prime des ressources limitées, les équipes dirigeantes sont en quête d'outils et de médias éducatifs « efficients ». Or, dans le cas des logiciels éducatifs et notamment adaptatifs, la question de l'efficacité est pour le moins ambiguë.

Cette complexité transparaît dans la volonté fédérale de soutenir les établissements dans leurs procédures de sélection, en incitant les éditeurs de technologies éducatives à produire des études d'efficacité. Le *What Works Clearinghouse*, portail créé par l'*Institute of Education Sciences* en 2002, centralise les recherches d'efficacité préalablement jugées rigoureuses par un comité de lecture, suivant les quatre « niveaux de preuves » (« *four tiers of evidence* ») prévus par l'*Every Student Succeeds Act* (What Works Clearinghouse, 2024). Les critères sont la mise en place d'études randomisées, la profondeur des échantillons et de faibles taux d'attrition (What Works Clearinghouse, 2024). Toutefois les travaux de cette initiative fédérale ne font pas consensus dans la communauté académique, à cause notamment d'une tendance à la surinterprétation de résultats préliminaires, telle que soulignée par une équipe d'évaluation de la fondation Arnold Ventures (Straight Talk on Evidence, 2018).

Les référentiels de l'ESSA ont par ailleurs quelques difficultés à imprégner les pratiques des éditeurs. Il est apparu en 2022 qu'une minorité d'éditeurs conduisaient des études d'efficacité robustes. Learnplatform a ainsi étudié les recherches disponibles pour les cent technologies éducatives les plus utilisées dans les classes américaines. 39% d'entre elles avaient publié des études d'efficacité, et 26% de ces recherches se référaient aux « *tiers of evidence* » de l'ESSA (Learnplatform, 2023). Parmi elles, seules cinq services avaient le grade le plus élevé (« *Strong Evidence* »).

D'aucuns dénoncent également les discours infondés et partisans des acteurs gouvernementaux, selon lesquels les districts auraient tout intérêt à implémenter le *personalized learning*, et en particulier les ITS, étant données les promesses de parcours d'apprentissage personnalisés et d'informations fiables sur la qualité des apprentissages. Comme le relatait un article du média spécialisé *Education Week*, dans le cas du programme *Race to the Top-District*, par lequel plus de

[67] Parmi les dépenses des établissements publics du primaire et secondaire compilées par le *National Center for Education Statistics*, celles en logiciels éducatifs rentrent dans la catégorie « *Instruction* », et plus précisément dans la sous-catégorie « *All other expenditures* » qui représente 55,2 milliards de dollars, soit 12,1% de la catégorie. Les deux autres sous-catégories (« *Salaries and wages* » et « *Employee benefits* ») occupent la majeure partie du poste (61,5% et 26,4%, National Center for Education Statistics, 2023c).

500 millions de dollars sont allés à 21 districts en 2012 et 2013, « les propres rapports du *Departement* ont parfois tiré des conclusions qui ne semblent pas justifiées par les preuves disponibles »[68] :

> « Par exemple, une série d'études de cas affirmaient que le *personalized learning* avait déclenché des "changements culturels et transformé l'apprentissage des étudiants" dans le comté de Miami Dade en Floride, et de New Haven en Californie. Pour étayer cette affirmation, le rapport citait en grande partie « des résultats flous et auto-déclarés d'enseignants affirmant qu'ils se sentent dorénavant plus à l'aise pour prendre des risques en classe »[69] (Herold, 2016).

En découle une défiance de la part des acteurs de terrain. La « recherche-action » est avancée par de nombreux experts comme une manière pour les districts, les établissements et les professionnels de bâtir une stratégie efficace d'évaluation. Comme le relate la chercheuse australienne Renée Crawford, cette méthode allie l'enquête scientifique avec l'action pratique en vue de résoudre des problèmes concrets (Crawford, 2022). Inspirée de principes pragmatistes comme ceux de John Dewey, la recherche-action est théorisée en 1946 par le psychologue Kurt Lewin comme un processus en plusieurs étapes, comprenant la planification, l'action, l'observation des résultats et l'ajustement des pratiques en conséquence. En éducation, elle peut être menée par des équipes de recherche impliquant des enseignants, qui collectent des données sur leurs pratiques et mènent des expérimentations pédagogiques en classe. Ces équipes analysent *in fine* les résultats pour tenter d'ajuster ensuite les stratégies pédagogiques (ibid.). Dans cette veine, nombreuses sont les institutions éducatives ayant recours, dans le processus d'achat (« *procurement* »), à une phase préalable de « *pilot testing* », ou expérimentation d'une version *beta* dans un nombre restreint de classes, en partenariat avec une université, pour juger de l'utilité d'implémenter un outil donné (Berger et Stevenson, 2007).

Cité dans l'article d'*Education Week*, Bart Epstein, directeur du Jefferson Education Accelerator – un réseau de chercheurs, d'éducateurs et d'entrepreneurs basé à l'Université de Virginie –, recommande ainsi que les districts s'engagent, notamment financièrement, dans ces démarches : « Les petits districts […] peuvent […] s'adresser à une université locale ou s'engager dans ce qu'on appelle une évaluation en « cycle court », pour obtenir un retour formatif au fur et à mesure. Tout district qui lance un nouveau programme majeur devrait absolument prévoir un véritable soutien à la recherche dans son budget »[70].

Néanmoins ces phases sont longues et coûteuses pour les districts qui disposent, on l'a vu, de budgets déjà contraints. Autre obstacle : le système traditionnel de *licensing* tend à profiter aux éditeurs qui vendent des lots de plusieurs de leurs produits aux districts, sans possibilité pour ces derniers de sélectionner celui qu'ils utiliseront réellement. Sans intervention de l'État fédéral dans ces marchés, les agences éducatives de certains États fédérés et certains districts tentent d'imposer leurs procédures d'achat de façon isolée (Encadré 1).

---

[68] *« …the department's own reports have sometimes drawn conclusions that don't seem warranted by the available evidence ».*

[69] « *For example, a set of recent case studies claimed that personalized learning had sparked "cultural shifts and transformed student learning" in such places as Miami-Dade County, Fla., and New Haven, Calif. To back that up, the report cited largely soft and self-reported findings, such as teachers saying they are now more comfortable taking classroom risks* ».

[70] *« Small districts (…) can (…) reach out to a local university or engage in what's called "short cycle" evaluation, to get formative feedback as they go ».*

Encadré 1. À défaut de moyens supplémentaires, certains États mettent en place des procédures d'achat alternatives

Pour réduire la proportion d'outils inutilisés, l'État du Utah accorde, à travers son *Utah STEM Action Center,* une subvention pour les districts s'équipant en logiciels de *personalized learning* en mathématiques (le « *K-12 Math Personalized Learning Software Grant »*). Il est demandé aux candidats d'identifier les besoins des écoles en amont (Utah STEM Action Center, 2024)[71]. En outre, le programme pose plusieurs exigences auprès des entreprises sélectionnées, comme le relate un article du média spécialisé *The Hechinger Report* : pour être payées, celles-ci doivent au préalable proposer des formations aux enseignants de chaque école. Ensuite elles doivent restituer toute licence inutilisée, pour une réutilisation l'année scolaire suivante. Une forme de garantie de remboursement pour les districts, qui aurait cependant dissuadé plusieurs éditeurs (Yoder, 2023).

Certains districts posent également des exigences nouvelles sur le modèle de *licensing* ainsi que le service après-vente. Le *Baltimore County Public Schools* (BCPS) oppose aux éditeurs, habitués à ce que les districts achètent des *packs* incluant plusieurs de leurs outils, une approche « iTunes », selon l'expression de la directrice de son département informatique, citée par Edsurge (2016, p.49) : « Nous ne [voulons] pas acheter tout l'album – juste les chansons »[72]. Une exigence de sélection d'outils spécifiques que refuseraient de nombreuses entreprises. Dissuasive pour certains petits acteurs est aussi la demande d'une assistance privilégiée aux heures « de pointe » ou durant le week-end, quand de nombreux enseignants ont le temps d'analyser les données collectées sur l'activité des étudiants. « Il y [a] beaucoup d'inflexibilité quant à ce que les entreprises [feront] et ne [feront] pas »[73] confie à ce propos une responsable du district (p.49). Cette stratégie du BCPS est un exemple isolé de politique ferme et proactive d'un district face aux éditeurs. La plupart ne disposent cependant pas de stratégies claires et homogènes, ce qu'ont révélé crûment la pandémie de Covid-19 et les achats en urgence qui en ont découlé.

Face à ce manque de liquidités, les fondations privées, pour beaucoup issues d'entreprises du secteur technologique (Bill and Melinda Gates Foundation, Dell Foundation, Hewlett Foundation) et les associations à but non-lucratif (LEAP Innovations) subventionnent de longue date les organisations éducatives comme les districts, écoles privées (*charter schools*) et universités d'État dans leurs stratégies d'équipement. Ils collaborent notamment avec les pouvoirs publics dans des programmes de financement d'ampleur. Ce fut le cas en 2009 lors du programme *Investing in Innovation* (i3), voté dans le cadre de l'*American Recovery and Reinvestment Act* (ARRA), visant à soutenir le développement, la validation et la mise à l'échelle de pratiques éducatives jugées innovantes. Douze fondations ajoutèrent 500 millions de dollars aux 650 millions prévus par le *Department of Education*[74].

Ainsi les politiques publiques gouvernementales répliquent-elles les incohérences mises en exergue dans le chapitre introductif de cette partie. Elles manquent de données fiables sur l'efficacité de telle initiative menée par le passé ou de tel média éducatif. Elles éparpillent leurs financements entre une politique de l'offre visant à soutenir des firmes éducatives dans la conduite d'études d'efficacité non-consensuelles, et une politique de la demande, consistant à financer des districts souvent désorientés lors du processus d'acquisition de médias éducatifs. Leurs manques sont apparus d'autant plus crûment durant la pandémie de Covid-19, lorsque le distanciel a dû être généralisé.

[71] « L'objectif est de demander uniquement ce que vous allez utiliser et ainsi garantir que toutes les licences répondront à une « utilisation efficace » – c'est-à-dire que les étudiants utilisent un produit donné plus de 30 à 50 minutes par semaine en moyenne, et que les enseignants utilisent des données pour étayer leur enseignement chaque semaine – et que l'on atteigne autant d'étudiants que possible grâce au financement fourni par le législateur » *(« The goal is to ask only for what you will use and thus ensure that all licenses will meet "effective usage" – students using a given product more than 30-50 minutes a week on average, and teachers using data to inform instruction weekly – requirements and that we reach as many students as possible with the funding that has been provided by the legislature »).*

[72] « *We didn't want to buy the whole album - just the songs we wanted. That blew their minds »*

[73] *« There was a lot of inflexibility on what many of the companies would and would not do ».*

[74] Les douze fondations étaient l'Annie E. Casey Foundation, la Bill & Melinda Gates Foundation, la Carnegie Corporation of New York, la Charles Stewart Mott Foundation, la Ford Foundation, la John D. & Catherine T. MacArthur Foundation, la Lumina Foundation for Education, la Robertson Foundation, la Wallace Foundation, la Walton Family Foundation, la William & Flora Hewlett Foundation et la W.K. Kellogg Foundation.

Dès lors le projet *personalized learning* apparaît non seulement comme un horizon lointain, mais aussi comme une énième prophétie soutenue par des acteurs industriels et gouvernementaux. De l'utopie à la déception, il n'y a qu'un pas, déjà franchi par d'autres technologies éducatives avant l'intelligence artificielle, ainsi que l'a systématisé Larry Cuban (Cuban, 1993).

Cependant, en amont des initiatives des acteurs publics états-uniens, qu'en est-il des stratégies des firmes ? S'accorderaient-elles autour du projet du *personalized learning*, de sa vision connexionniste de la société et de son approche additive et ouverte de la technologie ? S'en détourneraient-elles au contraire ? On a observé dans la partie précédente que les firmes avaient des profils extrêmement variés, et étaient engagées dans cinq industries éducatives aux problématiques spécifiques. Avant d'observer la manière dont elles codent leurs conceptions éducatives dans leurs produits, il convient de caractériser leurs rapports de force au sein d'une socio-économie intégrée des médias adaptatifs, pour jauger la cohésion et les valences industrielles et politiques de leurs stratégies.

2. Les stratégies des firmes

Notre hypothèse formulée dans le chapitre précédent est que les algorithmes d'apprentissage adaptatif contribueraient à la centralité des acteurs des technologies dans cette socio-économie, ainsi qu'à la convergence entre les industries culturelles et éducatives. En découlerait un projet techno-éducatif potentiellement complémentaire avec le *personalized learning.*

Pour éprouver cette hypothèse généraliste, nous adoptons un point de vue socio-économique sur une filière intégrée des médias adaptatifs. Celle-ci est telle que présentée dans le chapitre précédent, divisible en trois étapes principales : en amont, la conception d'algorithmes d'apprentissage adaptatif et la production de contenus pédagogiques, et en aval, la distribution du média auprès d'un public éducatif en « *Software as a Service* ». Nous étudions les stratégies des firmes de statuts et d'activités très divers, impliquées dans ces trois étapes : les structures de production informatique de type « start-ups », les acteurs plus anciens et concentrés du numérique éducatif, les éditeurs scolaires proposant des offres numériques, les gestionnaires de plateformes d'accès aux médias numériques, et les multinationales des technologies ou *Big Tech*. Il convient, comme nous l'avons expliqué plus haut, d'observer les manières dont des tendances nouvelles à la fédération ou à l'éclatement de ces acteurs concourent à des stratégies rénovées, à l'aune des modèles socio-économiques issus de la théorie des industries culturelles.

Le développement des algorithmes d'apprentissage adaptatif

Trois types d'acteurs jouent un rôle structurant dans la conception d'algorithmes d'apprentissage adaptatif : les start-ups, les *Big Tech* et les grands groupes du numérique éducatif.

La conception d'algorithmes d'apprentissage adaptatif, cœur de métier des start-ups

Les start-ups sont généralement créées par des ingénieurs, des informaticiens, des chercheurs et doctorants issus de laboratoires de *Learning Sciences*, ou des anciens enseignants. Elles investissent la majorité de leurs ressources dans leur recherche et développement, et en particulier dans leurs algorithmes, dans l'espoir d'en tirer un avantage concurrentiel et des revenus futurs.

Le développement de ces technologies nécessite des fonds importants, notamment pour attirer et rétribuer des ingénieurs qualifiés, mais aussi pour distribuer, en *front-end*, un média attractif auprès des apprenants, permettant aux algorithmes de *machine learning* d'être entraînés sur une base d'utilisateurs suffisante.

Les fonds de capital-risque (« *venture capital* ») représentent le principal guichet de financement des start-ups. Les États-Unis représentent un des principaux foyers mondiaux de capital-risque pour l'éducation. Les fonds levés par des entreprises de l'*EdTech* ont atteint un record en 2021 avec 8,2 milliards de dollars (Wan, 2024), contre 487 millions d'euros en France par exemple (BrightEye Ventures, 2024). Plusieurs facteurs favorables aux levées de fonds discriminent les start-ups.

Un premier facteur est le secteur d'activité et sa rentabilité telle que perçue par les investisseurs. Les levées les plus importantes sont à ce titre réalisées dans le *K-12* (éducation primaire et secondaire), où la décentralisation des chaînes d'achat et la multitude de districts créent un marché porteur. En 2020, année de la pandémie, de la fermeture des écoles et de leur passage généralisé au distanciel, le marché du numérique éducatif scolaire aurait représenté 35,8 milliards de dollars (Cauthen, 2021), contre 4,9 milliards de dollars pour l'enseignement supérieur (Grand View Research, 2022). Ainsi en 2023, trois des cinq levées les plus importantes ont été réalisées par des firmes spécialisées dans le *K-12* : Amplify avec 350 millions de dollars, ClassWallet et HOMER avec 95 millions de dollars (Wan, 2024).

Un second facteur est le recours à l'intelligence artificielle comme argument commercial et marketing auprès des investisseurs. Les start-ups liées à l'intelligence artificielle, en particulier générative, auraient amassé près de 50 milliards de dollars en levées de fonds dans le monde selon la base de données Crunchbase (Metinko, 2024), près de neuf fois plus que l'ensemble des start-ups de l'*EdTech* mondiale et leurs 5,6 milliards de dollars levés (BrightEye Ventures, 2024). En découle une stratégie d'éligibilité courante chez les start-ups de l'éducation, consistant à présenter leurs services comme basés sur de l'intelligence artificielle, alors que pour certaines les technologies en question sont peu innovantes. On peut l'observer chez des firmes prétendant offrir des prestations adaptatives alors que leurs méthodologies consistent en des traitements statistiques communs, notamment en matière de répétition espacée.

Des dizaines de fonds spécialisés existent, comme LearnLauch Fund+Accelerator, Rethink Education, New Markets Venture Partners, Owl Ventures ou GSV Ventures. D'autres fonds généralistes ont des participations dans des entreprises du numérique éducatif comme Reach Capital ou SJF Ventures.

Cependant la volatilité des levées est un facteur de bouleversement pour le secteur. Après le record de 2021, la somme totale chute à 5,2 milliards en 2022, puis 2,8 milliards en 2023 (Figure 9., Wan, 2024). De nombreux facteurs sont en cause : la montée des taux d'intérêt, la fin de la pandémie de Covid-19, ou la défiance accrue de certains investisseurs quant aux technologies d'intelligence artificielle.

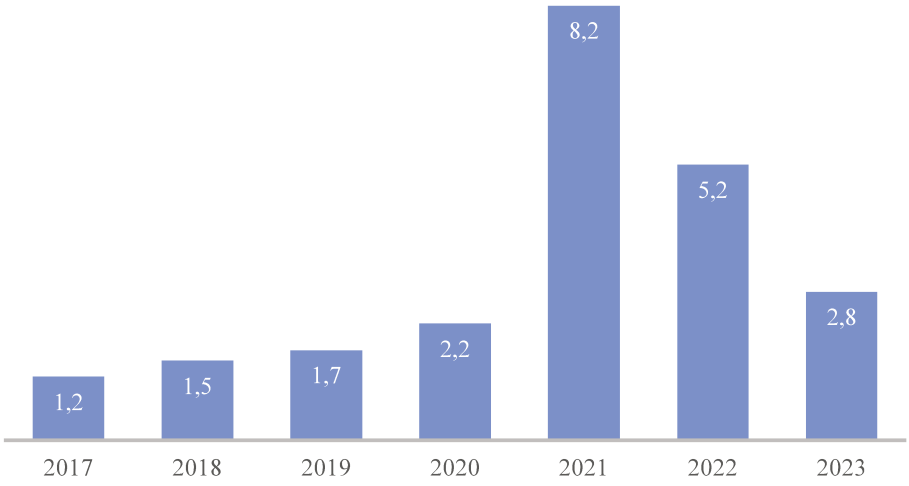


Figure 9. Levées de fonds totales réalisées par les start-ups EdTech *aux États-Unis (en milliards de dollars). Source : Wan, 2024.*

Les fondations philanthropiques représentent un autre guichet de financement pour les start-ups et les groupes plus anciens en quête de financements, dans des domaines d'activité en phase avec leurs missions sociétales. Ces financements sont moins substantiels que les participations des fonds de capital-risque, et s'élèvent à des dizaines voire des centaines de milliers de dollars pour une même firme. Ils prennent cependant la forme de dons et de subventions sans dilution du capital de l'entreprise. The Bill & Melinda Gates Foundation a par exemple accordé deux *grants* à DreamBox Learning, éditeur proposant depuis 2006 un EIAH adaptatif pour l'apprentissage des mathématiques en primaire. La première subvention date de 2015. D'un montant de 748 100$, elle consistait à « soutenir le développement d'un prototype de module de développement professionnel pour les enseignants » (Bill & Melinda Gates Foundation, 2024). La dernière en date, accordée en 2023 pour 97 250$, consistait à « tirer parti de l'intelligence artificielle dans la conduite de recherches et de développements à l'appui des résultats en mathématiques des étudiants noirs, latinos et/ou issus de milieux défavorisés » (Bill & Melinda Gates Foundation, 2024). On peut également citer les concours organisés par des fondations accordant des prix à des projets entrepreneuriaux (Encadré 2).

Encadré 2. La *Learning Tools Competition*

Ce concours international a été créé en 2020, en pleine pandémie, par le cabinet de conseil The Learning Agency. Soutenu par des fondations, agences fédérales et fonds de capital-investissement, son but est de « développer le domaine de l'ingénierie de l'apprentissage en stimulant les innovations technologiques dédiées à l'exploitation du *Big Data* pour soutenir la recherche en sciences de l'apprentissage et les besoins des apprenants du monde entier »[75] (Tools Competition, 2023). En 2023, plus de 4 millions de dollars issus de fondations et d'agences fédérales ont ainsi été répartis entre 32 gagnants, lors de deux concours. Le premier, axé sur l'enseignement primaire et secondaire, a été financé par les fondations Schmidt Futures, Walton Family Foundation, the Bill & Melinda Gates Foundation, ainsi que les fonds de capital-risque Citadel et AlleyCorp. Le second concours était consacré à l'apprentissage des adultes et soutenu par la *Defense Advanced Research Projects Agency* (DARPA). Si des firmes du monde entier ont remporté des prix – à l'image du français Evidence B en 2023 – la grande majorité des initiatives soutenues sont états-uniennes – 92 sur 130 lauréats depuis 2020 (Tools Competition, 2024).

[75] « *It aims to grow the field of learning engineering by spurring ed tech innovations dedicated to leveraging big data to support learning science research and the needs of learners worldwide* ».

Les inégalités entre start-ups demeurent très fortes, alors que les taux de mortalité, notamment lors de la première année, sont très élevés. Les « trajectoires » sont dès lors très diverses, entre les entreprises réalisant des levées auprès de fonds de capital-risque, celles touchant soit additionnellement soit exclusivement des subventions fédérales (*EIR grants, SBIR program*) ou de fondations, ou celles capitalisant sur des fonds privés (personnels ou provenant de proches). L'aboutissement idéal pour ces firmes est l'introduction en bourse, comme est parvenu à le faire Duolingo en 2021 (Encadré 3).

Encadré 3. Les évolutions successives de Duolingo

Ses fondateurs sont un professeur de *computer science*, Luis von Ahn, et son doctorant Severin Hacker, de l'université de Carnegie Mellon. Le premier a participé au développement du protocole d'identification en ligne ReCAPTCHA en 2007. La firme s'est financée à ses débuts par un prix de la fondation McArthur décerné à Luis von Ahn ainsi qu'une subvention de la *National Science Foundation*. En plus de ses algorithmes d'apprentissage adaptatif, une clé de son succès est sa pédagogie voulue « fun », inspirée du modèle BLAP – pour « *Badges, Levels/Leaderboards, Achievements, Points* », ou littéralement « Badges, Niveaux/tableaux de bord, Succès, Points » –, théorisé par Scott Nicholson (2012) : sur le modèle vidéoludique, l'application récompense l'élève ayant réussi un exercice par des points d'expérience, lui permettant de récolter des badges, monter de niveaux, et concourir dans des ligues face à des apprenants du monde entier.

La firme a connu un succès économique fulgurant, prenant rapidement les rênes du marché aux États-Unis, tout en levant des dizaines de millions de dollars tout au long des années 2010 auprès de fonds de capital-risque tels que A-Grade Investments (3,3 millions de dollars en 2011), Union Square Ventures (15 millions de dollars en 2012) et CapitalG, le fonds d'investissements de Google (45 millions de dollars en 2015 puis 30 millions de dollars en 2019). Se rémunérant à ses débuts par la traduction de textes par ses utilisateurs pour des entreprises clientes, l'entreprise a depuis adopté un modèle *freemium* : une offre gratuite avec publicités est complétée par deux offres *premium, Super Duolingo* et *Duolingo Max*. Elle a diversifié son offre dans la certification à distance en anglais, l'apprentissage des mathématiques ou de la lecture. La firme de Pittsburgh s'est ensuite lancée en bourse à l'indice du Nasdaq en juin 2021. En termes d'indicateurs économiques, elle domine le secteur par son chiffre d'affaires de 531,1 millions de dollars en 2023, en hausse de 44% par rapport à 2022 (Duolingo, 2024a, p.72). Elle possède également la plus large base d'utilisateurs et d'abonnés dans le monde, avec 88,4 millions d'usagers et 6,6 millions d'utilisateurs payants (ibid., p.65).

Les dépenses en recherche et développement destinées notamment à perfectionner son produit et son IA se sont élevées à 194,4 millions de dollars en 2023, soit environ 36,6% de son chiffre d'affaires (ibid., p.72). Le ratio est particulièrement important, même pour les standards traditionnellement élevés du secteur technologique : en comparaison, pour l'année fiscale 2021, Apple avait un ratio avoisinant 6 %, Amazon 12 %, Microsoft et Alphabet 12.3 % et Facebook 21 %, selon leurs rapports annuels respectifs.

Ces firmes se développent par des financements extérieurs et une mise sur le marché à partir d'une proposition de valeur centrée sur leurs algorithmes. Elles développent ces derniers soit par une expertise interne – lorsque les fondateurs sont ingénieurs de formation, ou que des informaticiens ont été recrutés dans les premières années de l'entreprise –, soit par leur liens avec des laboratoires.

Le concours de Big Tech

La plupart des start-ups capitalisent cependant sur les offres techniques des *Big Tech*, à commencer par leurs services de *cloud computing*, cruciaux pour la collecte, le stockage voire l'analyse de leurs données d'apprentissage. Le principal apport pour les *EdTech* est de pouvoir entraîner leurs algorithmes sur ces serveurs tout en réduisant les coûts de maintenance, et de s'adapter aux pics et aux creux de fréquentation. Microsoft et Google sont deux des principaux acteurs mondiaux de ce type d'hébergement. Selon une étude du Synergy Research Group, Amazon et son offre *Amazon Web Service* (AWS) dominaient le marché mondial des infrastructures *cloud* en 2023 avec 31 %

de part de marché, suivis de *Microsoft Azure* (24 %) et *Google Cloud* (11 %), quand les offres des poursuivants (Alibaba, Salesforce, IBM, Oracle) ne dépassaient guère les 5% (Synergy Research Group, 2024).

Les « géants des technologies » ne se contentent guère de gérer la collecte et la maintenance des services de start-ups *Edtech*. Au contraire, ces offres *cloud* se trouvent être des portes d'entrée pour la vente d'algorithmes œuvrant au traitement et à l'analyse des données d'apprentissage. Face aux difficultés de nombreuses start-ups pour développer et entraîner des algorithmes sophistiqués dans un temps limité, les *Big Tech* proposent une sous-traitance algorithmique concourant à l'adaptativité finale.

Sur son site, AWS donne l'exemple de son partenariat avec Knewton, start-up emblématique de l'*adaptive learning*, spécialisée aujourd'hui dans l'enseignement supérieur. Amazon explique lui avoir licencié diverses fonctionnalités liées à AWS (Figure 10.). On voit sur ce schéma que les fonctionnalités d'Amazon sont nombreuses dans l'infrastructure de Knewton. Elles alimentent directement ses modules d'apprentissage adaptatif situés dans le « *Private Subnets* » : « *analytics* » (« analytique ») ou encore « *Recs* » (« recommandations »). Témoin l'*Amazon Elastic Map Reduce* (ou « *Amazon EMR Cluster* » sur le schéma), servant à « analyser de grands ensembles de données sur la plateforme en parallèle afin d'adapter ses modèles de recommandation et analytiques » (Amazon Web Service, 2024b).

Amazon mentionne aussi sa sous-traitance technologique pour Duolingo, qui emploie l'architecture *PyTorch* pour développer ses propres algorithmes de *machine learning* :

> « Pour développer ces algorithmes personnalisés (de la reconnaissance vocale non native à la classification pour la notation automatisée), Duolingo utilise le *framework* de *deep learning PyTorch* sur *Amazon Web Services* (AWS). […] La rapidité et la capacité de mise à l'échelle sont essentielles pour entraîner les modèles. En effet, ces derniers peuvent utiliser entre 100 000 et 30 millions de points de données à la fois, selon le problème, afin de réaliser plus de 300 millions de prévisions par jour » (Amazon Web Service, 2024c) ».

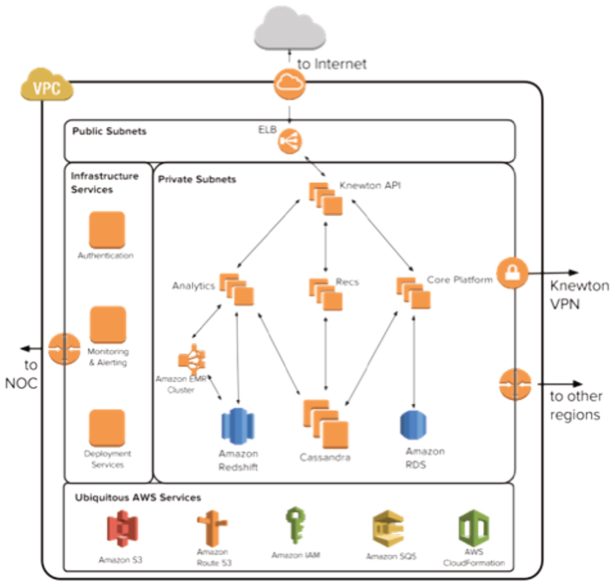


Figure 10. Schématisation de la plateforme de Knewton *selon Amazon. Source : Amazon Web Service, 2024b.*

Cet extrait montre par ailleurs le caractère de plus en plus stratégique du *cloud*, permettant de réaliser les très grands volumes de calculs nécessaires au *deep learning*. Amazon vend à Duolingo d'autres

outils d'IA aux fonctionnalités éducatives plus palpables pour l'utilisateur, tels qu'*Amazon Polly,* « outil de synthèse vocale basé sur le *deep learning* qui s'intègre facilement aux applications pour donner une voix au test et à de nombreux cours » (ibid.).

Le site de Google Cloud donne quant à lui l'exemple de Schoolytics, éditeur d'une plateforme de fourniture de *Learning Analytics*, qui « aspire » les informations des LMS pour fournir aux acteurs de l'éducation des tableaux de bord faciles à utiliser. Google lui a fourni son algorithme *Big Query* « pour optimiser l'accessibilité des données et étendre ses services afin de desservir davantage d'écoles » (Google Cloud, 2024b).

Les *Big Tech* visent ainsi à accroître leur influence sur les marchés éducatifs en ciblant les start-ups *EdTech*, initialement en demande de services d'hébergement. Ils constituent également des offres d'incubation proposant des crédits pour leurs services *cloud* et l'accès à des réseaux d'acteurs constitués autour d'eux. Témoin le *Cloud Google for Startups,* proposant jusqu'à 350 000 dollars pour les start-ups d'IA (Google Cloud, 2024c), et l'*AWS Education Accelerator* offrant quant à lui des crédits de 100 000 dollars pour tout type de start-ups (Amazon Web Service, 2024d). Ces « programmes d'accélération » doivent non seulement permettre aux *Big Tech* de trouver de nouveaux clients pour leurs suites technologiques, mais aussi d'observer les avancées du secteur et cerner les opportunités de rachat. Certains d'entre eux ont acquis des entreprises d'IA éducative. En 2018 Google a fait l'acquisition de Socratic. Cette application mobile de soutien scolaire permet à des élèves de poser des questions ou d'envoyer des photographies de leurs exercices à une IA, qui lui explique en retour les étapes de résolution[76].

Les acteurs natifs du numérique au défi de la réinvention

De grands groupes états-uniens du numérique éducatif, antérieurs à l'engouement pour l'*adaptive learning*, jouent également un rôle dans le développement de ce type de technologies. Pour ces groupes concentrés surplombant une multitude de petites entreprises fragiles, l'engouement pour l'IA éducative est devenu à la fois une opportunité et une menace de « disruption », sur laquelle il a fallu se positionner. Ceux-là partageaient déjà, avant les années 2010, une proposition de valeur à rebours de l'enseignement collectif et standardisé, centrée sur l'individualisation de l'éducation. Leurs produits numériques initiaux n'ayant pas été conçus suivant un agenda adaptatif, et leurs équipes d'ingénieurs n'étant pas spécialistes de ce type d'algorithmes, leur transition vers l'adaptativité s'est faite en externe. Ces groupes ont en effet pratiqué des rachats de start-ups pour absorber leurs produits dans leurs catalogues. Représentatifs sont les positionnements de deux acteurs à quelques centaines de millions de dollars de chiffres d'affaires : Renaissance Learning et Cambium Learning Group. Tantôt cotés en bourse, tantôt rachetés par des fonds de *venture capital*, ceux-ci se sont constitués par vagues de fusions-acquisitions au gré d'une financiarisation accélérée.

Renaissance Learning a été créé en 1986. La firme vend alors des livres numériques pour enfants – un marché émergent grâce à la démocratisation des ordinateurs personnels. Ces livres ont pour spécificité d'inclure un test de compréhension pour l'enfant à la fin de sa lecture ainsi qu'un diagnostic détaillé. Cette proposition, non-adaptative, est cependant une forme d'individualisation qui confère à cet acteur un positionnement sur ce créneau. Celui-ci sera réinvesti deux décennies plus tard lors de sa transition vers l'IA. Entre-temps, l'entreprise ouvre son capital,

[76] Basé sur une ontologie de concepts, ce système envoie à l'élève les étapes de résolution ainsi que des contenus audiovisuels de remédiation voire d'approfondissement. L'intégration est aujourd'hui réalisée avec les services de Google, l'entreprise employant ses algorithmes d'OCR (*Optical Character Recognition*), de NLP (*Natural Language Processing*), et des contenus comme des vidéos YouTube.

d'abord par une entrée en bourse en 1997. La firme est ensuite rachetée successivement en 2011 et 2014 par deux fonds d'investissements. Parmi les actionnaires, Capital G, le fonds de Google, prend des participations significatives. La manne financière accumulée par Renaissance Learning lui permet de mener ensuite des rachats de start-ups positionnées sur l'*adaptive learning*. La firme commence par cibler des services modulaires de *Learning Analytics* :

- D'abord Subtext en 2013, plateforme permettant, selon l'analytique, de recommander à l'enseignant des groupes d'élèves ;
- Schoolzilla en 2019, outil de visualisation de *Learning Analytics* dans des tableaux de bords adaptatifs ;

En 2021, le groupe franchit un palier avec le rachat de deux firmes éditrices de systèmes de tutorat : Freckle pour l'apprentissage des mathématiques et de l'anglais en classe, et la start-up française Lalilo pour celui de la lecture.

Autre grand groupe du numérique éducatif des années 2000, Cambium Learning a été créé plus tard que Renaissance, en 2003. Cette holding acquiert l'année suivante Sopris West, l'éditeur d'un *Learning Management Systems* spécialisé dans l'enseignement de l'algèbre. En 2005 elle jette son dévolu sur Kurzweil Education, entreprise fondée par le chantre du transhumanisme Raymond Kurzweil, éditant un logiciel de conversion de supports écrits en documents sonores pour les personnes malvoyantes. En 2018, la société, devenue Cambium Learning Group suite à sa fusion avec le groupe d'*e-learning* Voyager Learning, est rachetée par le fonds d'investissements Veritas Capital.

En 2020, grâce à ces apports de capitaux, l'entreprise rachète Lexia, entreprise pionnière des médias adaptatifs, fondée en 1984 au début des recherches sur les *intelligent tutoring systems*, et spécialisée dans l'enseignement de la lecture aux enfants.

Ainsi la conception d'algorithmes d'apprentissage adaptatif est-elle un poste de dépense important, distinguant les start-ups selon leurs capacités à lever des fonds pour financer leurs recherches. Certaines ont recours aux divers services technologiques des *Big Tech*. Alors que Google a racheté l'une d'entre elles, la stratégie d'acquisition de start-ups est particulièrement adoptée par des éditeurs anciens du numérique éducatif, qui intègrent ainsi des produits adaptatifs à leurs catalogues.

Sur les contenus : des fédérations entre structures de production informatique et éditeurs de contenus encore timides

Comme on l'a vu en première partie, les algorithmes d'apprentissage adaptatif fonctionnent mal s'ils ne sont pas couplés à un éventail de contenus pédagogiques divers et variés, tant en termes de niveaux de difficulté que de formats. La production de contenus de qualité satisfaisante pour épouser les attentes des institutions et des publics éducatifs est un tout autre métier que celui du développement informatique dans lequel sont spécialisées les start-ups. Il suppose des compétences en termes d'écriture, d'édition et d'expertise pédagogique, et dès lors des coûts humains et techniques importants. De telles contraintes jouent en faveur du rapprochement des petites structures de production informatique avec des acteurs emblématiques des contenus pédagogiques : les éditeurs scolaires.

Ceux-ci constituent un oligopole à frange très concentré, parmi les plus importantes industries éducatives en termes de chiffre d'affaires. Après avoir opté pour la transition de leurs catalogues vers

le numérique, les grands groupes états-uniens sont désormais tentés de s'engager dans l'*adaptive learning*, qui représente un nouveau relai de croissance. Cependant le défi éditorial est de taille, car il s'agit non plus de numériser des manuels mais de modulariser l'ensemble des contenus, pour permettre à des algorithmes d'adapter des parcours individuels. Or ces groupes comptent peu ou pas d'ingénieurs rompus à ces technologies de pointe, et leur grande taille favorise une inertie organisationnelle incompatible avec la conception expérimentale d'algorithmes. Joue en revanche en leur faveur leur puissance financière, *a fortiori* pour les groupes dominants l'oligopole du haut de leurs milliards de chiffres d'affaires annuels. Les spécialisations respectives des start-ups et des éditeurs scolaires poussent donc à leur fédération. Celle-ci s'est réalisée aux États-Unis sous deux formes principales : la commande de « modules adaptatifs » par les éditeurs auprès de start-ups, et, comme dans le cas des éditeurs du numérique éducatif, l'acquisition de start-ups.

La sous-traitance de modules adaptatifs

Par « module », nous entendons une déclinaison d'un algorithme, adaptée et « branchée » au *Learning Management System* de l'organisation cliente afin de le rendre adaptatif sur un aspect pédagogique souhaité (les évaluations ou les *feedbacks* notamment). Les contrats entre des éditeurs développant leurs LMS et des start-ups sous-traitant leurs technologies ont été particulièrement courants dans les débuts de l'*adaptive learning* aux États-Unis. Pour les start-ups qui n'ont pas les moyens d'investir dans la création d'ITS intégraux, ce mode de vente a l'avantage de représenter une manne financière importante, de même que l'opportunité d'entraîner leurs algorithmes sur des publics vastes. Pour les éditeurs, cette sous-traitance leur permet de se positionner à moindre coût dans le marché et de suivre ses avancées techniques.

Dans la formation professionnelle, où les organismes de formation éditent des contenus pour des secteurs professionnels spécifiques, ce type de sous-traitance est courant. Ainsi l'entreprise danoise Area9 Lyceum, à travers sa division Area9 Learning spécialisée dans la formation professionnelle, distribue-t-elle ses algorithmes sous licence à des organismes de formation souhaitant rendre leurs LMS adaptatifs. Parmi les nombreux exemples recensés sur le site de la firme, il y a son contrat passé avec l'American Diesel Training Centers (ADTC), formant des mécaniciens pour le transport routier (Area9 Lyceum, 2024). Cette sous-traitance a débouché sur l'*ADTC Skills App,* un ITS d'évaluation formative des compétences du mécanicien, incluant un espace de *monitoring* pour le formateur.

Knewton s'est également développé comme sous-traitant pour plusieurs éditeurs à ses débuts. L'entreprise fondée en 2008 et spécialisée dans l'enseignement supérieur a connu un engouement financier au début des années 2010, levant au total plus de 180 millions de dollars. Figure de proue de l'*adaptive learning* dans l'espace médiatique, elle a alors contracté avec de grands éditeurs comme Pearson, qui a participé à la levée de 33 millions de dollars réalisée par l'entreprise en 2011 (McKenzie, 2019).

L'acquisition de start-ups

Dans l'enseignement primaire, secondaire et supérieur où le marché de l'édition est oligopolistique, cette voie s'est avérée être une impasse quand la plupart des grands éditeurs se sont détournés de la sous-traitance pour enclencher une autre forme de fédération : l'acquisition de start-ups. Le *modus operandi* est semblable à celui employé par les grands groupes du numérique éducatif vus plus haut,

mais avec, pour certains, des moyens plus forts, leur permettant de se positionner sur différents niveaux d'enseignement.

Cette stratégie est particulièrement visible outre-Atlantique, où six acteurs ont mené des acquisitions à partir de 2013. McGraw Hill est par exemple l'un des plus actifs dans ce type de rachats. C'est lui qui a effectué en 2013 le premier rachat d'une entreprise des ITS, ALEKS, opérant à la fois dans l'enseignement primaire, secondaire, et supérieur. En 2014, le groupe a acquis l'entreprise Area9 Lyceum évoquée plus haut, auprès de laquelle McGraw Hill a auparavant sous-traité le développement algorithmique de son ITS « maison », *LearnSmart*. Deux rachats sont advenus en 2021, cette fois pour un positionnement dans l'enseignement primaire : celui de la start-up Kidaptive, et celui du groupe d'*e-learning* Achieve 3000, propriétaire de la start-up SmartyAnts spécialisée dans l'alphabétisation.

Cette stratégie d'acquisition de start-ups liées à l'*adaptive learning* se démultiplie dans la seconde moitié des années 2010, comme le révèle notre tableau récapitulatif en annexe 2. Elle a affecté grandement les structures de production s'étant développées comme sous-traitants. En 2017, lorsque Pearson annonce ne pas reconduire la plupart de ses contrats avec Knewton, s'enclenche la fin de l'âge d'or de la start-up. Elle sera elle-même rachetée deux ans plus tard par un éditeur spécialisé dans l'enseignement supérieur, Wiley, pour un montant évalué à 17 millions de dollars – une somme bien en deçà des 180 millions de dollars levés par l'entreprise new-yorkaise depuis sa création (Wan, 2019).

Le géant de l'édition britannique Pearson entame pour sa part sa restructuration vers l'enseignement supérieur : il vend sa division *K-12* au fonds d'investissement Nexus Capital Management en 2019, jugeant que ses produits scolaires sont les moins rentables de son portfolio (Cavanagh, 2019). Prétendant se concentrer sur le développement de technologies adaptatives en interne – perspective illusoire au vu de son manque d'expertise en la matière –, Pearson noue en fait un partenariat avec IBM pour incorporer son agent conversationnel Watson à ses outils pour l'enseignement supérieur (Wan, 2017).

Se positionnant de plus en plus directement dans le marché de l'adaptativité, les éditeurs ont ainsi l'occasion de ne pas laisser les start-ups s'adjuger seules un ethos innovant. À des fins de *branding*, certains groupes adoptent une stratégie de production de discours à vocation scientifique sur l'*adaptive learning*. Ainsi Pearson édite-t-il en 2016 un livre blanc sur le secteur rédigé par le média spécialisé Edsurge, intitulé « *Decoding Adaptive* ». Pour présenter la collection « *Open Ideas* » dans laquelle l'ouvrage est édité, il est écrit :

> « Comment apprenons-nous et qu'est-ce qui nous motive à le faire ? Quel est l'ensemble des connaissances et des compétences dont les apprenants ont besoin alors que nous entrons dans la seconde moitié du XXIe siècle ? Comment les technologies numériques intelligentes peuvent-elles être mieux déployées pour réaliser l'objectif d'une éducation plus personnalisée ? Comment pouvons-nous construire des systèmes éducatifs qui offrent à tous des opportunités pour un apprentissage de haute qualité ? Ces questions sont trop importantes pour que les meilleures idées restent uniquement dans l'amphithéâtre, sur l'étagère ou dans une salle de classe. Elles méritent plutôt d'être découvertes et soutenues, partagées et débattues, adoptées et affinées. Nous espérons qu'« *Open Ideas* » vous aidera dans cette tâche et que vous vous joindrez à la conversation »[77] (Edsurge, 2016, p.3).

[77] *« How do we learn, and what keeps us motivated to do so? What is the body of knowledge and skills that learners need as we move into the second half of the 21st Century? How can smart digital technologies be best deployed to realize the goal of a more personalized education? How can we build education systems that provide high quality learning opportunities for all? These questions are too important for the best ideas to stay only in the lecture*

Ainsi l'entreprise se pose-t-elle en révélateur d'idées innovantes dans l'espace public, dont celles portant sur l'IA éducative. Publiant en licence libre ce livre blanc, la firme projette d'une part sur cette technologie un ethos innovant, au service du bien commun, liée au paradigme du *personalized learning*. Elle tente d'autre part, en se mettant en scène comme un acteur de sa diffusion sociale, de capitaliser sur cet ethos.

Les tentatives d'édition d'ITS par des start-ups

Face à l'impasse et la difficulté de ces fédérations, de nombreuses start-ups tentent de vendre directement aux organisations éducatives des systèmes de tutorat, incluant contenus, architecture algorithmique et interface. Les acteurs entrent alors dans une concurrence frontale avec les éditeurs et leurs LMS sur ce segment numérique. Cette diversification des métiers, en particulier dans la production et l'édition de contenus, entraîne des coûts d'autant plus grands. Les entreprises ont plusieurs solutions. Une première est de recruter des équipes pédagogiques en charge de la conception des contenus, grâce à des financements externes. C'est la voie choisie par Duolingo, leader des applications d'apprentissage des langues, par exemple.

Une deuxième solution est le recours aux ressources éducatives libres. Cela a été la voie choisie par Knewton pour mettre en branle son nouvel ITS « maison » en 2017. Devant le volte-face de son actionnaire et client Pearson, Knewton décide d'abord de vendre son propre outil aux organisations éducatives. Cet ITS fonctionne non pas avec des contenus écrits par des équipes pédagogiques, mais avec des ressources libres de droits glanées sur internet. Le travail humain de sélection de ces ressources et d'éditorialisation pousse toutefois à relativiser les économies réalisées. De plus, la faible variation des formats pédagogiques ne permet pas à l'outil d'apparaître comme un concurrent du manuel, comme tendent alors à le faire croire les dirigeants de la firme. Il s'agit plutôt d'un service d'aide aux devoirs à la maison relativement rigide et redondant dans les exercices qu'il assigne aux élèves (Young, 2019).

Une troisième solution est représentée par l'essor des grands modèles de langage (*large language models*), permettant de générer des contenus. Duolingo a annoncé au début de l'année 2023 le lancement de deux agents conversationnels, *Explain My Answer* et *RolePlay*. Le premier fournit à l'utilisateur des explications sur ses erreurs, et le second l'occasion de converser dans des situations variées. Ils ont été développés en partenariat avec l'éditeur OpenAI, à partir de son grand modèle de langage *GPT-4* (Duolingo, 2023).

Cette voie donne un pouvoir accru dans la filière numérique à un groupe restreint de multinationales des technologies ayant développé des LLM, dont certaines *Big Tech*. Les coûts fixes pour entraîner un LLM sont en effet particulièrement importants, notamment en termes d'équipement en processeurs graphiques ou de consommation d'énergie des serveurs. Seules quelques multinationales ayant une puissance financière ainsi qu'un savoir-faire technologique suffisants sont capables de développer ces modèles et de les commercialiser. Cet avantage les situe comme des prestataires de premier choix pour les start-ups désirant générer des contenus pédagogiques. Outre le partenariat d'OpenAI avec Duolingo, Google distribue également son agent Gemini dans son offre *cloud* à destination des start-ups.

theatre, on the bookshelf, or alone in one classroom. Instead they need to be found and supported, shared and debated, adopted and refined. Our hope is that Open Ideas helps with this task, and that you will join the conversation ».

Cependant, l'IA générative se heurte encore à des limites, qui recoupent en partie celles rencontrées par les chercheurs et évoquées en première partie. Du point de vue des acteurs socio-économiques, la première est la confidentialité des données entrées par les apprenants, enseignants et organisations éducatives dans les grands modèles de langage, devenant des données d'entraînement pour les LLM. Dans le cas de données potentiellement sensibles comme celles d'apprentissage, nombre d'organisations sont réfractaires à l'idée d'intégrer ces agents conversationnels à leurs pratiques pédagogiques. Dans une apostrophe aux *EdTech*, Google anticipe ces inquiétudes en leur faisant cette invitation menant directement à Gemini : « Appuyez-vous sur de grands modèles de langage (LLM) de haute qualité pour générer automatiquement du code, du texte, des contenus multimédias et plus encore, tout en respectant la confidentialité » (Google Cloud, 2024a).

Le deuxième obstacle est l'impossibilité pour ce type d'IA de produire des contenus conformes aux programmes ou aux attentes pédagogiques des institutions clientes si le modèle n'est pas entraîné sur une base de données spécialisée, comme un répertoire de manuels et de contenus certifiés. Le *fine-tuning* nécessite donc en amont la constitution d'une base de ressources éducatives fiables et mises à jour. Si cette technologie ne semble pas à court terme en mesure de ravir aux éditeurs scolaires leur avantage concurrentiel, elle reste toutefois un domaine d'investissement privilégié par des firmes de l'*EdTech*, accroissant l'emprise des *Big Tech* sur la filière.

La distribution des médias adaptatifs

En aval de la filière, la mise à disposition en SaaS des médias adaptatifs auprès des publics éducatifs est un enjeu crucial à la fois en termes commerciaux, pour entretenir un contact avec les usagers, et en termes techniques, pour contrôler les modalités de collecte de données nécessaires à l'entraînement des algorithmes. Selon les niveaux d'enseignement et les industries éducatives, deux circuits de distribution peuvent être classiquement distingués :

- Les licences de logiciels et services web, par usager ou formateur, vendues aux districts, établissements du supérieur et employeurs disséminés sur le territoire américain ;
- La mise à disposition d'applications, directement auprès des consommateurs sur leurs terminaux mobiles, soit gratuitement avec achats intégrés et/ou publicités ciblées, soit par abonnement.

Dans le circuit institutionnel, la vente de licences aux institutions s'apparente au modèle du club. L'acteur central percevant les revenus est l'éditeur. Dans le second circuit, la distribution d'application s'apparente au modèle du courtage : l'acteur central est la plateforme d'intermédiation (en l'occurrence le magasin d'applications), percevant des redevances sur les abonnements et achats intégrés réalisés entre les éditeurs et les consommateurs. Ce modèle tend cependant à gagner les circuits institutionnels avec l'arrivée de plateformes privées (*Clever* dans l'enseignement primaire et secondaire, *Microsoft Teams* dans l'enseignement supérieur et la formation professionnelle) cherchant à s'installer comme des places de marché entre les organisations éducatives et les éditeurs. Elles prétendent répondre aux besoins des publics éducatifs, notamment par des garanties en termes de confidentialité des données transmises aux éditeurs ou une promesse de simplicité via une connexion unique. Elles promettent également aux firmes éditrices une simplification de leur accès aux districts et aux organisations disséminées sur le territoire états-unien. Avec la multiplication des usages et des données collectées sur les médias adaptatifs, ces acteurs de l'intermédiation sont dès lors susceptibles de jouer des rôles importants et de structurer une socio-économie cohérente.

La reconduction sous contraintes de la vente de licences

Dans le premier circuit, les secteurs éducatifs institutionnels, de l'enseignement primaire à la formation professionnelle, sont souvent complexes à pénétrer pour les firmes éditrices. D'une part, la distinction entre l'acheteur (l'institution), l'usager (l'enseignant, l'apprenant voire le parent) et le prescripteur (comme les agences éducatives des États fédérés) crée une complexité aiguë pour l'éditeur, qui doit à la fois mobiliser des arguments commerciaux, éducatifs et politiques pour satisfaire ces pôles de réception.

Les niveaux primaire et secondaire offrent un exemple éloquent de cette configuration. L'acheteur est le « *district intendant »*, préoccupé par des objectifs financiers et d'efficacité pédagogique, alors que les usagers – élèves, enseignants voire parents – sont sensibles à des enjeux pédagogiques et pratiques. S'y ajoutent les prescripteurs, qui sont à la fois les *State Education Agencies,* formulant, dans le cas des *adoption states,* des listes d'éditeurs approuvés à destination des districts, et les enseignants, faisant remonter des besoins aux chefs d'établissements et de *districts*. Pour les offreurs, les attentes à prendre en compte sont donc variées.

D'autre part, l'organisation décentralisée de ces échelons éducatifs entraîne un éclatement des acheteurs que la firme doit convaincre personnellement, ce qui la conduit à engager des frais commerciaux importants. Aux niveaux primaire et secondaire, la dissémination des districts et des établissements est une contrainte logistique de taille pour les firmes les plus petites. Devant plus de 13 000 districts, des équipes importantes de vendeurs sont nécessaires pour sillonner l'immensité du territoire américain à la rencontre des *district intendants*, et jouer des coudes au sein d'une concurrence âpre pour des marchés seulement locaux. Cette contrainte commerciale avantage les acteurs aux capacités financières plus grandes, comme les éditeurs scolaires (McGraw Hill et consorts) ou les groupes concentrés du numérique éducatif (Renaissance Learning ou Cambium Learning Group).

Traditionnel dans les logiciels éducatifs, le modèle du « club », opéré via des licences annuelles par élève ou enseignant, permet aux firmes éditrices d'engranger directement des revenus substantiels « sécurisés » pour au moins une année. Ceux-ci peuvent atteindre plusieurs dizaines de milliers de dollars pour l'équipement de tous les élèves d'un établissement. La double difficulté consiste à concilier la satisfaction des publics éducatifs, pour que le district reconduise les licences l'année suivante, et la désirabilité de l'offre auprès des prescripteurs que sont les *State Education Agencies* et les enseignants, afin de rallier un plus grand nombre de districts et d'établissements clients.

Outre le modèle du club, la gratuité est également un levier socio-économique fort en éducation, valorisable sur deux plans :

- D'abord idéologique, les firmes insistant sur leur engagement politique au service d'un accès démocratique à l'éducation ;
- Et économique, lorsque la gratuité sert de porte d'entrée vers des licences payantes, suivant un modèle *freemium.*

Lalilo offre un exemple éloquent des arbitrages que doit effectuer une start-up pour s'adresser à ces divers prescripteurs, ainsi que des logiques de concentration qui régissent la filière. La firme s'est développée en 2016 à partir du *freemium* aux États-Unis (Encadré 4). L'ensemble de l'application était à l'époque gratuit et les barrières se situaient du côté de l'interface « enseignant », dans

les tableaux de bord où certaines informations n'étaient pas disponibles. Les enseignants devaient jouer un rôle de prescripteur au profit de Lalilo, en demandant des licences auprès de leurs chefs d'établissements. Autrement dit, le *freemium* visait à la fois à « séduire » l'enseignant par les fonctionnalités générales de l'outil, et à le « frustrer » à travers les barrières installées, pour *in fine* le conduire à demander un abonnement à sa hiérarchie. En cela, la firme se distinguait à l'époque de ses concurrents non-adeptes du *freemium*, comme Lexia, qui comptent convaincre directement les *districts intendants*. Le *freemium* lui permettait aussi d'économiser des frais de recrutement en commerciaux pour pénétrer les districts.

Depuis son rachat par Renaissance Learning en 2021, la firme a abandonné le *freemium* pour passer à une formule d'essai gratuit limité dans le temps de la version intégrale (dite aussi « *free trial* »). Par ailleurs, la firme a renoncé au *licensing* par enseignant pour vendre ses services par élève, le modèle en vigueur chez Renaissance, comme nous l'a confié le co-fondateur de l'entreprise Laurent Jolie (Entretien 4)[78]. Ainsi le rachat de cette start-up innovante par sa stratégie de distribution a conduit à sa standardisation commerciale au sein du catalogue de la maison-mère. Concession faite à son modèle centré sur les enseignants, la gratuité a été conservée au sein d'un essai gratuit, au terme duquel les enseignants peuvent continuer à prescrire l'outil auprès de leur hiérarchie. Cependant, l'entreprise profite de l'équipe de commerciaux étendue de Renaissance Learning pour s'adresser dorénavant aux districts. Son produit est depuis vendu en « *cross-selling »* avec d'autres médias éducatifs de sa maison-mère.

Encadré 4. Lalilo, une start-up française ancrée aux États-Unis

Cette start-up d'origine française est créée en 2016 par trois étudiants de l'école Polytechnique, Laurent Jolie, Amine Mezzour et Benjamin Abdi. Les trois camarades développent un outil d'apprentissage de la lecture d'abord en français puis en anglais. Ils migrent dès la fin de l'année à San Francisco pour participer au programme d'accélération *Refiners* fondé par des entrepreneurs français. Cela leur permet de lever 1,1 millions d'euros auprès des fonds d'investissements Kima Ventures et Nestadio Capital ainsi que des *business angels*, puis de remporter un prix de la BPI d'un montant de 500 000 euros (Lalilo, 2017). Incubés au sein de l'incubateur de la ville de Paris, *Agoranov*, ils recrutent une équipe partagée entre la France et les États-Unis, et commencent à commercialiser la première version de Lalilo en *freemium* à destination des professeurs. Cet espace en ligne consiste en un exerciseur mis en forme de façon ludique : l'enfant s'exerce en progressant dans des « mondes » différents et en remportant des récompenses diverses (badges, trésors, histoires). En 2019 la firme se développe fortement en France en remportant un appel à projets du Ministère de l'Éducation nationale axé sur l'IA (P2IA). Elle lève la même année 5 millions d'euros auprès des fonds d'investissements français Partech, Educapital et Citizen Capital. La pandémie de Covid-19 accélère grandement la croissance de sa base d'utilisateurs, tant aux États-Unis, son principal marché, qu'en France. Selon les chiffres que nous a confiés Laurent Jolie pour 2021, Lalilo compte alors environ 50% de ses utilisateurs aux États-Unis, 25% en France, 15% au Canada et 10% dans le reste du monde (Entretien 4). La même année, la firme est rachetée par Renaissance Learning.

En effet les groupes concentrés issus de l'édition scolaire ou directement du numérique éducatif capitalisent sur leurs portfolios de produits pour pratiquer cette stratégie de vente. Le *cross-selling* (ou « vente croisée ») consiste à vendre des licences de plusieurs produits, combinées dans un « *pack* », à un district. Le groupe McGraw Hill peut pratiquer par exemple une distribution groupée incluant des produits de *literacy* pour différentes classes – comme *Smarty Ants* pour l'alphabétisation des élèves de *Pre-K* (équivalent de la moyenne section de maternelle) à *K-2* (CE1) ou *Achieve3000 Literacy* pour la compréhension écrite des élèves de *K-2* à *K-12* (terminale) – et d'autres produits de mathématiques, comme ALEKS pour les élèves de *K-3* (CE2) à *K-12* (McGraw Hill, 2023).

[78] « On vendait par prof, puis depuis qu'on a été rachetés on s'est alignés sur un modèle par élève, qui est un modèle majoritaire de la boîte qui nous a rachetés » (ibid.).

Scholastic, éditeur spécialisé dans les livres jeunesse, opère également cette stratégie de vente incluant son ITS d'alphabétisation des enfants, *Scholastic FIRST* (Encadré 5).

Cette stratégie de vente permet aux éditeurs d'une part d'augmenter les revenus amassés lors d'une transaction, et d'autre part de réduire l'incertitude quant à la satisfaction des usagers, en mobilisant ce que les chercheurs des industries culturelles ont théorisé comme la « dialectique du tube et du catalogue ». Trait saillant des filières adeptes du modèle éditorial comme le cinéma (Miège, 1984), elle permet à des éditeurs de maximiser la rentabilité globale de leur catalogue à partir d'un ou de quelques « tubes » qui dégagent seuls des bénéfices. Ceux-ci finançant à la fois leur production et celle du reste du catalogue non-rentable. On peut interpréter le « *cross-selling* » comme une manière de « solidariser » les produits d'un catalogue. C'est cette logique que certains districts souhaitent défaire en choisissant des produits spécifiques au sein de portfolios, comme on l'a vu avec l'initiative du *Baltimore County Public Schools* (Encadré 1, p.34).

Encadré 5. *Scholastic FIRST*

Lancé en 2008 sous le nom « Ooka Island », ce produit vidéoludique propose à l'élève d'explorer une île peuplée d'elfes, qu'il contrôle au gré de mini-jeux. Des récompenses sont aussi prodiguées à travers un temps de « récréation » en fin de session, durant lequel l'enfant peut acheter des accessoires pour son avatar. La firme éponyme a été créée par deux enseignants dont une spécialiste de l'apprentissage de la lecture, Kay McPhee, dont la méthode est la base théorique du produit. En 2017 l'entreprise a été rachetée par Scholastic. Dès lors le produit a été renommé Scholastic FIRST (*Foundations in Reading, Sounds and Text*). Sa commercialisation suit le modèle socio-économique du « club » : il est vendu par abonnement aux districts, soit indépendamment, soit par *cross-selling* au sein d'un *pack*, *Scholastic Digital Manager*, incluant au choix les trente produits digitaux de la maison-mère. Il est disponible sous forme de SaaS ou d'applications mobiles et tablettes.

De nombreux médias adaptatifs « institutionnels » empruntent cette voie socio-économique traditionnelle, régie par le modèle du club, qui présente des barrières à l'entrée importantes pour des petites structures aux budgets commerciaux restreints. Convaincre les prescripteurs que sont les districts ou les enseignants est un impératif pour les firmes de l'*EdTech* mêlant des argumentaires divers résonnant avec les agendas de ces acteurs : efficience pédagogique, garantie en termes de protection des données personnelles, ergonomie et facilité d'utilisation, interopérabilité avec les systèmes numériques d'information, identification simple pour des élèves en bas âge notamment… Prétendant répondre à ces contraintes logistiques et pratiques, des intermédiaires se sont récemment imposés dans la filière. Ils tendent à faire advenir une autre socio-économie abaissant des barrières à l'entrée pour les plus petites structures, et instituant le modèle du courtage.

Une voie socio-économique consacrant les plateformes d'intermédiation

En 2015, l'IMS Global Learning Consortium publie le protocole *OneRoster*, permettant l'échange de données d'identification d'élèves et de leurs données d'apprentissage entre les systèmes d'information des établissements (*Student Information Systems* ou *SIS*) et les firmes éditrices. L'objectif de cette démarche est de simplifier la création et l'actualisation des comptes d'utilisateurs de part et d'autre – malgré le changement d'établissements ou de classes d'un élève par exemple (1EdTech, 2024). Intégrant ce protocole, des plateformes marchandes proposent de devenir des centrales de connexion unique (« *Single Sign-On* ») aux médias éducatifs achetés par un établissement. Elles offrent en outre des tableaux de bord réunissant les applications disponibles pour les élèves et les enseignants, ainsi que des interfaces pour les administrateurs, leur permettant de suivre les usages faits de chaque application. À partir de ces deux propositions de valeur tournées

vers l'interopérabilité et une administration voulue efficiente de l'enseignement, un marché dynamique s'est développé aux États-Unis, dominé par deux firmes privées.

La première est Clever. Celle-ci relie, en septembre 2024, près de 3 100 éditeurs *EdTech* à 77% des établissements du primaire et secondaire américains, selon les chiffres de la firme (Clever, 2024a). À l'origine, c'est une start-up créée en 2012 par trois camarades diplômés de Harvard en mathématiques et *computer science*. Ils constatent à l'époque la profusion de données d'identification à gérer par les districts, ainsi que la difficulté des éditeurs à raccorder leurs solutions aux systèmes numériques d'information de chaque district (Bosmeny, 2017). Ils conçoivent alors une plateforme de transmission de données d'identification d'élèves (ou « *rostering* ») et d'authentification unique (*Single Sign-On*). Les publics scolaires se connectent en passant par l'interface de Clever, qui collecte les données d'identification et d'apprentissage et les partage entre les éditeurs *EdTech* et les districts selon un cadre voulu sécurisé.

Clever propose aussi des interfaces de *monitoring* à destination des enseignants pour superviser en un seul endroit les métriques d'apprentissage de leurs élèves, et des administrateurs d'écoles et de districts, pour contrôler l'intensité des usages faits des médias contractés. Les éditeurs de ces médias peuvent en outre consulter ces données d'usage dans un cadre contrôlé par Clever : nulles autres données que celles d'apprentissage ne peuvent être transmises et vendues à des acteurs tiers, conformément aux engagements pris par la firme en matière de sécurité et de confidentialité de l'apprentissage (Clever, 2024b).

Clever s'est imposé comme intermédiaire principal de l'administration du *e-learning* en deux temps : d'abord en mobilisant l'argument de la gratuité à destination des districts, pour maximiser les usages des publics. Des formules payantes leur sont proposées en complément pour des aides à l'implémentation d'outils numériques ou la formation d'enseignants et de personnels administratifs. La deuxième stratégie de Clever est la vente de licences annuelles aux éditeurs pour leur permettre à la fois de profiter des services de partage d'informations avec les SIS (*rostering*) et de figurer dans son catalogue. Ces licences prétendent répondre au travail fastidieux que doivent effectuer les éditeurs pour connecter leurs produits aux SIS de chaque district. Surtout, Clever compte les convaincre d'intégrer son catalogue suivant sur un effet de réseau : le nombre de districts présents sur Clever doit rendre la présence des éditeurs indispensable et leur faire accepter de payer ces licences.

Clever propose en outre un *App Store* permettant aux districts d'effectuer des achats parmi le catalogue des éditeurs présents sur Clever, en passant par le système de paiement tiers de Clever. La firme ne précise pas si elle pratique des redevances sur ces licences, mais l'ajout de ce magasin d'applications tend à conforter la thèse de l'application du modèle du courtage au marché institutionnel.

En découle un positionnement de l'entreprise au service des éditeurs, avec la promesse de prendre en charge leur distribution auprès des districts de tout le pays. Eu égard aux contraintes du circuit traditionnel évoquées plus haut pour les petites structures, l'argument est de taille. L'un des cofondateurs confirme que le catalogue a d'abord intégré les petites start-ups de l'*EdTech* (Bosmeny, 2017). Selon lui, les éditeurs scolaires seraient arrivés dans un second temps, attirés par l'effet de réseau et la simplification entraînée par les outils d'identification unique.

ClassLink est le principal concurrent de Clever. La firme distribue une plateforme d'intermédiation à destination des districts et des publics éducatifs, *Launchpad*. Celle-ci permet une authentification unique ainsi qu'un tableau de bord pour les enseignants. Les administrateurs peuvent consulter un catalogue de produits *EdTech* adhérents à la plateforme, et contrôler les usages que font les élèves des applications déjà contractées. L'entreprise distribue également la fonctionnalité *StateLink* à destination des agences éducatives des *adoption states*, pour leur permettre de distribuer des médias agréés directement aux établissements et de superviser les usages à l'échelle de l'État. Ainsi, ClassLink se positionne davantage au service des institutions éducatives et de leurs publics (Encadré 6).

Son modèle d'affaires, différent de celui de Clever, y fait écho : la firme fait payer les licences aux districts et non aux éditeurs. Ce système inversé permet de compter davantage d'éditeurs partenaires que Clever, *a fortiori* chez des petites structures n'étant pas prêtes à grever leur rentabilité avec des frais d'inscription. En septembre 2024, son site affichait plus de 6 000 produits *EdTech* accessibles. En revanche, le nombre de districts clients est moins élevé que pour Clever (plus de 2 900 districts selon son site en 2024), étant donnés leurs budgets de fonctionnements limités comme on l'a vu précédemment.

Encadré 6. Clever et ClassLink, deux entreprises opposées jusque dans leurs trajectoires socio-économiques

Créée en 2012, Clever a connu une croissance fulgurante emblématique des start-ups états-uniennes à succès. Accompagnée par l'incubateur californien Y Combinator, elle a levé lors de ses premières années des dizaines de millions de dollars auprès de fonds de *venture capital* (Lightspeed Venture Partners, Sequoia Capital, GSV Capital) et de *business angels* comme Peter Thiel (Loeb, 2014). En 2021, Clever a été racheté par l'acteur norvégien des jeux éducatifs Kahoot ! pour un montant compris entre 435 et 500 millions de dollars (Bradley, 2021).

Classlink est en revanche une entreprise assez ancienne du numérique éducatif, créée en 1998. Avant l'avènement du *cloud* et des protocoles API, elle s'est positionnée de façon précoce dans la distribution des logiciels éducatifs et la fourniture de métriques aux administrateurs scolaires pour superviser le retour sur investissement de leurs achats. Elle compte deux financements connus dans son histoire : l'un en 2005 auprès du fonds *Ascend Venture Group*, pour 3 millions de dollars (Globe Newswire, 2005) ; et l'autre en 2007 auprès de l'*Economic Development Authority* (EDA), agence de développement du New Jersey, pour un montant de 750 000 dollars (New Jersey Economic Development Authority, 2007).

Devant épouser les standards d'interopérabilité et de sécurité des données pour figurer sur ces plateformes, les éditeurs n'en restent pas moins souverains en termes commerciaux. Tout au plus Clever prétend-il « travaill[er] […] avec [ses] partenaires éditeurs d'applications pour garantir que leurs prix ne soient pas plus élevés pour les districts qui utilisent Clever »[79], face au risque qu'ils répercutent les frais d'inscription payés à Clever sur le prix de leurs licences.

S'engage ainsi une redynamisation des marchés éducatifs du primaire et du secondaire par l'aval. Ces acteurs de l'intermédiation se positionnent à partir de la synchronisation des données d'apprentissage, cruciales pour les prestations algorithmiques des médias adaptatifs, vers la distribution commerciale des produits *EdTech.* Cela permet aux petites firmes éditrices de franchir les barrières à l'entrée qui se dressent traditionnellement sur les marchés institutionnels. Ils tendent progressivement à jouer le rôle de places de marché, susceptibles de redistribuer le montant des licences entre les éditeurs, conformément au modèle du courtage.

[79] « *We also work with our application partners to ensure their pricing is not higher for districts that use Clever* ».

Cependant le marché est encore peu mature et cherche son modèle d'affaires générique, comme en témoignent les divergences multiples entre Clever et ClassLink. D'autres acteurs exogènes à la force financière plus grande concurrencent ces intermédiaires sur la distribution : les *Big Tech*. Deux d'entre eux, Google et Microsoft, se spécialisent fortement dans la distribution de services numériques éducatifs. Conformément à la stratégie de développement par « cercles concentriques » mise en exergue par Olivier Thuillas et Louis Wiart (2023), ils se sont repositionnés sur l'aval des filières du numérique éducatif en établissant des plateformes éducatives multitâches. Vendues auprès des organisations éducatives sous formes de licences, ce sont à la fois des plateformes d'intermédiation ouvrant l'accès à un panel de services *EdTech* agréés, et des outils de travail collaboratif intégrant des fonctions pédagogiques, concurrençant par là même ces start-ups.

Le cas de Google est éloquent. La firme californienne est présente dans les institutions éducatives américaines depuis les années 2000, que cela soit par son produit initial, son moteur de recherche, ou ses ordinateurs Chromebook. En outre elle a conçu à partir de 2006 une offre plurielle de services web à usages spécialement éducatifs : la suite *Google Workspace for Education*. Celle-ci comprend des applications de travail collaboratif (*Google Docs, Sheets, Slides*), une messagerie (*Gmail*), un outil de visioconférence (*Google Meet*) et un LMS pour l'enseignement primaire et secondaire (*Google Classroom*).

Signe du succès mondial de ces outils, *Google Classroom* comptait 150 millions d'utilisateurs dans le monde en février 2021, au sortir d'une première année de pandémie ayant vu ses inscriptions bondir de 40 millions de nouveaux usagers (Lazare, 2021). Cette suite est aujourd'hui commercialisée suivant plusieurs formules (Google for Education, 2024a) :

- La première, gratuite, nommée « *Education Fundamentals* », offre un accès illimité aux fonctionnalités basiques de ces divers outils, et notamment, pour *Google Classroom*, les usages habituels d'un LMS (des outils d'édition de leçons, de planification de devoirs et des tableaux de bord pour les enseignants, un environnement de travail diversifié pour les élèves, et des messageries et espaces de communication entre les divers usagers).
- Les formules suivantes, distribuées sous licences aux districts comme tout produit *EdTech*, y ajoutent des fonctionnalités diverses. La formule « *Education Standards* » propose des garanties accrues en termes de sécurité et des visualisations des usages. Les deux formules suivantes incluent quant à elles des fonctionnalités pédagogiques étendues, notamment au sein de *Google Classroom.*
- La formule « *Teaching and Learning Upgrade* » permet selon la firme aux enseignants d'« améliorer leur impact pédagogique avec des outils qui enrichissent et personnalisent l'apprentissage, génèrent de l'efficience en classe et permettent d'enseigner et d'apprendre de n'importe où »[80] (Google for Education, 2024b). Le *personalized learning* est ici avancé comme un argument commercial au service d'une technologisation additive, d'assistance de l'enseignant par la technologie, dans une vision similaire à celle du *Department of Education*. Les fonctionnalités sous-entendues sont la possibilité pour le formateur de créer dans *Google Classroom* des exercices interactifs à partir de documents *pdf* et de vidéos *YouTube*, enrichis par une IA générative pouvant automatiser la fourniture d'indices. Il peut enfin accéder à des visualisations des progrès des élèves. *Google Classroom* devient ainsi un média adaptatif concurrent pour les éditeurs spécialisés. Il capitalise sur l'IA générative propriétaire de

[80] « *The Teaching and Learning Upgrade builds on the functionality of Education Fundamentals and helps educators enhance instructional impact with tools that enrich and personalize learning, create classroom efficiency, and enable teaching and learning from anywhere* »

la maison-mère, et l'enrichissement des outils de création par des modules d'éditeurs *EdTech* partenaires, des « *add-ons* », comme la génération de quizz ludiques avec Kahoot!.

- *« Education Plus »* ajoute à ces diverses fonctions une sécurisation accrue des données et un *monitoring* plus étoffé à destination des enseignants et des administrateurs (*Classroom Analytics*). Avec cette formule maximale, Google fait la promesse aux districts de « réduire le besoin d'investissements dans d'autres *EdTech* »[81], de leur faire « économiser en moyenne 410 000 dollars et [d']avoir un retour sur investissement de 191% en éliminant le besoin d'outils de sécurité, de collaboration, de détection du plagiat, d'*adaptive learning* et de stockage »[82] (Google for Education, 2024a). En réunissant tous ces services dans son portfolio, l'entreprise se dresse en concurrent frontal des éditeurs *EdTech* et notamment des médias adaptatifs, avec cette promesse d'efficience faite aux districts selon une rhétorique gestionnaire empreinte des préceptes du *Social Efficiency Movement*. *Education Plus* propose également une plateforme d'intermédiation par authentification unique (*Single Sign-On*). Ainsi, elle promet aux districts une facilité d'accès à leurs diverses licences. L'entreprise concurrence cette fois les acteurs de l'intermédiation que sont Clever et ClassLink. Cependant, pour le partage des données d'élèves entre le SIS et *Classroom*, Google a recours au service de *rostering* de Clever.

En définitive la firme adopte une rhétorique ambivalente, se faisant à la fois concurrente des firmes *EdTech* avec son LMS multitâche, et acteur de leur distribution à travers l'intégration de services tiers dans ses fonctionnalités.

Un autre exemple est Microsoft, qui distribue sa suite *Microsoft 365 Education*. La plateforme donne accès, comme *Google Workspace for Education*, à des applications de travail collaboratif (*Word, Excel, PowerPoint*), une messagerie (*Outlook*), et un LMS (*Teams*). Elle comprend trois formules (Microsoft 365 Education, 2024) :

- la première, « *A1 »*, est une formule gratuite comprenant *Teams* et toutes les autres applications de travail collaboratif sur internet. Elle donne accès à des *Learning Accelerators,* exercices générés par l'IA de Microsoft en lecture et recherche d'informations, offrant *in fine* des analytiques au formateur. Cette formule comprend également une plateforme d'intermédiation, *School Data Sync (SDS),* pour intégrer les données d'identification des systèmes d'information des établissements via les applications Microsoft 365. Distribuant gratuitement cette fonction d'intermédiation, la firme va plus loin que Google en concurrençant directement Clever et ClassLink sur cet aspect central de leurs offres marchandes. En revanche, nulle interface d'authentification unique vers des applications tierces n'est proposée dans la suite, à la différence de Google ou Clever qui offrent cela gratuitement.
- la deuxième, « *A3 »*, comprend les fonctionnalités de *A1*, ainsi que des applications de bureautique, de sécurité et d'analyse des usages. Dans cette formule payante, elle propose une interface d'authentification unique vers les applications partenaires (*Entra ID*). La formule comprend également une IA générative, *Microsoft Copilot*, permettant aux enseignants de créer automatiquement des plans de cours ou des présentations.
- *« A5 »* est la formule la plus complète qui comprend tout ce qu'offre *A3*, ainsi que des outils de *monitoring* et de sécurité enrichis.

[81] « *Education Plus reduces the need to invest in other EdTech* »

[82] « *Experience an average $410k savings and 191% return on investment by eliminating the need for legacy security, collaboration, plagiarism detection, adaptive learning and storage tools* »

La suite de Microsoft constitue donc à la fois un LMS multitâche concurrent de nombreuses applications *EdTech* à la manière de Google, mais se positionne différemment de son homologue californien sur les offres de distribution, en proposant son propre outil de *rostering.* Par ailleurs, les apprenants ont accès à des services *EdTech* via une place de marché intégrée à *Teams*, le *Microsoft Teams Store*. Pour les start-ups qui souhaiteraient en être, il leur faut d'abord obtenir une certification par les équipes de Microsoft, selon des critères d'utilisabilité, mais également de marchandisation. Celles-ci doivent en effet opter pour l'un des deux modèles de tarification, *au forfait* ou *par utilisateur*. Microsoft définit ainsi ces deux options : « Forfaitaire : autorisez l'accès à votre offre pour un prix forfaitaire mensuel ou annuel. […] Par utilisateur : autorisez l'accès à votre offre pour un prix basé sur le nombre d'utilisateurs pouvant accéder à l'offre ou occuper des sièges » (Microsoft Learn, 2024a). Microsoft applique pour sa part un taux de redevance de 3% sur les ventes réalisées, suivant le modèle du courtage. Ce type de dispositif contribue au statut d'intermédiaire incontournable recherché par Microsoft sur le marché *EdTech*, offrant aux start-ups de l'*EdTech* une distribution simple auprès de toute institution éducative ayant souscrit des licences à *Microsoft Teams*.

Dans l'enseignement supérieur, ces acteurs se posent enfin en adjuvants des établissements dans leurs stratégies d'équipement en technologies éducatives et de mise en place du *personalized learning*. Tissant de longue date des rapports étroits avec des institutions éducatives, ils leur prodiguent des offres spécifiques de formations et d'outils pédagogiques. Présentant leurs offres en recours à des problèmes structurels (échec scolaire, coûts des frais d'inscription, disparition d'emplois), ils formulent des « promesses socio-techniques » (Joly, 2010) devant leur conférer, outre un pouvoir économique, une aura sociétale. Un exemple est *Student Success Services* de Google, un service d'accompagnement à destination des établissements d'enseignement supérieur. La firme californienne présente ainsi son offre :

> « Les établissements d'enseignement travaillent plus que jamais pour offrir des opportunités d'apprentissage au nombre croissant d'étudiants qu'ils accueillent. Mais il existe encore des lacunes dans les compétences dont les employeurs ont besoin une fois les étudiants diplômés. Une récente enquête de McKinsey & Company a révélé que 87 % des entreprises ont déclaré qu'elles connaissaient des déficits de compétences maintenant ou dans l'année à venir. Et l'impact de la pandémie n'a fait qu'ajouter à l'urgence pour les écoles et les entreprises de technologies éducatives de trouver de nouvelles façons d'aider encore plus d'apprenants »[83].

Face à ce problème structurel, arrive alors la nouvelle offre de Google, permettant à des institutions d'implémenter un paradigme éducatif « à la demande » via une « nouvelle plate-forme d'apprentissage basée sur l'intelligence artificielle *Google Cloud*, qui comprend [un] tuteur interactif »[84]. Ces établissements pratiquent principalement des formations à distance, recherchant des économies d'échelle à travers leurs technologies. Google mentionne ainsi sur son blog sa collaboration avec la Walden University, établissement de formation à distance dans les métiers de la santé, ayant développé avec Google son « tuteur intelligent » expérimental nommé *Julian* (Google Cloud, 2024). Entraîné à partir des modèles de langage de la firme*,* le système analyse le contenu saisi par les instructeurs et génère des questions à choix multiples, répond aux questions des étudiants sur un contenu complexe et pose des questions pour tester leurs connaissances. Google

[83] *« Education institutions are working harder than ever to provide learning opportunities to the increasing number of students they serve. But gaps still exist in the skills needed by employers once students graduate. A recent McKinsey & Company survey found that 87% of companies reported that they are experiencing skills gaps now or in the coming year. And the impact of the pandemic has only added to the urgency for schools and education technology companies to find new ways to support even more learners ».*

[84] *« …a new Google Cloud artificial intelligence powered learning platform, which features the interactive tutor ».*

justifie cette collaboration par « l'objectif de développer les talents en élargissant l'accès à la technologie […] à une époque où de nombreuses régions du pays sont aux prises avec une pénurie de personnel de santé »[85] (Google Cloud, 2024). La firme californienne se pose ainsi en alliée des universités face aux difficultés budgétaires qu'elles rencontrent, et en acteur politique répondant à des problèmes socio-économiques.

Amazon n'est pas en reste sur le marché de la distribution. Outre la possibilité de profiter de ses algorithmes, l'adhésion des éditeurs *EdTech* au *cloud* AWS leur ouvre l'accès à une place de marché, *AWS Marketplace*. Celle-ci réunit face à eux des organisations éducatives également adeptes d'AWS pour l'administration de leurs systèmes informatiques. Dans un article publié sur le blog de la firme, deux cadres du groupe se félicitent de compter en 2023 plus de 2 000 institutions éducatives de tous niveaux effectuant des achats sur cette plateforme (Raj Purohit et Abowd, 2024). En octobre 2024, la plateforme réunissait 124 produits d'éditeurs allant du niveau *K-12* à la formation professionnelle. Comme dans *Microsoft Teams*, le taux de redevance sur les ventes de licences est de 3% (Amazon Web Service, 2024e).

Ainsi les acteurs issus de la filière des systèmes de gestion de l'information, ou exogènes comme les *Big Tech*, font-ils advenir une socio-économie nouvelle dans le circuit institutionnel, en lançant leurs plateformes d'intermédiation. Le modèle du courtage tend à représenter une alternative au modèle du club, qui permet traditionnellement aux éditeurs d'amasser directement des revenus mais favorise les acteurs les plus concentrés. Cette alternative est en revanche le modèle dominant dans la distribution d'applications directement au consommateur, où les *Big Tech* ont des positions hégémoniques.

L'emprise des Big Tech *dans la distribution des produits éducatifs non-formels*

Apple et Google dominent depuis la fin des années 2000 le marché de la distribution d'applications sur les terminaux mobiles. En amont, les deux *Big Tech* se partagent le marché américain des systèmes d'exploitation, avec respectivement 51,6% de part de marché pour *iOS* et 48% pour *Android* en juin 2021 selon Kantar (Kantar Worldpanel, 2024). Ils installent sur leurs systèmes des magasins d'applications, respectivement l'*App Store* et *Google Play Store*, assortis de systèmes de paiement propriétaires. Suivant le modèle du courtage, les deux *Big Tech* perçoivent ainsi les abonnements et les « achats intégrés » réalisés par les internautes dans les applications, et y prélèvent des commissions substantielles avant de les redistribuer aux éditeurs.

Dans les années 2010, les taux de redevance avoisinaient les 30%. En 2020, un événement perturbateur a été la décision des deux groupes de retirer le jeu vidéo *Fortnite* de leurs magasins d'applications, après que le studio qui le développait, Epic Games, a renvoyé ses utilisateurs vers son propre système de paiement. Ce dernier a alors intenté des procès contre Apple et Google pour dénoncer leurs abus de position dominante à l'encontre des éditeurs d'applications. Justifiant leurs systèmes de paiement par des enjeux de sécurité, les deux *Big Tech* ont néanmoins revu leurs politiques suite aux décisions de la Cour Suprême.

Apple a d'abord baissé en 2020 son taux de redevance à 15% pour les entreprises générant moins d'un million de dollars de chiffre d'affaires. Si, deux ans plus tard, la Cour Suprême a blanchi la firme

[85] *« Both Google and Walden share the goal of developing talent by expanding access to technology. At a time when many parts of the country are struggling with healthcare worker shortages, Julian can help Walden provide upskilling and reskilling training to all types of learners ».*

de Cupertino de tout abus de position dominante, elle l'a contrainte à laisser les éditeurs d'applications contourner son système de paiement. Apple s'est exécuté, tout en instituant pour ces éditeurs une taxe de 27% de leurs recettes. La Cour Suprême a jugé en décembre 2023 que Google avait enfreint le droit de la concurrence, sans préciser encore les conséquences pour la firme de Mountain View, qui a d'ores et déjà fait appel. Ainsi les deux *Big Tech* se partagent à deux le marché de la distribution d'applications, en exploitant leur ascendant sur les éditeurs d'applications malgré des critiques croissantes.

En dépit des ponctions importantes sur leurs recettes, les éditeurs tirent des opportunités commerciales de ce circuit de distribution où l'acheteur et l'usager ne font qu'un. Il leur est ainsi plus facile de relever le double défi de recruter de nouveaux abonnés et de les fidéliser. Usuel pour constituer et monétiser des réseaux d'utilisateurs, le *freemium* est un modèle d'affaires plus rentable dans la vente directe au consommateur que dans les circuits institutionnels. Il ne s'agit plus en effet de laisser l'usager naviguer dans l'application dans l'espoir qu'il convainque l'acheteur institutionnel de payer une licence, à l'image de la stratégie première de Lalilo. Dans les marchés de l'éducation non-formelle comme l'autoformation, l'usager peut lui-même s'abonner.

Duolingo, leader du marché de l'apprentissage des langues en ligne, offre un exemple de ces perspectives commerciales et des contraintes que posent les redevances aux magasins d'applications. Après s'être rémunéré, comme on l'a vu dans l'encadré 2 (p.87), par la vente de traductions effectuées en *crowdsourcing* par les apprenants, la firme de Pittsburgh a adopté en 2017 un modèle *freemium* complété par une gamification mercantile du dispositif.

Elle justifie cette décision auprès du grand public par une volonté de démocratiser l'éducation. Dans l'onglet « Mission » de son site, la firme explique ainsi : « Duolingo est utilisé par l'homme le plus riche du monde et de nombreuses stars d'Hollywood, mais aussi par les élèves des écoles publiques des pays en développement. Pour nous, l'égalité, c'est de ne pas devoir dépenser plus pour s'offrir une meilleure éducation » (Duolingo, 2024b). Une assertion mise en doute par les options pédagogiques proposées dans ses offres *premium.*

Dans son rapport délivré en 2021 aux investisseurs avant son introduction en bourse, la firme offre une explication plus prosaïque : « Notre modèle *freemium* est le cœur de notre succès, car il permet un recrutement significatif d'utilisateurs »[86] (Duolingo, 2021, p.112). Il replace ainsi son produit dans un schéma de développement vertueux fait de deux « *flywheels »* décrivant le développement de l'entreprise en jouant sur les deux terrains économique et éducatif (Figure 11).

Ainsi ces roues, l'une pédagogique, l'autre financière, s'entraînent mutuellement vers une augmentation toujours plus forte de la base d'utilisateurs. Pour y parvenir, un moyen est mentionné de manière centrale : l'accroissement du duo « efficacité-engagement ». Y réside le projet stratégique de développement de Duolingo tel que communiqué aux investisseurs : la fourniture d'une « expérience » toujours plus engageante et efficace, et attirant toujours plus d'apprenants – ce dont témoigne la flèche verte au centre du schéma. La promesse de l'effet de réseau découlant d'un duo « efficacité-engagement » est explicitement postulée : « Plus nos produits sont engageants et plus nous enseignons efficacement, plus nos apprenants parlent de Duolingo à leurs amis et plus nous

[86] « *Our freemium business model is core to our success because it enables significant user scale ».*

continuons à développer notre base d'apprenants »[87] (Duolingo, 2021, p.3). Le modèle *freemium* est dès lors le moyen, du côté de l'*investment flywheel*, de monétiser la qualité de l'« expérience ».

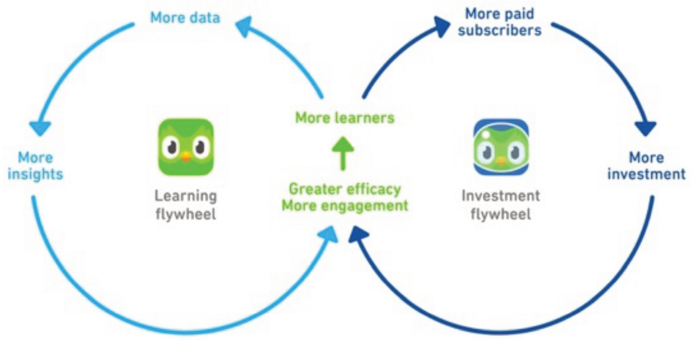


Figure 11. La stratégie de développement organique de Duolingo. Source : Duolingo, 2021, p.113.

L'application gratuite consiste ainsi en une version comprenant des publicités personnalisées grâce à la collecte de *cookies*. Cette source de financement est toutefois secondaire dans le cas de Duolingo, les 49,9 millions de dollars amassés par ce biais en 2023 représentant 9,4% du chiffre d'affaires total (Duolingo, 2024a, p.74). La mise en compétition ludique des apprenants et la marchandisation de bonus divers, « à l'acte », apportent une source de revenus supplémentaire, pour 34,7 millions de dollars, soit 6,5% des recettes annuelles (ibid.). En cumulé, la version gratuite représente donc environ 16% du chiffre d'affaires. Ce sont bien les formules par abonnement qui constituent la manne principale de l'entreprise, avec 404,7 millions de dollars, soit 76,2% des revenus totaux (ibid.).

C'est pourquoi, dans ses *flywheels,* la firme vise surtout la conversion de l'effectif croissant d'utilisateurs en abonnés. Aux côtés de la formule gratuite sont proposées à ce jour deux offres d'abonnement aux États-Unis : *Super Duolingo* et *Duolingo Max*. Le défi de la conversion est cependant difficile à relever, sachant que sur les 88,4 millions d'utilisateurs mensuellement actifs au 31 décembre 2023, seuls 6,6 millions étaient abonnés – soit un ratio d'environ 7,5% (Duolingo, 2024a, p.65). En cause, la forte volatilité des utilisateurs sur le marché des applications d'apprentissage des langues, étant donnés des coûts de sortie quasi-nuls.

L'engagement de l'utilisateur est donc surveillé par l'entreprise pour ses retombées commerciales. Ainsi, sur la formule gratuite, il indiquerait la satisfaction des usagers quant à l'application, dont découle leur éventuelle propension à s'abonner, sinon leur rétention en tant que prospects pour des annonceurs. Pour maximiser la conversion, un second chantier consiste à rendre l'abonnement plus désirable que la formule gratuite aux yeux du grand public. On peut dénombrer trois principaux arguments d'appel : l'absence de publicités, la permanence d'avantages initialement vendus à l'acte, tels que les gemmes – la monnaie virtuelle de Duolingo, permettant de prolonger des *streaks* ou d'obtenir des essais temporairement illimités –, et enfin des technologies de personnalisation pédagogique supplémentaires.

L'abonnement *Super Duolingo* propose ainsi, dans l'onglet *Practice Hub*, un algorithme reposant à l'apprenant les questions sur lesquelles il a commis des erreurs par le passé. Cette capitalisation sur l'IA a été poussée plus loin dans *Duolingo Max*, offre d'abonnement commercialisée aux États-Unis

[87] *« The more engaging our products are, and the more effectively we teach, the more our learners tell their friends about Duolingo and the more we continue to grow our learner base ».*

et au Royaume-Uni depuis mars 2023. Deux agents conversationnels, *Explain My Answer* et *RolePlay*, fournissent pour le premier des explications à l'utilisateur sur ses erreurs, et pour le second l'occasion de converser dans des situations variées. Ils ont été développés en partenariat avec l'éditeur OpenAI, à partir de son grand modèle de langage *GPT-4*.

Finalement, les revenus croissants amassés par Duolingo lui permettent de se démarquer de ses concurrents, via des opérations marketing multicanales, ayant coûté 75,8 millions de dollars en 2023, soit 14,8% de ses revenus (Duolingo, 2024a, p.72). Surtout la firme se distingue en investissant des sommes très importantes dans sa recherche et développement comme on l'a vu plus haut, à hauteur de 36,6% de son chiffre d'affaires en 2023. Celle-ci vise à développer ses algorithmes mais également à tester de nouvelles fonctionnalités sur des millions d'apprenants en *A/B testing*.

Cependant les contraintes économiques à court ou moyen terme sont nombreuses pour des entreprises de l'éducation non-formelle comme Duolingo, principalement à cause du poids des redevances auprès des magasins d'applications. Pour la firme de Pittsburgh, passant jusqu'ici par les systèmes de paiement d'Apple et Google et excédant de loin le statut de petit éditeur concédé par le premier, le prix à payer est important. En 2023 il s'est élevé à 142,1 millions de dollars, soit 26,7% de ses recettes totales (Duolingo, 2024a, p.72).

La firme se doit donc de parvenir à un équilibre financier pour être rentable en fin d'exercice comptable. Cependant le coût des investissements dans la recherche et développement, et en particulier dans l'intelligence artificielle, pèse lourd dans les comptes de l'entreprise. L'EBITDA ajusté de Duolingo – « *Earnings Before Interest, Taxes, Depreciation and Amortization* », équivalent de l'Excédent Brut d'Exploitation (EBE) français, indiquant la performance financière de son activité[88] – était déficitaire de près d'1 millions de dollars en 2021 (Duolingo, 2022, p.62), avant de redevenir positif l'année suivante en s'élevant à 93,7 millions en 2023. « Nos coûts continuent d'augmenter, et plusieurs de nos investissements ont pour effet de réduire notre marge opérationnelle et notre rentabilité. Si nos investissements n'étaient guère couronnés de succès, nos performances commerciales et financières pourraient en être affectées »[89] (Duolingo, 2024a, p. 4). Étant donnés les risques d'« accident industriel », le danger se fait pressant pour l'entreprise, poussée à diversifier ses abonnements (avec *Duolingo Max*), ou à investir dans d'autres marchés (comme celui de la certification en anglais). Malgré le récit d'une croissance organique et autonome de Duolingo, la firme est pour l'instant dépendante des apports de fonds extérieurs pour continuer à financer son modèle d'investissement.

L'exemple archétypal de Duolingo nous montre les contraintes diverses qui, dans la distribution des applications adaptatives directement au consommateur, s'imposent aux éditeurs. Une constante demeure le modèle par abonnement lié à la distribution en SaaS – norme permettant aux entreprises de s'assurer un revenu récurrent, mais à partir de laquelle elles doivent innover pour attirer de nouveaux utilisateurs et espérer devenir rentables. Si Apple et Google y impriment fortement leur mainmise, la tendance duopolistique du marché de la distribution attire l'attention croissante des pouvoirs fédéraux.

[88] Elle équivaut à la marge nette sans les intérêts de la dette, les impôts, les provisions, les amortissements ni les dépréciations, entre autres dépenses.
[89] « *Our costs are continuing to grow, and some of our investments have the effect of reducing our operating margin and profitability. If our investments are not successful, our business and financial performance could be harmed* »

Les trois domaines générateurs de valeur dans la filière états-unienne des médias éducatifs, à savoir le développement d'algorithmes, la production de contenus et la distribution des médias éducatifs, ont fait apparaître de nouveaux rapports de force structurants. Les structures de production de type « start-ups » sont fragiles, mais jouissent pour les mieux dotées d'une flexibilité et de guichets de financements variés pour investir la plupart de leurs ressources dans leurs algorithmes, attisant les convoitises de groupes concentrés du numérique éducatif et de l'édition scolaire.

En matière de production de contenus, les contrats de sous-traitance entre les structures de production et les éditeurs de contenus certifiés tendent à laisser la place à des rachats et des voies d'autonomisation pour les start-ups comme les LLM. Malgré leurs défauts et les doutes qui les entourent, ceux-ci technologisent un peu plus la filière au profit des *Big Tech* qui possèdent ces infrastructures.

La distribution des médias adaptatifs, cruciale pour la rentabilité des firmes et l'entraînement de leurs technologies, est dictée dans l'autoformation par le modèle du courtage, consacrant les magasins d'applications d'Apple et Google. Ce modèle tend à gagner les circuits institutionnels traditionnellement régis par le modèle du club, à la faveur des intermédiations proposées par des gestionnaires de plateformes privés comme Clever et ClassLink, et des *Big Tech* comme Microsoft, Google et Amazon. Ces divers rapports de force incitent les firmes éditrices à réinventer leurs formules par abonnement ou leurs licences, en instillant des formes de gratuité ou d'achats de fonctionnalités à l'acte.

S'il est prématuré de conclure qu'une socio-économie nouvelle et stabilisée entoure la valorisation des médias adaptatifs, on peut en revanche constater que trois tendances plus ou moins anciennes sont puissamment réimpulsées : l'importance des start-ups en matière d'innovations, la dynamisation de la filière par des acteurs de l'intermédiation, et les tentatives d'imposition des *Big Tech* sur l'ensemble de la filière. Ces tendances homogènes entre circuits institutionnels et informels sont cependant conditionnées par la constitution de compétences et de solidarités professionnelles parmi les praticiens des diverses filières.

3. Derrière l'*EdTech* états-unienne, des communautés renforçant l'influence des acteurs des technologies

La décision d'IMS Global Learning Consortium, l'alliance d'acteurs du numérique éducatif à l'origine du protocole *One Roster*, de se renommer « 1EdTech » en 2022 révèle deux tendances importantes : d'une part, le nouvel élan socio-économique entraîné par l'avènement des start-ups et de l'intelligence artificielle aux États-Unis dans les années 2010 ; d'autre part la volonté de « faire réseau » parmi des acteurs très variés : créateurs de start-ups, cadres de multinationales des technologies, associés de fonds de capital-risque, éditeurs scolaires, administrateurs d'institutions scolaires et cadres de l'administration fédérale. À travers des médias d'information communs ou des espaces de rencontre comme les salons annuels, ces communautés s'avèrent être des interfaces pour le dialogue entre les décideurs politiques et les acteurs des technologies, que ceux-ci soient des start-ups ou des *Big Tech*. Ces derniers sont particulièrement proactifs dans l'animation des réseaux, leur permettant d'accroître leur influence dans la définition des professionnalités liées à l'*EdTech.*

Des réseaux diversifiés et dynamiques

Ces communautés émanent de deux types d'acteurs. Les premiers sont des organisations à but non-lucratif s'étant imposées dans le paysage éducatif par la délivrance de normes et de protocoles pour implémenter les technologies dans l'éducation.

Le pouvoir fédérateur des organisations à but non-lucratif

C'est le cas de 1EdTech. La création d'IMS Global Learning Consortium en 1995 a été impulsée par le média Educause, créé en 1992, et devenu une source d'information privilégiée pour les professionnels de l'enseignement supérieur (Pearson, 2024).

Une autre organisation de ce type est l'*International Society for Technology in Education* (ISTE). Elle se concentre depuis 1979 sur l'intégration des technologies informatiques dans l'éducation. Elle fournit différents services aux professionnels de l'éducation (éducateurs, administrateurs d'établissements), pour les former et faire évoluer leurs pratiques. Surtout, elle édite les « *ISTE standards* », ayant pour but de formaliser auprès des professionnels de l'éducation des compétences nécessaires pour apprendre, enseigner et administrer une institution éducative avec des outils numériques. Certaines *State Education Agencies* recommandent des outils conformes à ces normes, à l'image de celle de l'État de New York (New York State Education Department, 2020).

ISTE propose des ressources d'apprentissage professionnel, notamment des webinaires, des cours en ligne, des services de conseil, et un média, Edsurge. Acquis en 2019, ce dernier est une source d'information technologique de premier plan pour les professionnels de l'enseignement primaire, secondaire et supérieur. ISTE organise également chaque année une convention annuelle, l'*ISTE Live*. Dans son appel aux exposants en 2024, l'organisation vante cet événement comme « une place de marché internationale de l'*EdTech* inégalée »[90], réunissant plus de 16 000 personnes chaque année (International Society for Technology in Education, 2024a). Pour l'année 2023, elle rapporte des visiteurs issus d'organisations très diverses, avec une surreprésentation des institutions d'enseignement primaire et secondaire : des écoles, collèges et lycées (41%), des districts (28%), des *colleges* et universités (5%), des employeurs (5%) des organisations à but non lucratif (5%) et des vendeurs de technologies (7 %, ibid.).

Quant aux professions des visiteurs, la diversité était également de mise : des enseignants des niveaux primaire et secondaire (29 %), des dirigeants et professionnels d'entreprises technologiques (15%), des responsables technologiques d'établissements (12%), des directeurs ou directeurs-adjoints d'établissements (7%), des cadres-dirigeants de districts (5%), des ingénieurs pédagogiques (4%), des consultants et formateurs (3%) ou des bibliothécaires et spécialistes des médias (3% ; ibid.). Cette convention hétéroclite représente une opportunité commerciale pour les exposants, rencontrant des usagers et acheteurs nombreux.

Les réseaux de fonds de venture capital *et d'incubateurs privés*

Le second type d'acteurs à l'initiative des réseaux *EdTech* est constitué par des fonds de capital-risque et des incubateurs privés. Ceux-ci réunissent dans leurs programmes des start-ups jouissant des conseils de professionnels variés. C'est le cas de StartEd, réseau fondé par Ash Kaluarachchi.

[90] « *Attendees spend an average of six hours exploring the expo hall, making it an unequaled global edtech marketplace* »

Ancien consultant dans le domaine financier et technologique puis entrepreneur, il s'est lancé dans l'incubation de start-ups en dirigeant TechStars, l'un des incubateurs les plus importants des États-Unis, avant de créer StartEd en 2016. Basé à New York, ce réseau soutient des start-ups dans l'ensemble des secteurs éducatifs, avec ses programmes *Hyper Accelerator* et *USA EdTech Accelerator*. Il permet ainsi l'adhésion à une communauté internationale d'entrepreneurs en technologies éducatives. Il organise chaque année le salon *EdTech Week*, réunissant des responsables de l'éducation, des investisseurs axés sur l'*EdTech*, des fondateurs de startups, des fondations et des dirigeants d'entreprises.

Notre voyage aux États-Unis nous a permis d'assister à l'édition *EdTech Week 2023*. Celle-ci consistait en quatre journées de conférences et de *masterclasses* ayant lieu dans différents lieux, et notamment dans les sièges d'organisations partenaires aux profils variés : le siège new-yorkais de Google, l'école de formation pour enseignants Relay Graduate School of Education, ou encore dans le laboratoire *Learning Analytics Research Network* (LEARN) de l'université de New York. Ainsi ce réseau parvient-il à agréger des publics divers, allant jusqu'à des chercheurs en *Learning Analytics*.

Parmi les firmes observées précédemment, les start-ups apparaissent au cœur de ces communautés, que celles-ci soient impulsées par des organisations à but non-lucratif ou par des incubateurs et des fonds d'investissements. Elles en tirent des conseils pour leurs développements ainsi que des opportunités de *networking* et de financements.

Les grands groupes de l'édition scolaire sont quant à eux particulièrement actifs dans les réseaux anciens et à but non-lucratif tels qu'ISTE ou 1EdTech. Savvas Learning Company et McGraw-Hill figurent ainsi parmi les principaux sponsors de la *Learning Impact Conference* organisée par 1EdTech en 2021 (1EdTech, 2021). Des cadres de Pearson sont engagés quant à eux de longue date dans les ateliers et initiatives de l'IMS puis de 1EdTech (Pearson, 2024).

Les *Big Tech,* en particulier Microsoft, Google et Amazon, exercent une influence forte dans chacun de ces réseaux, par un sponsorat très actif. Les deux premiers sont les principaux sponsors de l'édition 2024 d'*ISTE Live.* Déjà investi aux côtés de StartED dans le financement de programmes d'accélération comme « *StartEd sponsored by Google Cloud* », la firme de Mountain View fait partie des co-organisateurs d'*EdTech Week* en accueillant dans ses bureaux new-yorkais des conférences de l'événement. Microsoft fait quant à lui partie des entreprises dont les cadres s'expriment régulièrement lors des *masterclasses* et débats. Ces sponsorats et animations d'événements s'inscrivent dans une stratégie de fédération des acteurs du numérique éducatif – allant des professionnels de l'éducation aux entrepreneurs et investisseurs *EdTech* – sous leur étendard.

Ils en viennent ainsi à constituer des réseaux d'acteurs autour d'eux, liés à leurs programmes d'incubation observés plus haut. Dans le *Cloud Google for Startups* est mise en avant une « communauté mondiale composée de start-up, d'experts, d'investisseurs et de partenaires pour découvrir des opportunités de croissance et développer [son] activité » (Google Cloud, 2024c). *AWS Education Accelerator* est pour sa part « un réseau d'anciens élèves et d'investisseurs pour le mentorat »[91], proposant des « ateliers pratiques, tables rondes et discussions de leadership inspirantes » ainsi qu'un « programme adapté aux startups *EdTech* »[92] (Amazon Web Service, 2024d). Microsoft a en outre créé sa « *Microsoft Education Partner community »* faite d'entreprises

[91] *« Tap into our alumni and investor network for mentorship [...]».*
[92] *« Hands-on workshops, roundtable discussions, and inspirational leadership chats. Benefit from curriculum tailored to EdTech startups [...]».*

et d'institutions éducatives utilisatrices de la suite *Microsoft Education*. La communauté s'est ainsi réunie lors d'un « *Global EdTech Summit* » pour « discuter collectivement de l'orientation du secteur de l'éducation, examiner les opportunités pour la *Microsoft Education Partner community* et partager des informations sur la stratégie commerciale de Microsoft en matière d'éducation »[93] (Microsoft Education Team, 2022).

La prescription de compétences professionnelles

Quel que soit le type d'institutions à leurs origines, les communautés *EdTech* prescrivent des compétences liées aux technologies dans un cadre compatible avec le *personalized learning.* L'*ISTE* et ses standards à destination des publics éducatifs en offrent un premier exemple.

Les standards pédagogiques de l'ISTE

L'organisation non-lucrative incite les élèves, enseignants et personnels administratifs à adapter leurs pratiques à un projet techno-éducatif similaire à celui esquissé par Washington (International Society for Technology in Education, 2024b). Ces normes encouragent notamment une posture très active et autonome des élèves lors de l'apprentissage : ceux-ci doivent fixer leurs propres objectifs d'apprentissage, « construire des réseaux et personnaliser leurs environnements d'apprentissage de manière à soutenir le processus d'apprentissage »[94], solliciter du *feedback* pour améliorer leurs pratiques, s'approprier et comprendre le fonctionnement des outils numériques et des algorithmes, « évaluer l'exactitude, la validité, les biais, l'origine et la pertinence du contenu numérique »[95] ou encore s'assurer du respect de leurs données personnelles (International Society for Technology in Education, 2024c).

Du côté des enseignants, il s'agit de se former en permanence aux nouveaux usages en matière de technologies et de mettre en place « des expériences d'apprentissage personnalisées qui favorisent l'apprentissage indépendant et tiennent compte des différences et des besoins des apprenants »[96]. Outre cet habitus de facilitateurs du *personalized learning*, l'ISTE les investit du devoir de veiller à la protection des données de leurs élèves, et, suivant une gouvernance horizontale, de collaborer avec leurs collègues et leurs élèves pour améliorer les usages pédagogiques de la technologie. Ils doivent enfin employer les technologies pour évaluer, et les analytiques pour « guider les progrès et communiquer avec les élèves, les parents et les acteurs de l'éducation afin de renforcer l'autonomie des élèves »[97] (International Society for Technology in Education, 2024d).

---

[93] *« Microsoft's EdTech partner ecosystem plays a key role in how Microsoft can successfully serve our education industry customers around the world, so each year we meet to collectively discuss where the education industry is going, consider opportunities for the Microsoft Education Partner community, and share insights into Microsoft's Education business strategy ».*

[94] « 1.1b. Environnements d'apprentissage personnalisés. Les étudiants construisent des réseaux et personnalisent leurs environnements d'apprentissage de manière à soutenir le processus d'apprentissage » (« *1.1b Customized Learning Environments. Students build networks and customize their learning environments in ways that support the learning process* »).

[95] « 1.3.b. Évaluer les informations. Les étudiants évaluent l'exactitude, la validité, les biais, l'origine et la pertinence du contenu numérique » (« 1.3.b. Evaluate Information. Students evaluate the accuracy, validity, bias, origin, and relevance of digital content »)

[96] « 2.5.a. S'adapter aux différences des apprenants : utiliser la technologie pour créer, adapter et personnaliser des expériences d'apprentissage qui favorisent l'apprentissage indépendant et tiennent compte des différences et des besoins des apprenants » (« *2.5.a. Accomodate Learner Differences : Use technology to create, adapt and personalize learning experiences that foster independent learning and accommodate learner differences and needs* »).

[97] « 2.7.c. Utiliser les données pour guider les progrès. Utiliser les données d'évaluation pour guider les progrès et communiquer avec les élèves, les parents et les acteurs de l'éducation afin de développer l'autonomie des élèves ». (« *2.7.c Use Data to Guide Progress. Use assessment data to guide progress and communicate with students, parents and education stakeholders to build student self-direction* »).

Les directeurs d'établissements doivent quant à eux assurer aux différents publics l'accès à des équipements et des formations suffisantes pour améliorer leurs pratiques. Sur le plan de la gouvernance, ils sont censés développer des visions stratégiques concernant l'équipement, mettre en place des évaluations cycliques des outils et médias éducatifs, « développer des évaluations qui fournissent une vue personnalisée et exploitable des progrès des élèves en temps réel »[98] ou encore assurer la protection des données des usagers sous leur autorité (International Society for Technology in Education, 2024e).

Ces standards comprennent également un métier nouveau, « *coach* », qu'on peut traduire par « expert pédagogique ». Celui-ci doit contribuer à la formation des autres publics et conseiller les directeurs lors de la mise en place des technologies numériques à l'échelle des établissements. Le *coach* devrait ainsi assister le directeur et les enseignants dans leurs pratiques en établissant avec eux des plans personnalisés d'action et de formations, « collaborer avec les éducateurs pour développer des expériences d'apprentissage authentiques et actives qui favorisent l'action des étudiants, approfondissent la maîtrise du contenu et permettent aux étudiants de démontrer leurs compétences »[99]. Il doit en outre « connecter les dirigeants, les éducateurs, les services pédagogiques et techniques, les experts du domaine et les fournisseurs de solutions pour maximiser le potentiel de la technologie pour l'apprentissage »[100] (International Society for Technology in Education, 2024f). Le *coach* formerait ainsi un maillon central dans le changement des pratiques éducatives au niveau de l'établissement.

On retrouve ainsi une projection socio-constructiviste voire connectiviste de l'apprentissage, similaire au *personalized learning.* En outre des compétences transversales liées à la *digital literacy* sont appelées à se diffuser parmi les professionnels de l'éducation et les élèves – comme la compréhension du fonctionnement de dispositifs algorithmiques, la capacité à interpréter des analytiques d'apprentissage, le contrôle de la qualité de contenus numériques, ou la protection des données personnelles. Un ethos d'individu entrepreneur de ses pratiques éducatives transparaît de façon transversale dans ces recommandations.

Les normes professionnelles dictées par les Big Tech

Dans leurs offres à destination des établissements éducatifs, les *Big Tech* encouragent également des pratiques éducatives empreintes de flexibilité, faisant écho à la conception additive du *personalized learning*. Témoin Google, qui projetait dans sa suite *Google Workspace for Education* des usages de l'analytique en renfort des pratiques des enseignants.

Les *Big Tech* s'efforcent par ailleurs d'orienter la définition de professionnalités dans l'espace public. Microsoft met pour sa part en avant son offre de formations auprès des particuliers pour les former au numérique, et lutter contre ce qui est communément nommé « l'illectronisme »

[98] « 3.3.e. Utiliser des analytiques pour visualiser les progrès des étudiants. Développer des évaluations d'apprentissage qui fournissent une vue personnalisée et exploitable des progrès des élèves en temps réel » (« *3.3.e Use Analytics to View Student Progress. Develop learning assessments that provide a personalized, actionable view of student progress in real time* »).

[99] « 4.4.a Collaborer sur la conception de l'apprentissage. Collaborer avec les éducateurs pour développer des expériences d'apprentissage authentiques et actives qui favorisent l'action des étudiants, approfondissent la maîtrise du contenu et permettent aux étudiants de démontrer leurs compétences » (« *4.4.a Collaborate on Learning Design. Collaborate with educators to develop authentic, active learning experiences that foster student agency, deepen content mastery and allow students to demonstrate their competency* »).

[100] « 4.1.e Maximiser le potentiel de la technologie. Connecter les dirigeants, les éducateurs, les services pédagogiques et techniques, les experts du domaine et les fournisseurs de solutions pour maximiser le potentiel de la technologie pour l'apprentissage ». *(« 4.1.e. Maximize Potential of Technology. Connect leaders, educators, instructional support, technical support, domain experts and solution providers to maximize the potential of technology for learning »).*

– une problématique sociétale dont les pouvoirs publics ont pris la mesure tardivement. L'entreprise propose sur son « centre de formations » aux outils *Microsoft 365* des ressources pédagogiques liées à l'IA avec cette aspiration : « Microsoft collabore avec des communautés, des associations, des institutions d'enseignement, des entreprises et des gouvernements pour permettre aux personnes exclues de l'économie numérique d'acquérir les compétences, les connaissances et l'opportunité d'obtenir un emploi et une source de revenu » (Microsoft Learn, 2024b).

Idem pour Amazon et son *AWS Academy*, qui vise à donner aux établissements d'enseignement supérieur les ressources pour préparer les étudiants à des certifications et des carrières dans le *cloud*. À travers leurs centres de formation « maison », ces multinationales se présentent avec les atours de l'université, faisant d'eux de potentiels concurrents des établissements qu'ils prétendent servir. Ainsi que le montrait Thomas Grignon dans son étude de la « *Google Analytics Academy* » (2016) :

> « Les multiples emprunts au lexique universitaire observables sur l'ensemble de la plateforme (« formation », « éducation », « qualification », « cours », « séminaires », « ressources pédagogiques »…) et la référence à un lieu de savoir archétypal n'ont bien sûr ici rien d'anodin. La métaphore vient souligner la qualité, l'utilité potentielle et le sérieux des contenus proposés tout en permettant à Google de s'afficher, non comme une entreprise servant des intérêts privés, mais comme un acteur impliqué dans la médiation des savoirs et dans l'insertion professionnelle de ses « étudiants » (Grignon, 2016).

On touche là à l'enjeu politique de leurs offres auprès des organisations éducatives : face aux injonctions des pouvoirs publics à implémenter des méthodes éducatives ambitieuses sans que des financements suffisants leur soient accordés, les *Big Tech* leur offriraient des outils concrets, tout en renvoyant les organisations éducatives vers leurs portfolios de services marchands.

Ces services permettent en outre à ces groupes d'instaurer des normes d'usages. Thomas Grignon observe ainsi une standardisation des compétences liées aux métiers de la communication autour des codes dictés par Google :

> « En diffusant, dans une pluralité de contextes socio-organisationnels, des programmes de certification toujours plus nombreux, inspirés par les mêmes valeurs et les mêmes prétentions, Google conduit finalement des usagers très divers à adopter des pratiques similaires. À travers eux, il opère la mise en circulation de logiques, de savoirs et de manières de faire issues du secteur communicationnel et aboutit, sinon à la standardisation progressive des pratiques d'écriture médiatique, du moins au brouillage graduel des frontières entre les espaces industriels, experts et ordinaires de la communication » (ibid.).

L'ambiguïté est de mise dans la posture des *Big Tech*, tantôt au service de la « transformation digitale » des établissements, et tantôt établissements éducatifs eux-mêmes. Cet opportunisme est renforcé par leur intérêt à dorer leur « marque employeur » auprès des étudiants et futurs travailleurs hautement qualifiés (dont des ingénieurs et des *data scientists)*.

Les tendances observées précédemment dans les stratégies des firmes concordent avec les professionnalités liées à l'*EdTech*. La montée en puissance des start-ups en matière d'innovations est validée par la constitution de réseaux d'incubation autour d'elles. Les tentatives d'imposition des *Big Tech* dans l'ensemble de la filière sont également visibles dans leurs stratégies de sponsorat et de participation active aux diverses communautés *EdTech*, ainsi que la création de leurs propres réseaux et de leurs centres de formation. Ces espaces d'échanges participent en outre d'une homogénéisation des compétences et référentiels pédagogiques au-delà des spécificités sectorielles. Les réseaux à but non-lucratif ou privés relaient auprès de leurs divers membres (publics éducatifs comme

entrepreneurs) des compétences compatibles avec la conception fédérale du *personalized learning*. Les *Big Tech* tentent aussi d'influer sur la construction de compétences techniques et pédagogiques, en imposant des normes d'usage propres à leurs outils.

Conclusion provisoire

Il est temps de revenir sur notre hypothèse esquissée dans le chapitre introductif de cette partie, selon laquelle les médias adaptatifs feraient advenir une socio-économie faite de complémentarités et de clivages nouveaux au profit des acteurs des technologies, et fédérée autour d'un projet éducatif transversal. Le cas états-unien, décliné sur ces trois niveaux interdépendants que sont les politiques publiques, les stratégies des firmes et les professionnalités, a précisé les implications socio-économiques de l'adaptativité. Au niveau des politiques publiques, une volonté d'implémentation d'un projet éducatif cohérent, le *personalized learning,* s'est construite à l'échelle fédérale selon un agenda politique lié à l'*Evidence-Based Education*. Celui-ci revisite la forme scolaire dans une veine socio-constructiviste et consacre une technologisation additive. La réappropriation de ce projet par les établissements éducatifs publics apparaît cependant erratique, en raison de la défiance de ces institutions quant à l'efficacité de ces technologies et leurs budgets contraints. Les stratégies des firmes, prisme crucial pour l'observation d'un hypothétique paradigme éducatif, font voir de nouvelles dynamiques socio-économiques. Des perspectives de croissance sont apparues pour les start-ups, jouissant de guichets de financements variés pour investir dans leurs algorithmes. Les acteurs plus anciens et concentrés issus du numérique éducatif et de l'édition scolaire optent majoritairement pour des acquisitions de start-ups. Les *Big Tech* sont quant à eux positionnés sur les trois principaux domaines générateurs de valeur dans la filière des médias adaptatifs et y exercent une influence importante : ils sous-traitent auprès des firmes éditrices des technologies servant au développement algorithmes d'apprentissage adaptatif et à la génération de contenus pédagogiques, et commercialisent leurs propres plateformes d'intermédiation dans les marchés éducatifs institutionnels comme non-formels. S'il est prématuré de conclure à la formation d'une nouvelle socio-économie stable, la montée en puissance des start-ups en matière d'innovations, la dynamisation de la filière par l'aval, et les tentatives d'imposition des *Big Tech* dans l'ensemble de l'industrie constituent trois tendances de fond. Elles se retrouvent dans la constitution de compétences et de solidarités professionnelles parmi les praticiens des diverses filières, qui tendent à s'accorder sur un projet proche du *personalized learning*.

Un de nos questionnements de départ refait alors surface : dans quelle mesure ces enseignements sont-ils inhérents à la culture éducative et à l'environnement industriel des États-Unis, ou constituent-ils des effets structurels du déploiement des médias adaptatifs dans l'éducation et la formation ? Notre méthode contrastive nous amène à présent à éprouver notre hypothèse de départ sur le cas français. Celui-ci présente des spécificités sur les trois niveaux conditionnant l'industrialisation des médias adaptatifs : les politiques publiques, les stratégies des firmes et les professionnalités. Elles sont susceptibles de jeter des éclairages différents sur les implications sociales et économiques de ce type de médias éducatifs.

Par exemple, sur le plan des valeurs politiques attenantes à l'éducation, le pays perçoit traditionnellement son École (au sens large) comme un vecteur culturel d'égalité des chances et d'émancipation plutôt que comme un instrument de préparation des individus à la vie sociale et économique comme aux États-Unis. Sur le plan des déclinaisons de la forme scolaire, la France se signale par une autonomie des enseignants plus poussée, avec une garantie de « liberté

pédagogique » jouant en faveur d'une forme d'artisanat, alors qu'aux États-Unis leurs pratiques sont bien davantage cadrées et propices à l'industrialisation. Les effets du déploiement des technologies adaptatives sur ces traditions méritent d'être questionnés. Enfin, sur le plan organisationnel, les politiques publiques hexagonales, fortes d'une tradition jacobine centralisatrice, encadrent étroitement le système éducatif, contrairement à la tradition libérale états-unienne. En découlent des marchés plus administrés en France et moins développés en termes de capitaux privés. Ces configurations socio-économiques très différentes constituent des ingrédients heuristiques en vue de démêler les enjeux structurels des contingences culturelles dans le déploiement de l'adaptativité.

Chapitre VI. La différenciation pédagogique française

Comme pour le chapitre précédent, il est d'abord nécessaire de spécifier l'organisation du système éducatif et de formation hexagonal, pour comprendre le rayon d'influence des principaux échelons (Ministères, académies, collectivités territoriales et établissements) et les différences vis-à-vis de l'appareil décentralisé états-unien. Ainsi, nous pourrons jauger la cohérence et la portée des discours stratégiques ministériels visant à soutenir et encadrer le déploiement des dispositifs algorithmiques, au regard du *personalized learning* américain. Nous observerons dans un deuxième temps les stratégies des firmes concourant à l'industrialisation des médias adaptatifs, et notamment leurs spécificités par rapport à celles de leurs homologues états-uniennes, eu égard aux différences de morphologies des industries éducatives. Enfin nous nous pencherons sur les réseaux *EdTech* français, les professions qui les composent et les idéaux pédagogiques qui les animent, en rapport avec l'influence exercée par les *Big Tech* sur les communautés aux États-Unis.

1. L'idéal ministériel de différenciation pédagogique, une vocation culturelle et politique

Les modes d'organisation de l'appareil éducatif et de formation

En France, les niveaux d'enseignement primaire, secondaire et supérieur peuvent davantage être appréhendés comme des systèmes, du fait de leur centralisation autour des Ministères et de leurs relais territoriaux, les académies. L'État assure le plus fortement son contrôle sur l'Éducation nationale, formée par les enseignements primaire et secondaire. Il le fait d'abord à travers la voie législative. Le Parlement vote des lois encadrant le fonctionnement des marchés du numérique éducatif, ainsi que des textes de programmation pour l'éducation. Ceux-ci incluent, en matière de numérique, des financements pour l'équipement des établissements et la formation des personnels dans l'ensemble du système éducatif. L'État exerce ensuite un contrôle exécutif via le Ministère de l'Éducation nationale et de la Jeunesse. Ses services internes comme la Direction du numérique pour l'éducation (DNE) publient plans et circulaires fixant des obligations liées au développement des technologies éducatives dans les écoles, collèges et lycées. Les directives ministérielles peuvent être distinguées entre politiques de l'offre et de la demande, les premières visant à structurer des industries nationales, par des appels d'offres et subventions de l'amont (production de contenus) à l'aval (distribution des logiciels, plateformes et terminaux informatiques), et les secondes consistant à développer les équipements et la connexion des établissements ainsi que les usages des publics scolaires.

L'État conduit une action prescriptrice forte mais plus souple dans l'enseignement supérieur : le Parlement vote des lois de programmation et de financement du système universitaire, et le Ministère de l'Enseignement Supérieur et de la Recherche conduit des politiques de l'offre et de la demande via des appels à projets réguliers à destination des éditeurs comme des établissements.

La formation professionnelle et la place qu'y occupent les technologies éducatives sont davantage laissées à la discrétion des organismes de formation du côté de l'offre, et des employeurs et employés du côté de la demande. Néanmoins le Parlement vote des lois fixant des obligations pour tous ces acteurs et encadrant leurs rapports de force. Le Ministère du Travail, de la Santé et des Solidarités

en assure l'application. Des mesures structurantes décidées récemment sont par exemple le Compte personnel de formation (CPF), ayant individualisé en 2017 la commande de formations au profit des salariés et demandeurs d'emploi, ou la « loi pour la liberté de choisir son avenir professionnel » de 2018, qui prévoit une obligation de certification des organismes de formation pour bénéficier de subventions, sur la base d'un référentiel national unique (Qualiopi). Par ailleurs, l'État peut être un guichet d'appels à projets pour l'innovation en matière de formations – témoin ceux lancés dans le cadre du « Plan d'Investissement dans les Compétences » pour la période 2018-2022.

Le maillage territorial est particulièrement développé pour assurer l'application des décisions étatiques dans les établissements, tout en laissant aux acteurs territoriaux une capacité de décision adaptée aux problématiques de leurs territoires. Un premier échelon est la région académique. Créées en 2020, les dix-huit régions académiques doivent mettre en cohérence les politiques éducatives régionales dans l'ensemble des niveaux, de l'éducation primaire et secondaire à la formation professionnelle (Ministère de l'Éducation Nationale et de la Jeunesse, 2024a).

Vient ensuite l'académie, circonscription chargée de conduire l'action du Ministère dans le territoire sous son autorité. Chaque académie parmi les trente existantes doit conduire les politiques ministérielles dans son territoire, sous l'autorité de la région académique. Elle dispose de sa propre direction compétente en matière de numérique éducatif, la Délégation Académique pour le Numérique Éducatif (DANE). Celle-ci conseille le recteur dans l'établissement d'un projet académique déclinant les orientations ministérielles en termes de numérique éducatif. Du côté de l'offre, l'académie contribue à des partenariats pour la production d'outils, de ressources ou de services numériques, en lien avec des universités et des entreprises de la filière (Eduscol, 2023a). Du côté de la demande, elle est prescriptrice en matière de formations et d'usages, établissant notamment un cahier des charges à destination des directeurs d'établissements et des collectivités pour leurs équipements en médias éducatifs. L'académie constitue également un intermédiaire entre les Ministères et les universités. Son recteur, chancelier des universités, est chargé d'« assurer la coordination des établissements d'enseignement supérieur et d'exercer un contrôle *a posteriori* sur les établissements » de l'académie (Académie d'Orléans-Tours, 2024).

Viennent ensuite les collectivités territoriales. Dans les échelons primaire et secondaire, elles sont chargées d'assurer le financement et la maintenance des terminaux, réseaux, et médias informatiques selon la répartition suivante : les mairies pour les écoles primaires, les départements pour les collèges et les régions pour les lycées. Si elles participent au financement de l'enseignement supérieur, les collectivités investissent peu dans l'équipement des universités en numérique éducatif, mais plutôt, du côté de l'offre, dans les transferts de technologies issus de recherches universitaires (Béroud, 2024).

Les établissements jouissent enfin, on l'a compris, de degré d'autonomie divers dans leur mise en place du numérique éducatif. Les plus cadrés sont les écoles, collèges et lycées. Après identification des besoins pédagogiques par les directeurs et les enseignants référents, des procédures d'achat sont opérées selon les cahiers des charges académiques. Dans l'enseignement supérieur public, les établissements disposent de davantage d'autonomie dans la gestion de leurs achats. Néanmoins la majeure partie de leurs budgets de fonctionnement provient de l'État – à hauteur de 81% pour l'année 2017 selon la Direction générale de l'enseignement supérieur et de l'insertion professionnelle du Ministère (Calviac, 2019). L'équipement en technologies se fait d'abord par l'identification des besoins par les départements de formation, puis par des appels d'offres lancés

auprès des éditeurs. Le plus souvent, des agences mutualisent ces procédures pour plusieurs établissements, comme le Groupe Logiciel ou l'Agence de Mutualisation des Universités et des Établissements d'enseignement supérieur ou de recherche. Sous l'autorité du Ministère, celles-ci ont pour mission de négocier avec les éditeurs des conditions tarifaires avantageuses via des achats groupés.

La configuration des systèmes éducatifs diffère donc grandement entre la France et les États-Unis. La place de l'État est plus forte dans l'Hexagone, tant du point de vue des financements que des prescriptions en matière de numérique éducatif. Celles-ci descendent jusqu'aux institutions éducatives via un maillage territorial étroit. Dans ce cadre singulier et centralisé, les lois de programmation votées au Parlement ainsi que les plans et circulaires ministériels contribuent de longue date à l'intégration des technologies numériques à l'École et dans les universités. Depuis les années 2010 et à la faveur de l'essor de l'intelligence artificielle, l'idéal de la différenciation pédagogique a innervé des expérimentations localisées, conduisant aux premières généralisations de médias adaptatifs à l'échelle nationale. Les facteurs politiques ayant conduit à la constitution d'un projet techno-éducatif ainsi que ses contours paradigmatiques sont traités ci-dessous suivant une approche historique et contrastive avec le cas états-unien.

Contours d'un projet officiel

L'émergence de la différenciation technologique comme idéal culturel et politique

L'intégration de technologies de différenciation dans le système éducatif et de formation français procède d'une matrice culturelle et politique. La première caractéristique, culturelle, des initiatives ministérielles se distingue des conceptions économiques et politiques du *Department of Education*. Comme on l'a vu plus haut, celui-ci perçoit les institutions d'enseignement primaire, secondaire et *a fortiori* supérieur comme des instruments pour préparer les nouvelles générations à l'appareil économique national, selon des enjeux de politique intérieure ou extérieure. La technologie fait figure de moyen industriel efficace pour atteindre ces objectifs. En France, « la vocation culturelle de l'éducation passe traditionnellement avant sa fonction instructive et productive » comme le rappelle Pierre Mœglin (2016, p. 265) :

> « [Un] consensus quasi officiel se réalise sur le fait que le but de l'éducation générale […] est de familiariser les individus avec des valeurs, de les plonger dans une culture – le plus souvent, la culture savante – ou de les immerger dans un univers de connaissances dont ils ont à s'imprégner pour atteindre un niveau. Et ce, plutôt que pour acquérir des aptitudes et compétences directement monnayables et utilisables sur le terrain de l'emploi » (ibid.).

Bien que l'École ne soit pas exempte de tentatives d'instrumentalisation économique, le déploiement des technologies éducatives en son sein répond à cet agenda culturel. La reconnaissance de technologies comme des « médias éducatifs », utilisables dans les cursus scolaires et universitaires, doit ainsi obéir aux valeurs cardinales de l'École (égalité, laïcité) et à des exigences qualitatives liées à la « culture savante ». Qui plus est, avec le numérique, nombre de discours d'experts et décideurs politiques sont traversés par une lame de fond techniciste considérant que les technologies d'information et de communication feraient advenir une « révolution » culturelle. Les élèves doivent être initiés à leur utilisation en vue d'acquérir une culture technicienne, incluant notamment une distance critique vis-à-vis du fonctionnement de ces médias.

La seconde vocation, à dominante politique, des initiatives ministérielles d'intégration technologique émane de la démocratisation conjuguée des systèmes scolaire et universitaire après-guerre. Celle-ci a introduit une hétérogénéité sociale au sein des classes et des amphithéâtres. Le défi qui scande depuis lors les politiques éducatives est d'« éviter que l'école redouble les inégalités sociales par des inégalités scolaires » selon le chercheur en sciences de l'éducation Sylvain Connac (Connac, 2021). Le défaut d'adaptation de l'enseignement collectif et simultané aux élèves en difficulté apparaît d'autant plus fortement, faisant émerger des velléités de « différenciation pédagogique ». Sylvain Connac voit ainsi dans la réforme Haby de 1975, actant la suppression des filières dans les collèges, un événement décisif pour l'inscription de cet idéal à l'agenda politique français. La différenciation est alors attendue à la fois sous une forme « structurelle » – via le redoublement ou l'orientation vers des filières à la sortie du collège – et « pédagogique », au sein de la classe, par la constitution de groupes de niveaux, d'ateliers de soutien pour les élèves en difficulté ou d'approfondissement pour ceux en avance par exemple. La technologie apparaît dès lors dans les discours officiels comme un ingrédient possible au service de la différenciation.

Les principaux programmes ministériels d'intégration de technologies à des fins de différenciation pédagogique répondent à cette matrice culturelle et politique. Témoin le plan « Informatique pour tous » (IPT) de 1985, visant à équiper les établissements scolaires du pays en terminaux, en vue d'initier à l'informatique tous les écoliers, collégiens et lycéens. Il est à l'époque motivé par l'engouement social et médiatique pour la « micro-informatique », qui a suscité dès les années 1970 quantité de discours soulignant les implications anthropologiques de cette technologie[101]. À travers IPT, l'intérêt ministériel pour les micro-ordinateurs s'attache également à leur apport possible à la différenciation. Dans une analyse de ce plan, Bernard Miège relate ainsi « la publication de circulaires prescrivant l'utilisation du micro-ordinateur "tuteur inlassable et respectueux des rythmes de chacun" (sic) pour l'enseignement des langues, des arts plastiques, éducation musicale, comptabilité, etc. » (Miège, 1996, p.95).

Avec l'émergence du web, la « cyber-formation » devient la nouvelle priorité des politiques éducatives du début des années 2000, appliquant à nouveau une perspective politico-culturelle. Elles prennent alors appui sur les discours d'experts annonçant les effets déterminants qu'auraient les développements du numérique dans l'histoire des sociétés humaines[102]. Les acteurs politiques commencent ainsi à soutenir la diffusion d'une « littératie numérique » tout au long de la vie des citoyens, de certifications sanctionnant les « compétences numériques » à l'instauration du « service public du numérique éducatif » par la loi pour la Refondation de l'École de la République en 2013. Celui-ci vise à « former au numérique les citoyens et les professionnels de demain, qui ne pourraient sans cela ni exercer la plupart des métiers, ni même s'insérer dans la société » (Cour des comptes, 2019). Le numérique éducatif a également fait l'objet de discours préconisant son intégration comme outil de « transformation pédagogique ». La Cour des comptes, en préambule de son rapport sur le service public du numérique éducatif, explique qu'« il s'agit aussi d'utiliser les techniques numériques comme un puissant levier de transformation pédagogique, susceptibles de faire évoluer les méthodes d'enseignement, d'améliorer les apprentissages et de permettre l'évaluation en continu des résultats, *in fine* de favoriser la réussite des élèves » (Cour des comptes, 2019, p.11).

[101] En témoignent les descriptions d'une « société de l'information » qui découlerait du déploiement généralisé de ces technologies, comme dans le rapport d'Alain Minc et Pierre Nora de 1978 intitulé *L'informatisation de la société.*

[102] Cette refonte du social proviendrait pour beaucoup d'experts de la dissociation, introduite par le numérique, entre l'information et son support – « révolution fondamentale, peut-être encore plus importante à terme que l'imprimerie » selon l'informaticien et membre du Conseil Scientifique de l'Éducation Nationale Gérard Berry (Berry, 2008).

L'intelligence artificielle a dernièrement connu une légitimation semblable. Le rapport de mission parlementaire commandé au député Cédric Villani en 2018, ayant servi à esquisser la « Stratégie nationale pour l'intelligence artificielle », contient une rubrique intitulée « Enseigner à l'heure de l'IA ». Y est d'emblée décrit le tournant sociétal que représenterait l'intelligence artificielle, ainsi que les responsabilités qui en découleraient pour les systèmes éducatifs : « Dans une société automatisée, se former en permanence sera une nécessité. Dès lors, tous les acteurs de l'enseignement sont appelés à se transformer pour inscrire l'activité d'apprentissage dans une approche "tout au long de la vie" » (Villani, 2018, p.185). Et d'appeler à une refonte des méthodes vers davantage de personnalisation, grâce aux possibilités offertes par l'IA et le *Big Data* : « Dans cette perspective, l'intelligence artificielle ouvre de nouvelles opportunités pour former un grand nombre d'individus de manière personnalisée et adaptative » (ibid.).

À l'image de la micro-informatique et de la cyber-formation avant elle, l'intelligence artificielle s'inscrit dans le défi d'une démocratisation de l'enseignement en même temps qu'une élévation du niveau scolaire de tous les enfants, notamment défavorisés ou en situation de handicap. Néanmoins la différenciation technologique devient plus ambitieuse que celle rapportée par Bernard Miège à travers IPT, qui s'apparentait à de l'adaptabilité (adaptation du dispositif au rythme de l'utilisateur), ou que celle attribuée par la Cour des comptes aux technologies numériques, évoquant simplement leurs évaluations en continu. Les perspectives de renouvellement de la forme scolaire sont fondamentales : « Concrètement, les élèves n'auraient plus à travailler en même temps sur les mêmes exercices […]. Cela peut même aller jusqu'à des organisations davantage modulaires, c'est-à-dire qui poussent la personnalisation jusqu'à penser l'organisation des cours autour de ce principe » (ibid.).

Des menaces viennent cependant obscurcir le propos, comme la perte d'expériences collectives : « Ces services ne peuvent en aucun cas se substituer à l'expérience collective d'apprentissage vécue par exemple en classe, ni être perçus comme neutres d'un point de vue pédagogique » (ibid.). Cette posture à la fois volontariste et précautionneuse s'accorde avec celle de la Commission européenne, qui, dans son *Digital Education Action Plan* de 2018, souligne parmi les perspectives ouvertes par les *Learning Analytics* le besoin de mener des recherches sur leurs applications pédagogiques :

> « L'analyse de l'apprentissage permet d'améliorer l'apprentissage personnalisé, par exemple en repérant les élèves vulnérables, et d'évaluer l'incidence des différentes stratégies d'enseignement. Toutefois, étant donné que l'analyse de l'apprentissage en est encore à ses débuts en Europe, nous avons besoin de davantage d'actions pilotes pour mener des recherches et des expérimentations dans ce domaine »[103] (Commission européenne, 2018, p.12).

L'intérêt étant formulé, les Ministères en charge du système éducatif et de formation vont-ils jusqu'à reconnaître et encourager un renouvellement paradigmatique autour de ces technologies, comme leur homologue états-unien avec le *personalized learning* ?

[103] Ces dernières années ont vu une production abondante de la Commission européenne au sujet de l'IA dans l'éducation, notamment à travers son Centre commun de recherche : témoin les deux rapports intitulés *« The impact of Artificial Intelligence on Learning, Teaching and Education »* en 2018, et *« Emerging technologies and the teaching profession - Ethical and pedagogical considerations based on near-future scenarios »* en 2020.

Une stratégie ministérielle de recherche-action préalable aux généralisations

Avant de caractériser un projet éducatif, diverses expérimentations de formes de différenciation technologique émergent à partir des années 2010[104]. Dans le sillage des préconisations de la Commission européenne, les Ministères ont en effet adopté une stratégie de « recherche-action » en matière de numérique et d'IA éducative. Celle-ci consiste premièrement dans le lancement d'appels à projets à destination de structures de production informatique, de laboratoires et d'établissements pilotes, en vue de concevoir et expérimenter des prototypes numériques. Cette stratégie suit le modèle états-unien du « *demonstrator* », employé pour les *AI Institutes* comme on l'a vu plus haut. Ces expérimentations visent cependant à faire émerger, par des protocoles de recherche-action, des « bonnes pratiques » généralisables à l'échelle nationale. Ce choix de la recherche-action diffère ainsi de la stratégie fédérale états-unienne, qui encourage davantage, à travers l'ESSA et ses « *four tiers of evidence* », des études d'efficacité randomisées. Il s'en rapproche en revanche dans la mesure où les attentes ministérielles s'inscrivent dans l'*Evidence-Based Education.*

Des expérimentations aux généralisations de technologies de différenciation

La première expérimentation d'ampleur autour de la différenciation s'intitule « espaces de formation, de recherche et d'animation numériques » (e-FRAN) en 2015. Elle est le fruit de la mission interministérielle confiée auparavant au recteur Jean-Marc Monteil, qui encourageait le lancement d'un démonstrateur mobilisant à la fois l'État, les collectivités et les acteurs scolaires. Inclus dans le plan « Numérique pour l'éducation » consécutif à la loi pour la Refondation de l'École de la République de 2013, elle est financée par les Programmes d'investissement d'avenir (PIA). Sur le terrain, des Sociétés d'Accélération des Transferts de Technologies (SATT), créées dans le cadre du premier PIA, sont chargées de financer les transferts d'innovation des laboratoires universitaires vers les écoles participant à ce programme.

L'appel à projets lancé en amont met en exergue, parmi les innovations espérées, « le numérique comme outil pédagogique », et mentionne la personnalisation des enseignements comme l'une de ses facettes (Caisse des Dépôts et Consignations, 2015, p.8). Dans la vidéo de présentation de l'appel, le recteur justifie le lancement d'e-FRAN par une critique de la verticalité traditionnelle des appels et l'oubli des usages qui en découlerait :

> « Cet appel à projets n'est pas la conséquence d'un processus descendant, mais au contraire on s'adresse au terrain. […] Cette mission se déplace sur le terrain pour faire en sorte que les acteurs fassent remonter un certain nombre d'expériences, un certain nombre d'idées créatives sur l'utilisation du numérique pour éduquer la société dans son ensemble […] » (Monteil, 2016).

Il justifie également le programme par la volonté de mesurer l'efficacité des innovations avant d'investir dans celles-ci :

[104] Nous n'abordons pas ici la politique d'expérimentations de l'Union européenne, par ailleurs très dense en matière d'appels à projets visant des formes de personnalisation éducative. Quelques exemples d'appels en ce sens : dans le cadre du programme d'investissement Horizon 2020, *« Technologies for better human learning and teaching* » en 2014 et « *Technologies for Learning and Skills* » en 2015 ; dans le cadre d'Eramus+ « *Capacity building in the field of Youth* » en 2019 ; et dans le cadre du *Digital Education Action Plan 2022-2027*, l'appel « *Supporting the development of the Digital Education Hub* » en 2021. À noter également le projet Erasmus+ d'édition d'un manuel libre sur l'IA à destination des professeurs européens, intitulé AI4T *(Artificial Intelligence for Teachers)* en 2021.

> « Ce qui est nouveau c'est qu'on veut mesurer l'efficacité des apprentissages avec le numérique. Et pour ça il faut qu'on ait une approche scientifique, c'est-à-dire très méthodologique, très rigoureuse, qui permette de dire "en effet on a introduit un nouvel outil qui est très beau, très luxuriant mais est-il efficace pour les apprentissages ?". Il faut le savoir. Si on doit investir lourdement dans une technologie, il faut savoir si cette technologie est évidemment productive » (ibid.).

L'ambition est faussement novatrice au vu de la longue histoire du *Social Efficiency Movement* puis de l'*Evidence-Based Education,* notamment aux États-Unis. Vingt-deux projets sont au final soutenus, pour un montant total de 19,5 millions d'euros (Ministère de l'Éducation Nationale et de la Jeunesse, 2021a).

Dès 2017, l'intégration de l'IA adaptative a été programmée sur ce même modèle. Le Ministère de l'Éducation nationale et de la Jeunesse lance en juillet de la même année le « Partenariat d'Innovation Intelligence Artificielle » (P2IA). La présentation de l'appel à projets en résume ainsi le but : « développer des solutions, basées sur l'intelligence artificielle, à destination des enseignants pour les classes de l'école primaire et en particulier du cycle 2, afin de mieux accompagner leurs élèves dans les apprentissages du français et/ou des mathématiques » (Ministère de l'Éducation nationale et de la Jeunesse, 2017, p.4).

À leur création en 2014, la Direction des Affaires Juridiques du Ministère de l'Économie explique que « les partenariats d'innovation sont des marchés publics qui ont pour objet la R&D ainsi que l'acquisition de la solution innovante qui en résulte » (Direction des Affaires Juridiques du Ministère de l'Économie, 2014, p.2). Ils ont été créés pour « pallier les difficultés structurelles des […] marchés de recherche et de développement (R&D) qui imposent une remise en concurrence à l'issue de la phase de R&D pour pouvoir acquérir les produits, services ou travaux innovants qui en sont le résultat » (ibid., p.1). Le « partenariat d'innovation intelligence artificielle » stipule pour sa part que, loin d'une « solution sur étagère », est attendue une « solution disruptive, qui n'existe pas encore ou qui doit être très fortement adaptée » (Ministère de l'Éducation nationale et de la Jeunesse, 2017, p.9).

Au sein de la Direction du numérique pour l'éducation, le chef du bureau de soutien à l'innovation numérique et à la recherche appliquée, Axel Jean, est revenu en 2022 sur la chronologie de ce démonstrateur (Jean, 2022). Lors de la première phase de recherche et développement et de prototypage, la progression a dû s'opérer par itérations entre trois pôles : le laboratoire, l'entreprise *EdTech* et le terrain scolaire où ont été menées les expérimentations. Des délégations académiques au numérique éducatif (DRANE et DANE), des inspecteurs de l'éducation nationale chargés de la mission numérique départementale et des enseignants référents ont accompagné en outre les équipes pédagogiques des écoles volontaires. Est venue ensuite une phase de « pré-industrialisation » pour amender le produit en fonction des résultats et retours académiques. La dernière étape a été le « déploiement/exploitation en fonction de la maturité des services réalisés pour les professeurs au niveau national pour 4 années » (ibid., p.12). En définitive la personnalisation visée pour ces outils est « du niveau de personnalisation que peut apporter un professeur particulier » (ibid., p.16) – référence au préceptorat, esquissant une forme de technologisation substitutive.

L'objectif de faire de l'école et de l'université des laboratoires d'innovations coïncide avec celui de « scientificiser » l'éducation, pour stimuler la demande dans ces technologies. L'ambition est poursuivie en janvier 2018 par la création du Conseil Scientifique de l'Éducation nationale (CSEN). Doté d'un pouvoir consultatif auprès du ministre de l'Éducation nationale, il est « composé de chercheurs reconnus issus de différentes disciplines scientifiques » (Conseil Scientifique

de l'Éducation nationale, 2019, p.10), prenant part à onze groupes de travail, aux côtés de représentants du ministère et d'autres administrations, ainsi que de chercheurs invités. Ces groupes ont pour mission « d'identifier les atouts et les difficultés constatés sur [un] sujet et proposer des recommandations pour améliorer le système scolaire. Ces recommandations peuvent prendre la forme de rapports, conférences, outils pédagogiques, MOOC etc. disponibles. » (Réseau Canopé, 2024a).

Le président du CSEN, Stanislas Dehaene, est à l'origine un psychologue spécialisé en neuropsychologie, dont les travaux portent sur les représentations cérébrales des activités de calcul, de lecture, de langage, et plus globalement de conscience. Directeur depuis 2008 de l'unité de neuro-imagerie cognitive NeuroSpin, il est devenu professeur au Collège de France à la chaire de psychologie cognitive expérimentale en 2015. Dans des interviews accordées à divers médias avant la création du CSEN, il explique sa vision de l'éducation. Notamment son idéal d'objectivité et d'efficacité au service des valeurs républicaines :

> « Ce qui me motive, c'est l'idée d'agir pour l'éducation des jeunes, indépendamment de toute idéologie […]. Beaucoup d'enfants de milieux défavorisés pourraient avoir un avenir brillant, mais ils en sont privés parce qu'ils n'ont pas bénéficié des enseignements adaptés » (Rosier, 2018).

On retrouve dans ces propos le consensus culturel et politique des politiques d'intégration de technologies individuelles à l'École telle qu'esquissé plus haut. Dans une autre interview, Stanislas Dehaene place la quête d'intervention différenciée parmi les priorités du CSEN : « il faut à la fois avoir des outils d'évaluations et tout de suite, derrière, avoir des outils d'intervention » – « soit des outils informatisés soit des outils pédagogiques standards pour adresser le problème particulier de cet enfant-là » (Radiofrance, 2018). Un objectif traité par le groupe de travail 1 intitulé « Évaluations et interventions », et surtout, dans la voie de l'automatisation, par le groupe 6 « Intelligence artificielle et numérique pour l'éducation ».

En mars, peu après la création du CSEN, est publié le rapport de Cédric Villani devant permettre d'esquisser la « Stratégie nationale pour l'intelligence artificielle ». Outre les préconisations vues plus haut, le député formule le souhait d'expérimentations appuyées par la recherche : « Notre choix de mettre cette action en première place ne traduit pas une volonté de précipiter cette mise en œuvre ; elle traduit plutôt l'urgence à développer des expérimentations » (Villani, 2018, p.185). Il reprend ainsi l'argumentaire de la Commission européenne ou de Jean-Marc Monteil justifiant les expérimentations par le besoin de connaître les implications des technologies et les pratiques qui « fonctionnent » ou non. Selon lui, ce besoin ne serait pas suffisamment traité par les sciences de l'éducation :

> « Les expérimentations doivent être accompagnées de protocoles adaptés et surtout d'une évaluation systématique, qui fait trop souvent défaut. Si le développement des expérimentations doit être un axe majeur, les évaluations doivent pouvoir prévenir les désillusions et dégonfler les bulles, qui sont nombreuses dans ce domaine » (Villani, 2018, p.186).

L'appel de 2020 pour la deuxième vague des « groupements thématiques numériques » (GTnum) porte cette volonté. Lancés pour la première fois en 2017 pour trois ans par la Direction du numérique pour l'éducation, ces groupements suscitent en 2020 des attentes claires de la part de la DNE :

> « Les groupes thématiques numériques sont animés par des laboratoires universitaires ou autres opérateurs publics dont les missions comprennent un travail de recherche et/ou de veille et de médiation scientifiques.

> Leur travail doit s'inscrire dans une approche interdisciplinaire et systémique du numérique dans l'éducation » (Direction du numérique pour l'Éducation-TN2, 2020, p.2).

L'un des axes de travail s'intitule : « Intelligence artificielle et éducation : bien utiliser le potentiel d'assistance sans abandonner la décision aux machines ». La technologisation additive apparaît comme une attente officielle pour ces dispositifs : l'intelligence artificielle doit assister le professeur, et non se voir déléguer des décisions. L'axe se décompose en deux groupes : le GTnum 8 « Prédiction, anticipation, recommandation » et le GTnum 9 « L'IA pour de nouveaux modes d'interaction, de nouveaux modes d'évaluation et une hybridation des environnements d'apprentissage (empathie, examens, parcours, certification…) ». Le Ministère attend ainsi une expertise quant aux mutations à envisager dans des aspects structurels de l'éducation comme le diagnostic, l'évaluation ou l'interaction.

La pandémie de Covid-19 renforce le besoin de continuité pédagogique dans les écoles et les universités. En juin 2020, l'Agence Nationale de la Recherche publie l'appel « Hybridation des formations de l'enseignement supérieur » pour inciter les universités à préparer pour la rentrée de septembre des cursus flexibles entre présentiel et distanciel, mobilisant notamment les *learning analytics* (Agence Nationale de la Recherche, 2020).

Dans l'enseignement primaire et secondaire, face au manque d'intégration du numérique dans les pratiques éducatives – déjà souligné avant la Covid-19 par la Cour des Comptes (2019), puis révélé crûment lors de la fermeture des établissements –, le gouvernement organise les « États généraux du numérique pour l'éducation » en novembre 2020. Le but est, par une consultation de l'ensemble de la communauté éducative, de faire émerger une stratégie nationale en la matière. Quarante propositions en découlent, dont une appelant à une extension du démonstrateur « Territoires numériques éducatifs » (TNE). Débuté à la rentrée 2020 dans deux départements (Aisne et Val-d'Oise), il est dès lors étendu à dix nouveaux départements à la rentrée 2021 (Ministère de l'Éducation nationale et de la Jeunesse, 2020). Les TNE visent à tester, à plus grande échelle que celle des démonstrateurs précédents, la mise en œuvre de la continuité pédagogique. Le montant prévu dans le cadre de PIA4 puis de France 2030 totalise 27,3 millions d'euros pour l'accompagnement des TNE (Académie de Nancy-Metz, 2024), et 25 millions d'euros pour un marché public d'acquisition d'outils numériques (Ministère de l'Éducation Nationale et de la Jeunesse, 2022). En comparaison avec les 19,5 millions d'euros alloués par les pouvoirs publics aux vingt-deux projets e-FRAN, l'effort ministériel est bien plus prononcé. Derrière les deux ambitions phares avancées par le Ministère – « l'élévation générale du niveau et une plus grande justice sociale » (ibid., p.1) –, plusieurs objectifs de renouvellement méthodologique sont présents :

- Du côté des ressources, la « personnalisation » : « la mise à disposition d'une offre de logicielles et de ressources, pour préparer la classe, personnaliser son enseignement et les apprentissages des élèves » (ibid.) ;
- Sur le volet « formation », « permettre un changement des pratiques enseignantes, en particulier l'hybridation, et mesurer l'impact sur ces dernières » (ibid., p.2) ;
- Sur le volet « parentalité », « accompagner les parents pour co-construire une continuité pédagogique, s'acculturer aux usages numériques raisonnés, éclairés et responsables » (ibid., p.1).

Publié en décembre 2020, l'appel « Démonstrateurs de l'enseignement supérieur » (DemoES) vise à « expérimenter en vraie grandeur et dans une nouvelle approche globale, toutes les dimensions

de la transformation numérique et en les prenant en compte de façon simultanée » dans « un ensemble d'établissements représentatifs de la diversité de l'enseignement supérieur français » (Agence Nationale de la Recherche, 2021, p.5). Ces dimensions incluent, pêle-mêle, la « transformation des cursus pour intégrer cette nouvelle approche (approche compétences, modularité et individualisation des cursus, internationalisation, etc.) », le « changement dans le processus d'évaluation des étudiants », l'« équipement des salles et bâtiments, des enseignants, des étudiants pour qu'ils puissent étudier où qu'ils soient », la « formation des enseignants à l'utilisation du numérique, des étudiants à apprendre à apprendre avec le numérique », ou encore la « politique de vie étudiante » (ibid., pp.5-6). Le programme a accordé cent millions d'euros aux dix-sept démonstrateurs sélectionnés (Livingstone-Wallace et Welter, 2024). L'effort financier a presque doublé par rapport aux TNE, qui cumulaient 52,3 millions d'euros. DemoES vise une « transformation » de chaque territoire bien plus ambitieuse et fine que ce qu'auraient permis les politiques traditionnelles :

> « À ce titre, [les démonstrateurs] constituent une approche méthodologique très différente du déploiement massif sur l'ensemble du territoire de matériel ou de solutions, ne traitant à chaque fois qu'une partie des freins à cette transformation. La diversité des démonstrateurs et, par conséquent, des territoires vise à bien prendre en compte les différences qui peuvent exister entre les différents territoires en termes technologiques (accès et connectivité), géographiques (répartition spatiale des élèves), économiques et sociologiques (capacité des parents ou des collectivités territoriales à contribuer à la transformation), culturels (importance donnée à l'éducation par les familles ou les territoires), organisationnels (les régions, départements, communes et communautés de communes n'ont pas tous opté pour les mêmes schémas d'organisation) » (Agence Nationale de la Recherche, 2021, p.13).

Une critique du mode antérieur de déploiement, à la fois descendant, massif, spécialisé sur un aspect et indifférent aux problématiques territoriales, justifie l'annonce de ces démonstrateurs désormais capables de prendre en compte les besoins multiples des territoires, pour diffuser des « bonnes pratiques » à l'échelle nationale. Pour parvenir à cette prise en compte, le partenariat d'innovation est une nouvelle fois « particulièrement encouragé » avec des « entreprises de la *Edtech*, de façon à améliorer l'efficience de l'enseignement au sein du démonstrateur et développer l'écosystème français » (ibid., p.8). Outre l'« efficience » que sont censés apporter ces acteurs économiques, l'idée est d'investir dans une industrie française souveraine en matière de données : « il s'agit non seulement de soutenir les filières française et européenne, mais également d'assurer la souveraineté nationale sur les données produites et sur les solutions proposées, ainsi que de favoriser, à terme, l'exportation de solutions françaises ou européennes à l'international » (ibid.). Pour reprendre la distinction opérée plus tôt, des objectifs endogènes et exogènes à l'éducation sont présents : des volontés de transformation pédagogique se mêlent à une stratégie industrielle et politique, dans un projet global particulièrement ambitieux.

Résultant de ces « couches successives » de soutiens publics à la filière et aux usages, un cadre stratégique s'est peu à peu cristallisé au sommet de l'État. Le Ministère de l'Éducation nationale a publié en 2023 sa stratégie du numérique pour l'éducation pour une période courant jusqu'en 2027 (Ministère de l'Éducation nationale et de la Jeunesse, 2023a). Qui plus est, les fruits des expérimentations menées dans les démonstrateurs tendent aujourd'hui à être généralisés nationalement, à partir de ce que le Ministère de l'Éducation Nationale considère être des « preuves d'efficacité ».

C'est ce qui est advenu en décembre 2023 lorsque le Ministre de l'époque, Gabriel Attal, a annoncé la généralisation du produit « MIA Seconde » pour une remédiation et un accompagnement en français et mathématiques auprès des lycéens de seconde. Il s'est agi d'une distribution gratuite

à partir de février 2024 dans plusieurs académies, avant une généralisation à l'ensemble des élèves de seconde à la rentrée 2024, selon le dossier de presse accompagnant cette annonce (Ministère de l'Éducation nationale et de la Jeunesse, 2023b, p.26).

C'est l'aboutissement d'une procédure de développement démarrée en 2021, sous le ministère de Jean-Michel Blanquer. Dans le « cahier des clauses techniques particulières » préalable à l'appel à projets, il est alors précisé que l'outil doit s'inscrire dans le sillage du test de positionnement organisé depuis la rentrée 2018 par la Direction de l'évaluation, de la prospective et de la performance (DEPP) du Ministère pour tous les élèves de seconde, « pour permettre à chacun de se situer en français et en mathématiques » (Ministère de l'Education Nationale, de la Jeunesse, 2021b, p.5). Il est ajouté que « ce positionnement a vocation à aider les professeurs à mettre en place l'accompagnement personnalisé, adapté à chaque élève, pour consolider ou approfondir la maîtrise de compétences en français et en mathématiques » (ibid.). Le « service de remédiation et d'accompagnement » espéré doit donc être un outil à disposition des élèves et des enseignants pour poursuivre la personnalisation à partir de ce test (ibid., p.7).

En août 2022 a été annoncée la nomination d'un consortium pour le développer. Il était composé de Docaposte pour l'hébergement, le pilotage, le développement, le maintien en condition opérationnelle et l'assistance-utilisateur, rétribué à hauteur de près de 750 000 euros, et Evidence B – start-up déjà lauréate du P2IA – pour la conception des modules pédagogiques, pour près de 2 millions d'euros (Clavey, 2024).

Un an plus tard, Gabriel Attal annonce finalement la généralisation de l'outil comme une des mesures de la mission « Exigence des savoirs ». Lancée le 5 octobre 2023, celle-ci est coordonnée par quatre personnalités dont Stanislas Dehaene, président du conseil scientifique de l'éducation nationale (CSEN) – une présence marquant une continuité idéologique dans la politique ministérielle. La mission est motivée comme aux États-Unis par les mauvais résultats des élèves à l'enquête PISA : entre 2018 et 2022, la performance des élèves en France ont en effet baissé de 21,5 points en mathématiques (Bernigole et al., 2023), de 19 points en compréhension de l'écrit et de six points en culture scientifique (Bret et al., 2023). Alors que la généralisation de l'outil est prévue depuis 2021, le ministre lui donne un nouveau sens, en l'inscrivant dans une posture proactive mais précautionneuse de la France vis-à-vis de ces technologies. Le pays doit ainsi devenir un pionnier dans la course mondiale à l'IA selon les propos du ministre, empreints des préceptes de l'*Evidence-Based Education* :

> « Je souhaite que l'intelligence artificielle puisse être pleinement utilisée, au service de l'élévation du niveau, et dans les apprentissages de nos élèves. […] En matière d'intelligence artificielle, […] je préfère choisir plutôt que subir. Beaucoup d'outils ont été développés – certains de nos élèves les utilisent – et ça ne veut pas dire que ce sont toujours les meilleurs outils, ceux qui ont été prouvés par des chercheurs, par des enseignants, par des scientifiques, comme étant les plus efficaces au service de l'élévation de leur niveau. […] La France sera le premier pays au monde à mettre à disposition et à généraliser un outil fondé sur l'IA à tous les élèves d'une classe d'âge » (Ministère de l'Éducation nationale et de la Jeunesse, 2023c, 28:46 – 29:36).

De cette stratégie de recherche-action aux débuts d'une généralisation nationale, apparaissent des permanences constitutives d'un projet, au croisement d'une personnalisation de l'éducation et de l'usage de l'intelligence artificielle. Des similarités avec la stratégie fédérale états-unienne apparaissent dans la mise en place de démonstrateurs, la volonté d'employer les technologies de manière efficiente et de soutenir une industrie nationale de l'IA.

Des divergences affleurent en revanche dans les modes d'organisation horizontaux propres à la « recherche-action » et son débouché, la généralisation d'outils à l'échelle nationale, permise par la centralisation du système éducatif autour du Ministère. En outre, il nous est apparu que les Ministères n'adoptaient pas la même stratégie de production de normes techno-éducatives que le *Department of Education*. Alors que celui-ci s'est efforcé d'expliciter un projet paradigmatique cohérent à destination des organisations éducatives du pays, les Ministères français n'ont pas encore entrepris de « graver dans le marbre » un projet alternatif à la forme scolaire. Nous tentons néanmoins de reconstituer les permanences paradigmatiques d'un projet officiel latent, à partir des appels à projets et des dernières publications stratégiques évoquées à l'instant.

Les contours d'un projet techno-éducatif latent

Nous synthétisons les éléments paradigmatiques contenus dans les différents documents ministériels évoqués plus haut, ayant en commun de traiter directement de différenciation pédagogique. Ce corpus est donc composé :

- des quatre appels à projets ministériels « Espaces de formation, de recherche et d'animation numériques » (e-FRAN) et « Partenariat d'Innovation Intelligence Artificielle » (P2IA) pour l'enseignement primaire et secondaire, et « Démonstrateurs numériques dans l'enseignement supérieur » (DemoES) et « Hybridation des formations de l'enseignement supérieur » pour le supérieur.
- des documents (verbatim, appels à projet) liés à la création, par le Ministère de l'Éducation nationale, de deux instances productrices de savoirs sur l'IA éducative : le Conseil Scientifique de l'Éducation nationale en 2018, et les Groupements Thématiques Numériques en 2020.
- du cahier des charges du service de remédiation et d'accompagnement en seconde.
- de la stratégie « Numérique pour l'éducation 2023-2027 ».
- du dossier de presse de la stratégie « Choc des savoirs ».

Il s'agit de circonscrire dans ces discours pluriels un projet cohérent axé sur la différenciation pédagogique, qu'elle soit permise par les technologies numériques ou l'intelligence artificielle spécifiquement, et ce pour l'enseignement primaire, secondaire et supérieur. Pour prolonger notre méthode contrastive avec le cas états-unien, nous reprenons les cinq piliers donnés par *l'Office of Educational Technology* pour le *personalized learning* : Apprentissage, Enseignement, Évaluation, Gouvernance et Infrastructure.

Section	Enseignements primaire et secondaire	Enseignement supérieur
Apprentissage	**Appel e-FRAN, contexte :** « Afin de préparer tous les enfants à vivre dans une société devenue numérique, [l'École] doit rendre chacun capable de comprendre, choisir, et utiliser de façon avertie et critique les technologies numériques pour mieux les maîtriser et même *in fine* apprendre à les concevoir » (p.5). **Appel P2IA, objectifs :** « Accompagner les élèves et leurs parents en dehors de la classe, pour démultiplier les opportunités d'apprentissage » (p.6)	**Appel DemoES, contexte :** « Apprendre à apprendre et à réussir avec le numérique : offrir à la diversité des apprenants (étudiants en formation initiale, adultes en reprise d'études, apprentis, étudiants en mobilité et étudiants internationaux) des dispositifs pédagogiques numériques diversifiés et adaptés à leurs besoins (personnalisation) et à leurs contraintes : hybride ou totalement distant, simulation de gestes professionnels ou d'expérimentation (TP : technologies immersives réalité virtuelle, augmentée, jumeaux numériques)... » (p.6)

	Stratégie « Choc des Savoirs », cas d'usages de l'intelligence artificielle pour l'apprentissage des langues étrangères : « Utilisateur débutant (A1-A2) : L'IA au service de la différenciation pédagogique ; Utilisateur indépendant (B1-B2) : L'IA au service des approches métacognitives et de l'évaluation formative ; Utilisateur expérimenté (C1-C2) : L'IA au service de la créativité et du dépassement de soi » (p.25) **Cahier des charges du service de remédiation et d'accompagnement en seconde, contexte :** « Ce service permet de nouvelles modalités d'apprentissage qui engagent tous les élèves vers les compétences et les savoirs nécessaires pour développer leur créativité et leur citoyenneté et permettre leur insertion professionnelle dans la société du XXIe siècle » (p.5)	**Appel « Hybridation des formations de l'enseignement supérieur », projets attendus :** « La création d'activités pédagogiques en ligne collaboratives (projets de groupes à distance, webinaires de présentation de travaux d'étudiants, ateliers distanciels d'évaluation par les pairs, wiki poly de cours, wiki exercices, ...) » (p.6)

Concernant l'apprentissage aux niveaux primaire et secondaire, notre corpus reprend la dialectique du changement culturel auquel l'École devrait répondre. Celle-ci devrait préparer les élèves à l'avènement généralisé des technologies numériques en leur permettant de comprendre, sélectionner et utiliser celles-ci de façon éclairée, comme le souligne l'appel e-FRAN. À la focale culturelle du changement pédagogique s'ajoute un objectif utilitariste d'insertion professionnelle, à consonance états-unienne, que révèle le cahier des charges du service de remédiation et d'accompagnement en seconde. L'IA est dépeinte comme un outil pouvant occuper des fonctions pédagogiques variées, et notamment permettre à l'élève de développer sa créativité et de « se dépasser », selon la stratégie du « Choc des savoirs », pour mieux s'insérer socialement. Transparaissent là des valeurs néolibérales de performance personnelle. Dans l'enseignement supérieur sont mis en avant deux enrichissements apportés par le numérique : d'une part la démultiplication quantitative des « opportunités » d'apprentissage grâce à la levée des clôtures spatiales et temporelles de la classe, et d'autre part leur diversification qualitative, selon les besoins éducatifs et les types d'apprentissage, *a fortiori* collaboratifs.

Enseignement	**Appel P2IA, objectif :** « Aider les enseignants à identifier les difficultés précises de chacun de leurs élèves […], aider l'enseignant à constituer des regroupements d'élèves, selon la stratégie pédagogique qu'il souhaite mettre en œuvre » (p.6) **Appel GTnum 2020-2022, axe 4 :** « Intelligence artificielle et éducation : bien utiliser le potentiel d'assistance sans abandonner la décision aux machines » (p.5) **Cahier des charges du service de remédiation et d'accompagnement en seconde, « exigences fonctionnelles du service » :** « Ce service assiste le professeur dans son action continue de suivi des apprentissages et de remédiation, sans le remplacer. Pour cela, le service l'informe d'une part de l'utilisation par ses élèves ainsi que de leur progression, et d'autre part lui recommande des pistes pédagogiques complémentaires permettant d'orienter	**Appel DemoES, objectif :** « Former les enseignants et les enseignants-chercheurs ainsi que développer et professionnaliser les fonctions d'appui en matière d'ingénierie pédagogique, en prenant appui sur les résultats de la recherche » (p.6).

	le choix des activités à proposer sous forme numérique ou non, en fonction d'un objectif donné : remédiation ciblée, entraînement spécifique, consolidation, approfondissement… » (p.9). **Stratégie « Numérique pour l'éducation 2023-2027 » :** « les enseignants doivent pouvoir à la fois s'appuyer sur des outils souverains, sécurisés, libres et communautaires, c'est-à-dire des « communs numériques », leur permettant de co-construire et partager entre pairs leurs productions…». (p.24) **Stratégie « Choc des savoirs », enjeu** : « L'implication des enseignants pour concevoir les outils pédagogiques » (p.27).	

Les implications du numérique et de l'IA sur le métier d'enseignant sont surtout documentées dans l'enseignement primaire et secondaire. Le recours à la technologie y est souhaité sous forme de technologisation « additive », comme dans la conception états-unienne : la technique doit assister l'enseignant et non le déposséder de son pouvoir de décision. L'enseignant doit pouvoir suivre la progression et les difficultés de chaque élève, pour venir en aide à ceux en difficulté, grâce aux recommandations de la machine. Les textes confirment ainsi le principe de « liberté pédagogique » de l'enseignant inscrit dans le Code de l'Éducation[105]. L'appel du P2IA explique vouloir aider l'enseignant à constituer des regroupements d'élèves, mais nuance son pouvoir de prescription en rétablissant sa liberté de choisir sa méthode : «… selon la stratégie pédagogique qu'il souhaite mettre en œuvre » (p.6). Comme aux États-Unis, les dernières stratégies ministérielles mettent l'accent sur la mobilisation des enseignants dès la conception des outils et la production collaborative de contenus devant s'apparenter à des « communs numériques », dans la mouvance des ressources éducatives libres. Dans l'enseignement supérieur sont évoqués de manière vague les besoins de formation des enseignants au numérique, suivant les résultats de la recherche, en écho aux principes de l'« *Evidence-Based Education* ».

Évaluation	**Cahier des charges du service de remédiation et d'accompagnement en seconde, contexte :** « Ce service doit offrir aux professeurs une assistance dans les choix d'activités et dans l'analyse des apprentissages des élèves. Il doit permettre une évaluation continue ainsi qu'un suivi régulier des apprentissages par les professeurs » (p.6) ; **« exigences fonctionnelles du service » : «** Cet outil doit permettre une utilisation à l'initiative des élèves et donc leur proposer un parcours personnalisé dynamique (évaluation formative) de remédiation, d'entraînement, de renforcement, voire d'approfondissement de notions ou de compétences » (p.9). « Ce service repose sur les principes de responsabilité et d'autonomie de l'élève, notamment à travers : […] la capacité	**Appel DemoES, objectifs :** « changement dans le processus d'évaluation des étudiants qui évolue vers l'évaluation des compétences (et moins des connaissances) et pour lutter contre les nouveaux types de fraudes (plagiat, fraudes numériques etc.) » (p.5) **Appel « Hybridation des formations de l'enseignement supérieur », projets attendus :** « Le développement de nouveaux modes d'évaluation »… » (p.6) ; « La création d'activités pédagogiques en ligne collaboratives (…ateliers distanciels d'évaluation par les pairs) (p.7)

[105] Le Code stipule : « La liberté pédagogique de l'enseignant s'exerce dans le respect des programmes et des instructions du ministre chargé de l'éducation nationale et dans le cadre du projet d'école ou d'établissement avec le conseil et sous le contrôle des membres des corps d'inspection » (Art. L. 912-1-1).

	à s'autoévaluer, à effectuer seul un bilan en fonction de jalons donnés (par mois, par trimestre...) » (p.14). **Stratégie « Choc des savoirs » :** « Utilisateur indépendant (B1-B2) : L'IA au service des approches métacognitives et de l'évaluation formative » (p.25).	

Les textes officiels soulèvent des pistes très variées de refonte des modes d'évaluation par la technologie. La commande d'un service de remédiation et d'accompagnement adaptatif pour les lycéens a été l'occasion d'évoquer des modes d'évaluation formative. Ne visant guère à sanctionner des acquis comme son homologue sommative, l'évaluation formative intervient, selon un bulletin de veille de Réseau Canopé, « dans le processus d'apprentissage avec pour but de sécuriser l'apprenant » :

> « Elle permet un regard singulier sur les étapes d'apprentissage de chacun et offre la possibilité d'adapter les suivis pédagogiques. Elle se centre sur l'élaboration de critères de référence. Ces critères sont partagés avec l'apprenant pour lui permettre de se situer dans les compétences à acquérir, de prendre conscience de ce qu'il a acquis et de ce vers quoi il se dirige » (Coudray, 2022).

Une telle évaluation de l'élève par des algorithmes sert autant l'enseignant en vue de superviser ses progrès que l'élève lui-même, qui doit à la fois profiter d'un parcours adaptatif, s'auto-évaluer et développer des compétences métacognitives. Les perspectives de l'IA générative en matière d'évaluations formatives commencent également à être rapportées – notamment la possibilité pour l'élève d'évaluer lui-même l'IA, comme l'explique la stratégie « Choc des savoirs ». Dans l'enseignement supérieur, l'appel « Hybridation des formations de l'enseignement supérieur » vise l'expérimentation de nouveaux modes d'évaluation et envisage des « ateliers distanciels d'évaluation par les pairs » (p.7). Aussi divers soient-ils, les nouveaux modes d'évaluation soulevés par les Ministères placent les apprenants en posture active. Néanmoins l'appel DemoES en revient à un mode d'évaluation sommative lorsqu'il appelle les démonstrateurs à tester un « changement dans le processus d'évaluation des étudiants » (p.5) pour passer de l'évaluation des connaissances à celle des compétences. Cette transition est motivée par une posture défensive face aux possibilités de fraudes accrues avec le numérique et *a fortiori* l'IA générative, rendant l'évaluation des connaissances inutile.

Gouvernance	**Appel e-FRAN, objectifs :** « Le numérique comme ouverture à de nouveaux acteurs et à de nouveaux savoir-faire : développement de contacts avec des entreprises du secteur numérique, des associations, des acteurs de la culture scientifique, technique et industrielle ou des laboratoires de recherche ; conduite d'actions pédagogiques par projet en partenariat avec ces acteurs… » (p.9) **Stratégie « Numérique pour l'éducation 2023-2024 », « Les enjeux » :** « Créer une impulsion pour renforcer la coopération des acteurs ; construire une ambition partagée et une	**Appel DemoES, objectifs :** « Transformer les métiers de l'enseignement supérieur par un projet d'établissement global visant les meilleurs standards internationaux (campus numérique : scolarité, ENT, accueil et attractivité nationale et internationale, diplomation et certification, etc.) ; […] développer et professionnaliser les fonctions d'appui en matière d'ingénierie pédagogique » (p. 6)

	gouvernance efficiente qui structure la coopération des acteurs nationaux et locaux autour de projets pédagogiques mobilisant le numérique là où il est pertinent » (p. 14).	

Le numérique est vu comme une opportunité d'ouverture de l'École sur les acteurs socio-économiques liés à ces technologies. Doivent coopérer acteurs nationaux et locaux, dans ce qui serait un mode de gouvernance collaborative et « par projet » d'inspiration néolibérale. Elle serait ainsi « efficiente » selon la Stratégie « Numérique pour l'éducation 2023-2027 », réinvestissant les préceptes du *Social Efficiency Movement.* Dans l'enseignement supérieur, la mise en place de ce type de technologies doit transformer les métiers, et notamment développer celui des ingénieurs pédagogiques. L'avènement de « campus numériques » doit s'inscrire dans un projet à l'échelle de l'établissement, lui permettant d'être compétitif dans un contexte de concurrence mondialisée – témoin la référence aux « meilleurs standards internationaux » dans l'appel DemoES.

Infrastructure	**Cahier des charges du service de remédiation et d'accompagnement en seconde, exigences éthiques :** « Éthique dès la conception (*Ethic by design*) et en exploitation : Les chercheurs, ingénieurs et entrepreneurs qui contribuent à la conception, au développement et à la mise à disposition de systèmes d'IA sont amenés à jouer un rôle décisif dans la société numérique. Il est essentiel qu'ils agissent de manière responsable, en prenant en considération les impacts sociaux de leurs activités et en tenant compte systématiquement du droit et des avis des utilisateurs enseignants et des équipes de formateurs et d'encadrement académiques et nationales pour développer des solutions au service de tous. Il est ainsi attendu que soit systématiquement adoptée une approche technologique centrée sur l'utilisateur et ses besoins (*user-centric approach to the technology*). […] Une vigilance permanente et des études d'impact doivent être portées sur la sécurité, la qualité et la nature des données traitées, les risques de biais de données, de discrimination (*discrimination impact assessment*), etc » (p.17). « Transparence : Il s'agit de développer « l'interprétabilité » des propositions, « éclairer les boîtes noires » pour renforcer le pouvoir des utilisateurs (*empower users*). […] Le déploiement de l'IA passe par une meilleure « interprétabilité » des résultats obtenus. Le niveau d'explication à fournir peut être variable suivant les publics mais des explications doivent être disponibles pour tous, y compris le grand public comme par exemple les parents d'élève » (p.18).	**Appel DemoES, objectifs :** « Concevoir des solutions avec les communautés du logiciel libre et de l'*EdTech*. Favoriser la recherche-développement et la co-construction d'une offre technologique de l'*EdTech* et des communautés du logiciel libre adaptées aux problématiques de l'enseignement supérieur » (p.6) ; **critères de sélection :** l'utilisation et/ou l'intégration des solutions et ressources déjà proposées sur des plateformes existantes et/ou présentant des garanties en termes de souveraineté et de confidentialité » (p.9)

	Stratégie « Numérique pour l'éducation 2023-2024 », « Les enjeux » : « Accompagner et renforcer la sécurisation, la résilience et la fiabilité des données et du système d'information ministériel, au service des utilisateurs ; garantir l'interopérabilité des outils et des ressources, développer une approche utilisateur des produits et promouvoir un numérique écoresponsable » (p.14).	

Le développement des technologies doit se faire en lien avec les acteurs privés de l'*EdTech* et les communautés du logiciel libre, pour prendre en compte les besoins spécifiques des acteurs de terrain et leur mettre à disposition gratuitement ces outils – témoin l'appel DemoES. L'infrastructure technologique est néanmoins essentiellement décrite à travers un prisme régulateur. La commande d'un outil de remédiation et d'accompagnement au lycée donne à voir un agenda ministériel dès la conception de ces technologies : la prise de conscience de l'influence des ingénieurs sur la société et l'éducation pousse le Ministère à leur imposer une « vigilance » sur plusieurs aspects : la sécurité et le risque de discrimination émanant des données, mais aussi la « transparence » des décisions algorithmiques, qui doivent être explicables auprès de chaque type d'utilisateur. La fréquence des expressions anglaises (« *Ethic by design* », « *user-centric approach to the technology », « user empowers »)* montre l'influence des codes ingénieuriques états-uniens. Les autres textes ajoutent d'autres points d'attention : la souveraineté et l'interopérabilité des outils, ou l'« éco-responsabilité » eu égard à l'impact environnemental des technologies numériques – une spécificité par rapport au cas états-unien.

On retrouve là la sensibilité éthique très forte des pouvoirs publics français comme européens. Le respect de la confidentialité des données est devenu un enjeu particulièrement fort après le vote du Règlement Général de Protection des Données (RGPD) en 2016 à l'échelle de l'Union. Entré en application en France en 2018, il oblige les sites à informer ses visiteurs des finalités de collecte de leurs données et à s'assurer de leur consentement. Qui plus est, la Commission européenne a formulé à partir de 2018 son ambition de devenir une puissance pionnière de l'« IA éthique », comme en a témoigné le vote de l'*AI Act* en 2024, qui encadre le développement de l'intelligence artificielle et bannit les systèmes les plus néfastes selon une typologie de risques sociétaux liés à l'IA.

Cependant, ces directives peuvent sonner comme des vœux pieux voire des postures, à l'instar de certaines exigences posées à l'outil de remédiation et d'accompagnement des lycéens de seconde. Devant la complexité de l'explicabilité à l'heure du *deep learning* et l'idéalisme du concept de transparence, cet agenda éthique apparaît peu crédible, d'autant que le Ministère n'a toujours pas publié le code source des algorithmes de MIA Seconde en octobre 2024 (Clavey, 2024).

Sont ainsi posées les briques d'un projet formé par couches successives, au gré des appels à projets et des directives ministérielles. Si sa spécificité vis-à-vis de la forme scolaire est patente, il reprend de nombreuses caractéristiques du *personalized learning* américain, tout en multipliant au fil du temps des pistes pédagogiques qui affectent sa cohérence. Parmi ses aspects réformateurs, on peut noter la focalisation sur l'apprenant, en particulier l'élève en difficulté dans l'ancien système, dont les besoins seraient enfin reconnus. À l'idéal culturel de l'« acculturation numérique » des citoyens s'ajoute un objectif utilitariste d'insertion professionnelle proche de l'approche états-unienne. De même, la posture active de l'élève est encouragée dans son utilisation des outils numériques suivant un agenda constructiviste et collaboratif.

En outre, la « technologisation additive » est promue du côté de l'enseignant, c'est-à-dire une implémentation de la technologie limitée à un rôle d'assistance. La figure de l'enseignant fait cependant l'objet de davantage de garanties, tenant sans doute à la « liberté pédagogique » accrue des enseignants français par rapport à leurs homologues américains. Demeure dans les deux pays un même cadre industrialiste inspiré du *Social Efficiency Movement* et de l'*Evidence-Based Education*, conforme au « nouvel esprit du capitalisme » théorisé par Luc Boltanski et Ève Chiapello autour de valeurs comme la collaboration, l'innovation et le changement (Boltanski et Chiapello, 1997). La technologie adaptative doit rendre plus efficace l'enseignement de même que les investissements publics. L'impératif apparaît cependant moins contraignant en France pour les enseignants et directeurs d'établissements, non-concernés par l'*accountability*.

Du côté de la gouvernance, l'ouverture des établissements sur les acteurs socio-économiques externes est encouragée, notamment dans leurs territoires. Autre levier d'industrialisation, la professionnalisation est encouragée à travers le rôle renforcé des ingénieurs pédagogiques. Le projet des Ministères insiste enfin sur les garde-fous éthiques à opposer aux techniques d'IA pour préserver les libertés individuelles, de façon plus marquée qu'outre-Atlantique. Ils développent en revanche moins de doutes sur les limites techniques et pédagogiques de l'IA que le pouvoir fédéral américain appelant à mettre « l'humain dans la boucle ».

Le projet apparaît encore élusif sur de nombreux aspects, *a fortiori* dans l'enseignement supérieur où les discours sont moins nombreux. Plusieurs options pédagogiques sont, au mieux, décrites sans qu'un choix clair soit opéré. Par exemple dans l'apprentissage, jusqu'où la collaboration entre les élèves doit-elle aller, au sein et en dehors des outils numériques ? Selon quelles modalités les enseignants doivent-ils se référer et contribuer aux recherches en sciences de l'éducation ? Concernant l'évaluation, quel est l'avenir des modalités sommatives traditionnelles dans l'enseignement primaire et secondaire ?

La vocation de ce projet techno-éducatif pose enfin question. Les Ministères comptent-ils ainsi renoncer à la forme scolaire établie et installer un nouveau paradigme, ou seulement l'amender selon les nécessités de chaque échelon éducatif et de chaque territoire ? Cette interrogation nécessite de nuancer une vision unifiant excessivement le système scolaire et universitaire français autour des volontés ministérielles. Au contraire, deux pierres d'achoppement demeurent sur le chemin du renouvellement paradigmatique : d'une part la faible marge de manœuvre financière dont disposent les établissements publics pour transformer leurs modes d'enseignement et de gouvernance, et d'autre part les doutes suscités par l'instrumentalisation de la recherche scientifique au sein de l'administration.

Des obstacles financiers et scientifiques

Un sous-investissement par rapport aux ambitions de refonte

Le premier obstacle, les budgets restreints de bon nombre d'établissements, crée une dissonance avec les ambitions fortes affichées par les Ministères. Celles-ci consistent, du côté des politiques de l'offre, à développer des innovations « disruptives », et, du côté de celles de la demande, à « réinventer » les pratiques pédagogiques. Pourtant les sommes investies par l'État sont éloignées de celles annoncées initialement et faibles par rapport aux budgets ministériels.

Ce hiatus transparaît notamment à travers le plan « Numérique pour l'éducation » de 2015, dans lequel l'appel e-FRAN était inclus. L'annonce d'un investissement inédit – un milliard d'euros sur trois ans – a été dénoncée en 2019 par le rapport de la Cour des comptes évoqué plus haut : « les moyens mis en place par l'État en faveur de la politique du numérique éducatif s'élèvent à environ 10 M€ par an sur le budget du ministère, auxquels il convient d'ajouter les 250 M€ de dépenses du PIA sur la période, soit environ 300 M€ au total de 2013 à 2017, en regard du milliard d'euros initialement annoncé au titre du plan numérique pour l'éducation » (Cour des Comptes, 2019). L'appel e-FRAN prévoyait quant à lui trente millions d'euros. Au final, ce sont 19,5 millions d'euros qui ont été accordés aux lauréats. Une somme largement rognée, et minime par rapport aux treize milliards d'euros de budget de l'enseignement supérieur en 2016. Au tableau d'un sous-investissement dans le numérique depuis longtemps brossé par les chercheurs (Mœglin, 2010, p.111), l'intelligence artificielle n'a jusqu'ici apporté que peu de nouveautés.

Dans l'enseignement supérieur, les budgets serrés des universités font en outre de l'ingénierie pédagogique et de l'achat d'outils technologiques des postes de dépenses secondaires. Les appels à projet ministériels fournissent aux quelques établissements lauréats des moyens importants pour co-construire de nouveaux outils. Néanmoins la faible marge budgétaire des universités les empêche d'acheter par elles-mêmes des technologies existantes et éprouvées par la recherche. Sur ce terrain, un écart semble se creuser avec les écoles privées, d'ingénieurs ou de commerce, qui font de la personnalisation de leurs cursus des chantiers stratégiques et des arguments d'appels auprès des étudiants. Nombreux sont les établissements à se doter ainsi de services algorithmiques analysant les données variées des élèves afin de leur recommander des parcours de formations « sur-mesure ». Certaines écoles de commerce vont jusqu'à mettre en place des incubateurs pour les start-ups *EdTech*, telles que l'EM Lyon ou Neoma Business School.

Les désaccords internes sur la recherche d'efficacité

Le second obstacle au changement de paradigme éducatif provient des dissensions entre les directives ministérielles et les textes émanant de chercheurs comme de hauts fonctionnaires sur la question de l'efficacité pédagogique du numérique. Les Ministères tendent à recourir aux sciences en vue d'obtenir des cautions métriques pour asseoir la légitimité de réformes éducatives. Les discours des décideurs politiques évoqués plus haut ont ainsi fixé les mesures d'efficacité, en termes de performances des élèves, comme une priorité. L'interdisciplinarité dans les protocoles de recherche, devant permettre d'appréhender les divers facteurs pesant sur la réussite éducative, est fréquemment brandie comme un moyen pour y parvenir. Lors de l'inauguration du CSEN, le ministre Jean-Michel Blanquer en a d'ailleurs fait un trait saillant de l'organe :

> « C'est un moment important, parce que d'affirmation extrêmement forte sur l'importance d'avancer en matière d'éducation, à la lumière des sciences. […] Quand on dit : les sciences, il y a évidemment toute une série de réflexions et de remarques qui viennent immanquablement à l'esprit ; d'abord, sur la nécessaire diversité et l'interdisciplinarité des sciences elles-mêmes, quand il s'agit de comprendre la complexité de l'Homme, du monde et de la société. […]. Il est évident que, sur l'éducation, se penchent une bonne partie des sciences dites « dures » et des sciences sociales et humaines. Et nous avons besoin de la complémentarité des sciences. La composition même de ce Conseil l'indique ». (Conseil Scientifique de l'Éducation nationale, 2019, p.11).

La promesse d'interdisciplinarité est cependant loin d'être tenue dans le temps, les initiatives ministérielles penchant surtout vers les sciences informatiques et des professionnels issus du monde

socio-économique. Dans la présentation de ses activités pour l'année 2019-2020 est annoncée la composition du groupe de travail « Intelligence artificielle et numérique pour l'éducation » : treize membres au total, dont six membres du CSEN et sept chercheurs invités (Conseil Scientifique de l'Éducation nationale, 2019, p.59). Les disciplines représentées étaient, par ordre décroissant de représentation, l'informatique, avec quatre chercheurs (Gérard Berry, Francis Bach, Anne Boyer, Vanda Luengo), les sciences politiques, avec deux chercheurs (Marc Gurgand, Patrick Savidan), puis l'économie, les SIC, les sciences éducatives et la neuropsychologie, avec un chercheur pour chaque discipline (respectivement Julien Grenet, Jean-François Cerisier, Pascal Bressoux et Stanislas Dehaene).

La publication finale du groupe, trois ans plus tard, révèle cependant un chamboulement du comité. Il ne compte plus que deux chercheurs parmi ceux annoncés initialement (Anne Boyer et Jean-François Cerisier). En outre, elle fait état d'une minoration paradoxale des sciences, *a fortiori* humaines. Sur les onze membres permanents, la majorité sont des professionnels issus de l'administration scolaire – deux de la DNE (Axel Jean, Alain Thillay) et trois de Réseau Canopé (Jean-Michel Perron, Melina Solari Landa, Romain Vanoudheusden). En outre, trois membres sont entrepreneurs ou ingénieurs (Anis Ayari, Aurélie Jean, Victor Storchan). Seuls trois chercheurs en activités sont présents, dont deux en informatique (Anne Boyer, Jill-Jênn Vie) et un en SIC (Jean-François Cerisier).

Dans l'appel des « groupes thématiques numériques » de 2020, la consigne est aussi de favoriser « une approche interdisciplinaire et systémique du numérique dans l'éducation » (p.2). Les deux lauréats de l'axe consacré à l'intelligence artificielle, les GTnum 8 et 9, ont très variablement tenu la promesse d'interdisciplinarité. L'annexe 3 synthétise les caractéristiques de ces deux groupes. Pour le GTnum 8 a été retenu le projet coordonné par le laboratoire LS2N de l'Université de Nantes et le LabSTICC de l'IMT-Atlantique Brest, deux unités de recherche en informatique. La discipline est de fait surreprésentée dans le groupe de travail, les trois chercheurs le coordonnant étant affiliés à cette discipline. Sa démarche instrumentale paraît cependant insuffisante pour traiter les questions de recherche annoncées dans le portfolio du GTnum, liées à l'utilité éducative des ressources éducatives libres (REL) et de l'IA :

> « Est-il raisonnable, utile, de développer une politique ouverte en matière de contenus pédagogiques ? L'IA peut-elle aider à faciliter la navigation dans l'écosystème des REL ? à mieux trouver une ressource adaptée à un besoin particulier ? Inversement, peut-on proposer des REL pour apprendre l'IA ? et, plus particulièrement, pour aider les enseignants ? » (GTnum LS2N #IA_EO, 2023b).

Le GTnum 9 concilie davantage les sciences humaines et informatiques. Le comité de pilotage et les porteurs des cinq équipes de travail est interdisciplinaire : les sciences de l'information et de la communication comptent trois chercheurs, de même que les sciences de l'éducation et de la formation, alors que l'informatique en compte deux. Le groupe reconnaît par ailleurs l'intérêt relatif et incomplet des mesures d'efficacité : « La conception pédagogique, susceptible de provoquer une bonne régulation des sociotechniques, doit être moins pensée en termes d'efficacité des apprentissages […] qu'en ceux d'émancipation (versus aliénation) ». Il obéit toutefois à l'objectif ministériel de diffusion de « bonnes pratiques » via « l'identification et le partage des pratiques et des usages créatifs du numérique pour une éducation adaptée aux enjeux des sociétés contemporaines liés à la diffusion des outils comportant des techniques d'IA » selon son portfolio (GTnum LINE #Scol_IA, 2023). Les questionnements psychologiques qu'attendait l'appel sont finalement peu traités, ce qui se vérifie dans l'absence de chercheurs en psychologie dans le groupe.

Selon la proposition de recherche soumise par le collectif en septembre 2020 (Romero et al., 2020), les trois chercheurs ayant des spécialisations liées à la psychologie sont tous trois des chercheurs en informatique[106].

C'est en définitive une démarche instrumentale, à dominante informatique, qui est la plus mobilisée par le Ministère. Elle sert l'objectif d'efficacité, mesuré par les performances des élèves, et mobilisé pour légitimer des mesures comme la généralisation de MIA Seconde au lycée. Dans le compte rendu de la mission « Choc des savoirs », le Ministère met ainsi en avant le fait que « certains logiciels et outils numériques, incluant ou non de l'intelligence artificielle, ont fait la preuve de leur efficacité dans des situations pédagogiques spécifiques, en complément du cours délivré par les professeurs » (Ministère de l'Éducation nationale, de la Jeunesse et des Sports, 2023c, p.25). Cette lecture vague et instrumentale de la recherche est pourtant remise en cause au sein même de l'administration, fragilisant la légitimité du soutien à l'IA.

La recherche de corrélations entre l'usage des outils numériques et les performances des élèves a notamment été remise en cause dans le rapport de l'inspectrice générale de l'éducation nationale Catherine Becchetti-Bizot intitulé « Repenser la forme scolaire à l'heure du numérique. Vers de nouvelles manières d'apprendre et d'enseigner », remis en 2017 au Ministre de l'Éducation nationale. S'appuyant sur les acquis de la recherche en sciences de l'éducation, Catherine Becchetti-Bizot soutient que la question de l'apport des technologies, prise sous l'angle des performances scolaires des élèves, est une gageure à cause de l'aspect multifactoriel de l'apprentissage. « La question dès lors est de savoir quelles stratégies doivent être mises en place, mais aussi quels contextes (au sens de milieu) il faut créer pour permettre les apprentissages et obtenir les progrès souhaités » (Becchetti-Bizot, 2017, p.8). Des considérations supposant à nouveau une interdisciplinarité dans les protocoles de recherche.

Néanmoins l'instrumentalisation politique de la recherche suscite des remous jusque dans l'organe ministériel du CSEN. Ainsi, en mars 2024, peu après l'annonce des résultats de la mission « Choc des savoirs » dont Stanislas Dehaene était coordinateur, trois économistes membres du CSEN démissionnent : Élise Huillery, Yann Algan et Julien Grenet. Chacun remet à l'époque en cause l'impartialité du Conseil et l'absence de consultations suite aux mesures annoncées par le ministère, dont le renforcement des redoublements et la généralisation des groupes de niveaux. Ainsi Élise Huillery, interrogée par le Café pédagogique (Gani, 2024), regrette l'absence de consultations du CSEN en amont : « une consultation scientifique sur les groupes de niveaux par exemple aurait permis de rendre compte "des résultats de la littérature scientifique en termes de composition sociale, pour voir à quel point les groupes de niveaux recoupent les origines sociales des élèves et recréent de la ségrégation à l'intérieur des collèges" ». Julien Grenet, ancien membre annoncé du groupe de travail sur l'intelligence artificielle, est également cité comme ayant « le sentiment que le CSEN a été instrumentalisé comme caution scientifique et […] dénonce des mesures qui vont à l'encontre des résultats de la recherche » (ibid.).

Souffrant déjà de plusieurs apories, le projet techno-éducatif mis en exergue plus haut s'inscrit donc dans une politique ministérielle d'« *Evidence-Based Education »* remise en cause par de nombreux chercheurs, jusqu'à ceux participant à ses organes consultatifs. Malgré la centralité forte

[106] Solange Cartaut pour « Psychologie de l'activité d'apprentissage et de développement », Pierre Crescenzo pour « Psychologie des êtres artificiels » et Anaïs Leroy pour « Psychologie cognitive ».

des Ministères dans le système éducatif français, les conditions ne semblent pas réunies pour que des usages conformes au projet officiel émanent des organisations éducatives.

Néanmoins, l'intégration des médias adaptatifs dans les salles de classe et les amphithéâtres tient aussi à la manière dont les acteurs économiques les industrialisent. Cette question nécessite de s'intéresser aux points de vue pluriels des acteurs de l'offre. Que ce soient des petites structures de programmation *EdTech*, des éditeurs scolaires concentrés ou des entreprises anciennes du numérique éducatif, quelles sont leurs contributions à l'impulsion d'un projet techno-éducatif ? S'accorderaient-ils autour des grandes lignes du projet ministériel, de son approche additive et ouverte de la technologie ? S'en détourneraient-ils au contraire ? Comme dans le cas états-unien, cette question nécessite d'étudier l'évolution des rapports de force socio-économiques entre les firmes, pour identifier les acteurs en mesure d'imposer leurs visions.

2. Les stratégies des firmes

Il est temps de reposer, à l'aune du cas français, l'hypothèse que nous avions formulée en préambule de cette partie, selon laquelle les médias adaptatifs refaçonneraient les positions des acteurs dans l'ensemble des industries éducatives, contribuant à leur fédération au sein d'une socio-économie dominée par les acteurs des technologies. Après avoir observé le volontarisme et les apories des politiques publiques françaises, il convient maintenant d'analyser les stratégies nouvelles des firmes. Comme dans le cas états-unien, nous adoptons un point de vue transversal aux cinq industries éducatives en considérant une filière intégrée des médias adaptatifs, organisée autour des trois étapes génératrices de valeur : en amont, la conception d'algorithmes et la production de contenus, puis en aval, la distribution auprès d'un public éducatif.

Le développement des algorithmes d'apprentissage adaptatif

La conception d'algorithmes d'apprentissage adaptatif, cœur de métier des start-ups

Le développement d'algorithmes est en France aussi un chantier principalement assumé par de jeunes structures de production informatique de petite taille, les start-ups. Tirant de leurs algorithmes leur proposition de valeur dans la filière, elles investissent l'essentiel de leurs ressources dans leur R&D, et en particulier leurs algorithmes d'apprentissage adaptatif, pour bâtir et consolider un avantage concurrentiel. Pour soutenir ce poste de dépense, elles doivent composer avec un environnement de financement différent de celui des États-Unis. De celui-ci découlent des inégalités importantes entre les structures, des stratégies d'éligibilité et des risques spécifiques.

Un rapport d'EY-Parthenon de 2022 offre un aperçu général de ces firmes. Il indiquait qu'il y avait environ 500 « *EdTech »* en France, et que la majorité d'entre elles étaient de petites entreprises qualifiables de start-ups : « 60 % des *Edtech* ont moins de 10 salariés et 60 % […] déclarent un chiffre d'affaires annuel inférieur à 500 K€ » (ibid., p.8). Par ailleurs, 54% d'entre elles auraient été créées entre 2017 et 2021 (ibid.). En découle un paysage encore peu mature, où les inégalités sont importantes entre les vingt plus grandes start-ups réalisant 70% du chiffre d'affaires du secteur, et les centaines d'entreprises restantes (ibid., p.4). Alors que les « taux de mortalité » sont très élevés, notamment lors de la première année, les inégalités entre start-ups sont importantes pour l'accès aux principaux guichets de financements.

Le premier d'entre eux est représenté par l'État et les collectivités territoriales. Selon un rapport d'EY-Parthenon de 2022, « la filière *Edtech* française a bénéficié d'un soutien public important : 45 % des startups ont perçu des subventions publiques au cours de leur développement » (ibid.). On peut distinguer deux types de soutien. Le premier concerne les dispositifs fiscaux allégeant les charges des entreprises, comme les crédits d'impôt recherche. Comme l'a montré Hugo Carcaly pour les start-ups du numérique et de la création, ce type de soutien est sollicité par un grand nombre de jeunes entreprises. Toutefois le surinvestissement dans la R&D et la dépendance aux dispositifs de financements de l'innovation expose les petites entreprises à des effets pervers :

> « Ces entreprises concentrent l'essentiel de leurs investissements dans le développement de briques technologiques, très onéreuses et sans aucune garantie de parvenir à générer des revenus grâce à ce travail. […] Or, le risque de concentrer les investissements dans des projets de R&D est de créer une forme de dépendance à l'égard des dispositifs de soutien à l'innovation, mais également de sans cesse devoir renouveler la technologie au risque de subir la concurrence d'acteurs mieux dotés et aux stratégies de conquête des marchés très agressives » (Carcaly, 2022, p.258).

Les start-ups risquent ainsi de poursuivre leurs investissements à perte dans leur R&D pour continuer à percevoir ces aides publiques à l'innovation, quand de grands groupes, notamment les *Big Tech*, créeraient une menace d'obsolescence.

Pour accélérer le développement des start-ups, le second type de soutien public à même de créer un saut dans leurs ressources financières concerne les financements ponctuels, tels que les prêts ou les apports de capitaux en fonds propres de BPI France, les programmes de subventions et les marchés publics lancés par l'État et les collectivités territoriales. Ces derniers permettent en outre aux start-ups lauréates d'accéder à un marché plus ou moins vaste et d'y entraîner leurs algorithmes.

Dès lors des stratégies d'éligibilité, visant à répondre aux attentes de ces divers acteurs publics et attirer ces différents types de soutien, sont observables *a posteriori* dans la trajectoire des startups « à succès ». L'une d'entre elles est le recours à l'étiquette de l'intelligence artificielle pour présenter leurs activités. Lalilo en offre un exemple éloquent. Si les trois fondateurs ont d'abord pris la direction de San Francisco en 2016 pour suivre un programme d'incubation, l'appel P2IA publié par le Ministère de l'Éducation nationale l'année suivante leur a ouvert les portes du marché français (Encadré 4, p.124). Avec cette condition lexicale de se présenter comme un acteur de l'intelligence artificielle, ainsi que nous l'a confié Laurent Jolie : « on ne tient pas particulièrement à ce vocabulaire, c'est le Ministère qui y tient » (Entretien 4).

Un autre lauréat du P2IA, Evidence B, basait quant à lui sa plateforme sur l'idée d'un lien constant avec la communauté académique, notamment en sciences cognitives. Si ce positionnement est régulièrement justifié par l'entreprise pour son apport pédagogique, l'utilité stratégique d'une telle communication, au regard des appels à projets gouvernementaux, est palpable tout au long de son histoire (Encadré 7). Cette proximité avec les chercheurs offre en outre à l'entreprise des relais d'influence valorisables auprès du Ministère. Ainsi Emmanuel Sander, professeur de psychologie du développement et de l'éducation et œuvrant à la conception des exercices au sein de la plateforme, est-il membre du Conseil Scientifique de l'Éducation Nationale (CSEN). Le président de cet organe, Stanislas Dehaene, était par ailleurs coordinateur de la Mission « Exigence des savoirs » à la fin de l'année 2023, à l'origine de la généralisation de l'outil d'Evidence B, MIA Seconde, dans tous les lycées de France.

Encadré 7. Evidence B, une croissance grâce aux appels à projets et à son réseau en France et à l'international

Evidence B, éditeur d'ITS axés sur les « savoirs fondamentaux » en français et mathématiques, a été créé en 2017 par Thierry et Catherine de Vulpillières. Le premier est ancien directeur du numérique éducatif chez Editis et ex-responsable du programme *Partners in Learning* de Microsoft, quand la seconde est ancienne enseignante en langues anciennes en classe préparatoire. Ils s'associent rapidement à Didier Plasse, CEO de Kidscode et « serial entrepreneur » impliqué par le passé dans des entreprises du numérique (ImmoStreet). Ils réalisent une première levée de fonds de 2 millions d'euros à la fin de l'année 2019 auprès des fonds Finorpa, Nord-France Amorçage et IRD. L'entreprise se rémunère d'abord en vendant des modules adaptatifs intégrables aux LMS d'éditeurs scolaires. Ce modèle permettra à l'entreprise de remporter le P2IA en 2019, de gagner un marché avec Hatier ainsi qu'un appel « Banques de Ressources pour le Numérique pour l'Education » (BRNE) en latin. À partir du Covid et de sa première levée de fonds, Evidence B développe sa propre plateforme qu'elle vend à des éditeurs. Elle recrute alors ses premiers développeurs et intègre le programme d'accélération londonien Educate. Son chiffre d'affaires, d'environ 240 000 euros en 2019 selon Infogreffe, triple alors chaque année selon les dires de Thierry de Vulpillères en 2021 (Langlois, 2021). La firme s'internationalise notamment grâce à l'arrivée en 2020 de Philippe Méro en tant que vice-président « Développement et Partenariats Stratégiques ». Elle noue un contrat avec Pearson Italie pour la livraison d'un ITS en latin, ou encore un en mathématiques avec la Rosyth Gifted School de Singapour. À l'été 2022 elle remporte avec *Adaptiv'Math* un concours d'éditeurs *EdTech* aux États-Unis, le *Learning Tools Competition*, qui lui ouvre les portes des écoles du School District de New York pour y développer son ITS en anglais. En 2022 la firme est sélectionnée par le Ministère de l'Education nationale pour développer un outil de remédiation et d'accompagnement des élèves de seconde en français et en mathématiques, en consortium avec Docaposte. Cet outil deviendra MIA Seconde, comme on l'a vu plus haut. Evidence B réalise enfin une levée de fonds de 4 millions d'euros auprès d'Educapital, Finorpa et d'IRD Gestion en novembre 2022.

Néanmoins les financements et subventions émanant des acteurs publics constituent une manne complexe à atteindre pour un très grand nombre d'*EdTech*. Des lenteurs administratives sont en outre pointées du doigt par de nombreux entrepreneurs. Même pour un programme national comme le P2IA et pour une entreprise « à succès » comme Evidence B revient ce reproche, formulé par son vice-président « Développement et Partenariats stratégiques », Philippe Méro :

> « Les P2IA sont des dispositifs très intéressants pour la filière, qui favorisent le développement de solutions françaises innovantes, même si on peut aussi regretter les délais pour le lancement des appels d'offres, l'analyse des réponses, la notification et le lancement effectif des marchés » (Livingstone-Wallace et Welter, 2024, p.97).

Un autre guichet dont l'accessibilité génère également des inégalités entre les start-ups est représenté par les fonds de capital-risque. Les levées représentent une manne importante mais bien moins développée en France qu'aux États-Unis, comme on l'a dit plus haut. En 2021, année record pour les levées de fonds mondiales, celles comptabilisées par les start-ups françaises ont atteint 487 millions d'euros, quand leurs homologues états-uniennes ont rassemblé 8,2 milliards de dollars, soit environ 6,95 milliards d'euros (Figure 12). Plusieurs facteurs expliquent cette différence. D'abord les marchés éducatifs sont globalement plus décentralisés et libéralisés aux États-Unis, créant un environnement favorable à la marchandisation et la croissance des entreprises. En outre la financiarisation de l'éducation y est plus développée : d'une part, le secteur du *venture capital* est animé par une multiplicité des fonds d'investissement spécialisés dans les industries éducatives aux États-Unis, alors que la France ne dispose que de quelques représentants, tels qu'Educapital ou Bright Eye Ventures. D'autre part, les financements de fondations philanthropiques y sont nombreux contrairement à la France.

Les principales levées sont cependant réalisées dans la formation professionnelle, où l'individualisation du marché via le Compte Personnel de Formation a créé un environnement favorable à la croissance d'entreprises *EdTech*. Celles positionnées sur la formation professionnelle

réalisent en effet 85% du chiffre d'affaires de la filière en 2021 selon le rapport d'EY-Parthenon (2022, p.8). Le même rapport confirme que « les levées importantes restent concentrées sur les champions de la filière et sur la formation professionnelle » et souligne par contraste que « les *Edtech* avec un chiffre d'affaires inférieur à 500 K€ tendent à manquer de financements en amorçage et en accélération, notamment sur le segment scolaire et le segment de l'enseignement supérieur » (ibid. p.3). À l'inverse des États-Unis où le *K-12* est le marché le plus développé, la centralité des pouvoirs publics et le faible nombre d'acheteurs affectent les levées des acteurs hexagonaux (Perreau, 2021).

Une stratégie d'éligibilité des start-ups auprès des fonds de *venture capital* est, en France comme aux États-Unis, le recours à l'intelligence artificielle comme argument marketing. Sur ce courant technologique converge l'horizon d'attente des investisseurs privés comme des pouvoirs publics. Tous y placent des enjeux économiques, financiers et politiques sur lesquels les start-ups capitalisent depuis plusieurs années. En découlent des dérives, certaines entreprises usant du syntagme alors que leurs outils et médias informatiques présentent des formes très relatives, voire inexistantes, d'intelligence artificielle, comme on l'a vu aux États-Unis.

Toutefois les levées sont soumises à une instabilité similaire à celles des États-Unis. L'histogramme en Figure 12 restitue la chute des levées de fonds en 2022, à la fin de la pandémie de Covid-19 et suite à la hausse des taux d'intérêt. Aux États-Unis, elles ont fondu d'environ 36,6% à 5,2 milliards en 2022, puis de près de 46,2% à 2,8 milliards en 2023 (Wan, 2024). En France, la chute a été plus vertigineuse encore : de 487 millions d'euros en 2021, les levées ne totalisaient plus que 93 millions en 2022, soit une baisse de près de 81%. Elles ont connu ensuite une légère embellie en 2023, en atteignant 162 millions d'euros, soit une hausse de près de 74,2% (BrightEye Ventures, 2024).

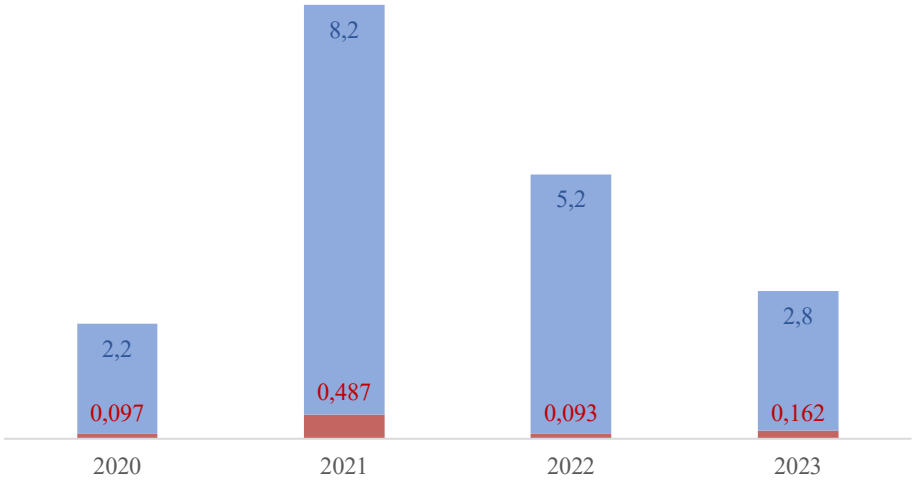


Figure 12. Levées de fonds totales réalisées par les start-ups EdTech *aux États-Unis (en bleu) et en France (en rouge), en milliards de dollars.*

Le développement d'algorithmes dans les premières années des entreprises est donc un impératif tant stratégique, en vue de développer et conserver un avantage concurrentiel, qu'économique, visant des financements publics comme privés.

La R&D s'opère de trois manières chez les start-ups. La première tient à des expertises en interne, notamment lorsque les créateurs se trouvent être des ingénieurs de formation. C'est le cas de Lalilo,

dont les fondateurs ont été formés à l'École Polytechnique. Un autre moyen, éventuellement complémentaire, est le recrutement d'ingénieurs hautement qualifiés. Cela entraîne des dépenses particulièrement importantes pour des start-ups, nécessitant souvent en amont le recours à des levées de fonds et la sollicitation d'aides publiques (comme les crédits d'impôt recherche).

Une deuxième manière de développer des algorithmes d'apprentissage adaptatif est la constitution de partenariats avec des laboratoires d'informatique, de sciences cognitives ou de psychologie cognitive. La forme la plus directe est le transfert de technologies depuis le laboratoire jusqu'à l'ITS. Evidence B a ainsi travaillé étroitement avec l'équipe Flowers de l'Inria Bordeaux, et notamment son directeur Pierre-Yves Oudeyer en tant que conseiller scientifique, pour incorporer l'algorithme *Zone of Proximal Development and Empirical Success* (ZPDES). L'entreprise a recruté ensuite un doctorant de Pierre-Yves Oudeyer, Benjamin Clément, comme *senior data scientist*, pour superviser son intelligence artificielle.

Domoscio, entreprise spécialisée dans l'enseignement supérieur et la formation professionnelle, a été créé en 2013, suite au transfert d'un algorithme issu des recherches de son co-fondateur au sein d'un laboratoire de l'école d'ingénieurs Grenoble INP-Ensimag (Encadré 8).

Encadré 8. Domoscio, une start-up de l'*adaptive learning* dans la formation professionnelle

Domoscio a été créée en 2013 par Benoit Praly et Ivan Ostrowicz, tous deux diplômés de Grenoble INP-Ensimag, école d'ingénieurs en informatique et mathématiques appliquées. Elle découle d'un transfert de technologie d'un laboratoire de Grenoble INP où Benoit Praly a contribué au développement d'un algorithme d'apprentissage adaptatif en 2012. Remportant des programmes de subventions pour soutenir et accélérer sa R&D, la firme commence en 2014 le déploiement de deux médias spécialisés dans l'enseignement supérieur et la formation continue : *Domoscio Hub*, ITS destiné à l'apprentissage via des parcours personnalisés, et *Domoscio Lock*, outil de révision consistant en un système de répétition espacée. Elle s'engage ainsi, en particulier avec *Domoscio Hub*, de manière précoce sur le marché de l'*adaptive learning*, en prétendant mêler les sciences cognitives, l'intelligence artificielle et le *Big Data*. Elle les vend sous formes de licences et remporte progressivement des contrats avec des établissements d'enseignement supérieur (Université Paris-Est Créteil, ESCP Business School, Junia HEI), des grands comptes (Ag2r La Mondiale, Enedis, SNCF) et des centres de formation (Action on Line). Entre 2014 et 2019, elle bénéficie du crédit impôt recherche et de subventions publiques pour environ 800 000 euros, selon le profil LinkedIn de Benoit Praly. L'entreprise lève notamment 250 000 euros auprès du fonds d'investissement belge Inventures en 2015. Domoscio remporte ensuite l'appel P2IA du Ministère de l'Éducation nationale en 2019, en consortium avec Hachette Education et Beneylu, pour concevoir Navi, outil d'apprentissage de la lecture. La firme atteint son premier million d'euros de chiffre d'affaires en 2019, toujours selon Benoit Praly. Annoncé en juin 2022, le rachat de Domoscio par l'éditeur belge Myskillcamp échoue en raison des difficultés financières de ce dernier. La firme a finalement été rachetée en avril 2024 par Rise up, autre ancienne start-up française spécialisée dans la formation professionnelle (Guerrier, 2024).

Une autre forme moins directe est le recrutement de doctorants en contrat CIFRE pour conduire des recherches appliquées sur leurs ITS. Lalilo a ainsi recruté Thomas Sergent, doctorant au LIP-6, laboratoire d'informatique de Sorbonne Université, sous la direction de Thibault Carron et François Bouchet, en vue d'investiguer l'autorégulation de l'apprentissage chez les jeunes utilisateurs de l'ITS.

Le développement de l'adaptativité peut s'opérer enfin par le recours à des sous-traitants. En particulier les *Big Tech*, qui sont, comme aux États-Unis, des acteurs centraux de la maintenance des services de start-ups françaises via leurs offres *cloud.* Lalilo a ainsi eu recours à des modules de divers acteurs des technologies pour son architecture adaptative. Parmi les applications tierces rapportées dans sa politique de confidentialité, l'on peut citer AWS comme prestataire *cloud,* ou encore de plus petites entreprises spécialisées dans l'analyse de données. Parmi les prestations

les plus cruciales pour l'adaptativité, l'on peut citer Mixpanel, qui contribue à l'analyse des comportements des utilisateurs via des traitements statistiques et des visualisations de données, ou encore Fullstory, qui permet l'enregistrement des sessions des apprenants pour une visualisation de leurs parcours en vue de contrôler la qualité des nouvelles fonctionnalités (Lalilo, 2024, p.6).

Les acteurs natifs du numérique au défi de la réinvention

On a vu qu'aux États-Unis les industries du numérique éducatif étaient animées de longue date par des tendances oligopolistiques, de grands groupes concentrés surplombant une multitude de petites entreprises fragiles – en particulier dans le *K-12* où les débouchés décentralisés sont plus développés et lucratifs. Pour les acteurs les plus puissants natifs du numérique, l'engouement pour l'IA éducative est devenu à la fois une opportunité et une menace, poussant la plupart d'entre eux à racheter des start-ups pour les intégrer à leurs catalogues, à l'image de Renaissance Learning et Cambium Learning Group.

En France, les filières du numérique éducatif n'ont pas suivi cette route oligopolistique et n'ont pas permis à des grands groupes concentrés d'émerger. En effet, les quatre industries de l'informatique pédagogique, du *e-learning*, de l'éducation non-formelle et des systèmes de gestion de l'information ont des débouchés moins nombreux qu'aux États-Unis. À l'exception de la formation professionnelle et du parascolaire où ont grandi les demandes d'acteurs privés, les marchés éducatifs sont souvent restés tributaires des subventions publiques. Ce dont découle une fragilité chronique des structures de production informatique (Mœglin, 2010, p.52). Cependant, avec l'engouement pour l'IA, les acteurs anciens étant parvenus à remporter des marchés publics ont pour certains entamé leur mue vers des technologies d'IA symbolique (des ontologies notamment), aux marges du *machine learning*.

Une stratégie de repositionnement idéal-typique de ces firmes émerge ainsi. Elle consiste d'une part dans leur candidature à des marchés publics, permettant aux lauréates de profiter de l'expertise de laboratoires pour incorporer des briques d'intelligence artificielle symbolique à leurs catalogues. S'y ajoute d'autre part une tendance de ces entreprises à capitaliser sur un ethos innovant lié à l'*adaptive learning* pour accroître leur financiarisation, par fusions-acquisitions ou levées de fonds. Les exemples de deux groupes français créés dans les années 2000, Educlever et Maskott, illustrent à la fois cette stratégie idéal-typique et des différences dans leurs trajectoires socio-économiques particulières.

Fondé en 2000, Maxicours est entré dans les industries éducatives par le *e-learning*, en devenant rapidement un acteur important du soutien scolaire, mêlant tutorat traditionnel et didacticiels en ligne. Devenue Educlever en 2007, l'entreprise se diversifie vers l'informatique pédagogique et prend le tournant de l'individualisation des apprentissages à l'école en développant Enseigno, *LMS* à destination des enseignants. En 2014, elle comptabilise un chiffre d'affaires de 2,44 millions d'euros selon Infogreffe.

La firme effectue un virage vers l'IA à la faveur des appels à projets lancés par le Ministère et évoqués plus haut, par des partenariats avec des laboratoires universitaires d'informatique, de sciences cognitives et de sciences de l'éducation. D'abord ceux consécutifs au plan « Numérique pour l'éducation » de 2015, et en particulier l'appel à projet « e-Education 2 ». Educlever conçoit à cette occasion le dispositif *Cartographie des savoirs* en consortium avec le Laboratoire informatique de Grenoble (LIG), le Laboratoire d'informatique en image et systèmes d'information (LIRIS) et le Laboratoire des Usages en Technologies d'Information Numériques (LUTIN Userlab).

Ce que l'entreprise proclame comme le « premier outil de pédagogie adaptative (*adaptive learning*) en France » (Educlever, 2014) est en fait un système situant le niveau de l'élève dans un référentiel de « profils » d'apprenants, et lui assignant des activités en fonction de règles prédéfinies (Figure 13).



Figure 13. Fonctionnement de Cartographie des savoirs selon Educlever (2014).

La firme poursuit sa moisson d'appels à projets ministériels en répondant à la « Banque de Ressources Numériques Éducatives » (BRNE), ce qui lui permet d'enrichir les contenus de son LMS. Educlever remporte ensuite le P2IA en 2019, en développant la plateforme d'enseignement des mathématiques en cycle 2 *Smart Enseigno*. Il s'agit d'un « univers ludique » où les élèves « se déplacent de planète en planète pour réaliser les activités et progresser en mathématiques », et « l'enseignant suit la progression de ses élèves depuis son tableau de bord » et « dispose d'un outil d'aide pour élaborer des parcours personnalisés » (Eduscol, 2023). Il s'agit d'un outil fortement inspiré de *Cartographie des Savoirs*, fonctionnant à partir d'un référentiel de compétences et assignant des activités à l'élève selon un profil identifié et des règles prédéfinies. Comme dans l'appel e-Education 2, l'entreprise a recours à un consortium avec des laboratoires universitaires d'informatique, de sciences cognitives et de sciences de l'éducation, en l'occurrence WIMMICS (équipe de recherche INRIA/Université Côte d'Azur/CNRS spécialisée dans l'intelligence artificielle et le web sémantique). L'attelage est également composé d'entreprises partenaires comme Cabrilog, la société de conception d'outils mathématiques, développant pour l'occasion des générateurs de contenus de remédiation pour les élèves, et LudoTIC, entreprise issue d'un laboratoire universitaire de psychologie de l'université de Nice, spécialisée dans l'ergonomie multimédia des interfaces homme-machine.

L'entreprise a en parallèle enclenché une diversification vers la formation professionnelle en 2017, en lançant *E-scope*, logiciel de remise à niveau dans les fondamentaux pour les salariés.

Ainsi, Educlever a tenté de reprendre le modèle de développement des start-ups *EdTech* : la firme a candidaté à des appels à projets ministériels et collaboré avec des laboratoires de recherche universitaire. En sont cependant ressorties des technologies d'IA symbolique, que la firme a présentées comme des produits d'*adaptive learning* pour se parer d'un ethos innovant. Elle a également diversifié ses activités et opté, sur les marchés aux entreprises, pour une distribution en SaaS par abonnement, afin de sécuriser ses revenus.

En 2020, Educlever fusionne avec Orthodidacte, éditeur de logiciels de certification d'expression écrite à destination des professionnels et principal concurrent du Projet Voltaire. Le groupe atteint alors entre 7 et 8 millions d'euros de chiffre d'affaires, pour devenir l'un des principaux groupes *EdTech* de France (Seveyrat, 2020). En 2022, la maison d'édition Editis rachète Educlever. La firme rejoint sa filiale SEJER, aux côtés des autres maisons d'édition « Jeunesse et Education » du groupe (Nathan, Bordas, Le Robert, Retz notamment).

Maskott a suivi une stratégie de repositionnement similaire, profitant de nombreux appels à projets ministériels pour se présenter en acteur de l'*EdTech*, en connaissant toutefois un processus de financiarisation différent. L'entreprise a pour sa part été fondée en 2004 par Pascal Bringer, alors enseignant de biologie du Puy-en-Velay. Elle visait d'abord l'édition de cédéroms éducatifs et la fourniture de matériels informatiques. En 2013, après l'équipement de plusieurs écoles de la ville en tablettes numériques et le constat d'une absence d'interfaces consacrées à ce type de terminal, la décision aurait été prise d'entamer un repositionnement vers le *e-learning* et la conception d'un *Learning Management System* (Authemayou, 2017).

Comme Educlever, Maskott répond à une série d'appels à projets, à commencer par e-Education 2. La firme auvergnate porte le projet *Tactileo*, consistant dans la « conception de ressources et [la] mise en œuvre en classe d'outils pédagogiques mobilisant les interfaces tactiles (tables interactives et tablettes essentiellement...) (Bouchigny et al., 2016). Comme Educlever, elle profite de l'expertise d'organismes de recherche publics comme le Laboratoire d'intégration de systèmes et des technologies du Commissariat à l'énergie atomique et aux énergies alternatives (CEA) ou l'Institut français de l'éducation (IFÉ). De ce projet, elle retirera l'idée et la marque de son outil, *Tactileo,* à la fois LMS et outil-auteur à destination des enseignants et formateurs. Ceux-ci peuvent y concevoir des exercices et contenus de cours et créer eux-mêmes des arbres de décision pour décider quels parcours proposer aux élèves selon leurs réponses.

On est loin de la promesse d'automatisation de l'*adaptive learning*. Cela n'empêche guère la firme de prétendre être partie prenante de ce mouvement technologique. Témoin un article de 2023 publié sur le blog *Experiences* de Microsoft, qui héberge le LMS sur son *cloud Azure*. Cet article, intitulé « L'*adaptive learning* change les codes de l'apprentissage », prend pour exemple la firme comme une représentante de ces technologies (Microsoft, 2023). Celles-ci sont entendues dans un sens très large, en tant que « méthode éducative qui utilise des programmes informatiques comme outils d'enseignement chargés d'organiser les ressources humaines et les supports d'apprentissage en fonction des besoins uniques de chaque apprenant ». Par ce type de communications, la firme présente sa technologie à la fois sous les atours innovants de l'*adaptive learning* et sous la caution d'une technologisation additive, renforçant le pouvoir d'action de l'enseignant. Pascal Bringer, le dirigeant de la firme, explique ainsi : « Nous ne sommes pas là pour remplacer le professeur, mais pour l'armer avec des outils d'aide à la décision. Nous allons lui remonter des informations plus fines que celles qu'il aurait pu avoir, pour l'aider à être plus précis dans sa pédagogie » (ibid.). La double rhétorique est utile pour se positionner sur des marchés publics de l'Éducation Nationale, qui sont multiples dans l'histoire de l'entreprise.

En 2016, Maskott répond ainsi à l'appel « e-FRAN » au sein du consortium « e.P3C » le liant cette fois au Laboratoire de psychologie sociale et cognitive (Lapsco) et au Laboratoire d'Informatique, de Modélisation et d'Optimisation des Systèmes (LIMOS) de l'Université de Clermont-Ferrand.

Le projet a consisté dans le développement et l'évaluation par *A/B Testing* de Tactileo auprès de 8000 élèves.

En 2022, l'entreprise prend part aux « Territoires Numériques Éducatifs » aux côtés de la start-up Edumalin pour développer, tester et généraliser l'usage de *Tactimalin* dans les établissements scolaires des douze départements partenaires. Dans ce consortium, Maskott a apporté son LMS *Tactileo* et Edumalin ses contenus pédagogiques.

La firme se diversifie en parallèle vers la formation professionnelle, en vendant auprès des organismes de formation son LMS *Tactileo Pro* par abonnement.

Ainsi l'entreprise est parvenue à remporter des appels d'offres ministériels, à partir d'un ethos axé sur l'*adaptive learning* et la technologisation additive, ainsi qu'une forte présence médiatique. Son dirigeant Pascal Bringer est notamment devenu président de l'Afinef (Association Française des Industriels du Numérique dans l'Education et la Formation), l'un des principaux syndicats français d'éditeurs de technologies éducatives, en 2018. Si elle a vu son carnet de commandes s'épaissir au fil des années, la firme a également accumulé des dépenses, notamment en frais de recherche et développement, qui l'ont amenée à une grande fragilité comptable.

Selon les comptes annuels de l'entreprise, elle a connu une croissance continue de son chiffre d'affaires, de 1,18 millions d'euros en 2015, année de la conception de *Tactileo*, à environ 4,11 millions d'euros en 2021. Toutefois l'entreprise a régulièrement rapporté des déficits, d'environ 186 000 euros en 2015, et de près de 85 000 euros en 2021. Elle a en effet investi chaque année des ressources très importantes dans sa recherche et développement : 696 000 euros en 2015 et 916 000 euros en 2021. Ces sommes n'ont pas été couvertes par les crédits d'impôt recherche sollicités chaque année, respectivement à hauteur de 98 000 et 80 000 euros. Ainsi le bilan de 2021 a révélé des dettes d'environ 6,41 millions d'euros. Malgré une levée de fonds à hauteur de 7 millions d'euros en 2022 auprès de plusieurs investisseurs institutionnels comme la Banque des Territoires ou le Crédit Mutuel Arkéa, et douze *business angels* (Crédit Mutuel Arkéa, 2022), le redressement judiciaire a été prononcé en octobre 2023 pour cessation de paiement (Le Figaro Emploi, 2023).

Une situation délicate qui rappelle la précarité de bon nombre d'entreprises du secteur, et ce même lorsqu'elles remportent des marchés publics et effectuent des levées de fonds importantes. En septembre 2024, Maskott a finalement annoncé « rejoindre la famille Cabri », catalogue d'outils numériques de Cabrilog, autre acteur ancien du numérique éducatif (Ludomag, 2024).

Au vu des exemples étudiés, une stratégie de repositionnement idéal-typique s'est vérifiée, consistant dans la candidature à des appels à projets ministériels, permettant aux entreprises de profiter de l'expertise de laboratoires pour développer des technologies essentiellement additives, qu'elles parent d'un ethos innovant lié à l'*adaptive learning*. Si les acteurs qui se sont ainsi positionnés parviennent à accroître leurs chiffres d'affaires, cela se fait au prix de frais en recherche et développement difficilement soutenables. Bien que l'on ne connaisse pas, dans le cas de Maskott, les termes de son union avec Cabrilog, ni, dans le cas d'Educlever, l'ampleur des dettes préalables à son rachat par Editis, on voit que de tels facteurs rendent ces firmes vulnérables face aux stratégies d'acquisition d'autre groupes éducatifs.

Sur les contenus : des fédérations timides entre structures de production informatique et éditeurs de contenus

Comme aux États-Unis, le besoin de contenus pédagogiques variés pour faire fonctionner les algorithmes du côté des start-ups, ainsi que l'absence de savoir-faire technique lié à l'adaptativité chez les éditeurs scolaires est un facteur de rapprochement entre ces deux types d'acteurs.

Les éditeurs de manuels constituent un oligopole à frange très concentré. Comme on l'a vu dans le chapitre introductif de cette partie, Hachette avec ses collections Hatier et Hachette Éducation, Editis à travers Bordas et Nathan, Albin Michel avec Magnard, et Humensis à travers Belin se partagent l'essentiel du marché. On a observé que les groupes français de l'édition tiraient toujours la majorité de leurs revenus du manuel imprimé. L'édition numérique ne représentait ainsi que 13,8% des ventes totales des éditeurs sur le segment scolaire en 2023 selon le Syndicat National de l'Édition (2024, p.12). Moins pressés que leurs homologues états-uniens d'avancer dans la transition numérique de leurs catalogues, ils développent tout de même des manuels numériques et des *Learning Management Systems* pour saisir des relais de croissance (Levoin, 2017). L'édition numérique connaît en effet une croissance continue, son chiffre d'affaires ayant crû de 5,8% en 2023, ce que le SNE explique notamment par le choix de régions de subventionner désormais exclusivement l'achat de manuels numériques.

L'adaptativité représente un nouveau secteur d'avenir que les éditeurs tendent à investir. Cependant, ils ne possèdent pas en interne d'ingénieurs spécialisés dans ces technologies de pointe, et doivent s'allier à des start-ups pour développer ces LMS. Pour les start-ups, le besoin de contenus pédagogiques certifiés pour entraîner leurs algorithmes, ainsi que la quête d'opportunités commerciales les poussent également à nouer des alliances avec ces grands groupes.

Des alliances entre contrats de sous-traitance et fourniture d'ITS « clé en main »

Comme aux États-Unis, ces alliances se sont concrétisées par des contrats de sous-traitance dans les premières heures du marché. Les éditeurs scolaires commandent ainsi des modules adaptatifs auprès de start-ups pour leurs LMS. Cela a été l'objet du contrat noué entre Hachette Education et Knewton en 2015. Comme lors des autres contrats de sous-traitance entrepris par la start-up, l'éditeur a apporté ses programmes et les a formatés, avec l'aide de Knewton, sous forme d'arborescences. C'est en fonction de cette modélisation du domaine que l'algorithme de l'entreprise américaine diagnostiquait le niveau de l'élève et lui attribuait des exercices de difficulté adéquate[107]. Dans ce type de contrats, les coûts en termes de formatage des contenus sont toutefois conséquents pour les éditeurs, qui doivent décliner leurs catalogues existants en niveaux de difficulté, en grains et en formats différents pour faire fonctionner les algorithmes.

Une autre solution est dès lors la commande, par l'éditeur scolaire, d'un ITS développé intégralement par la structure de production, et ce jusqu'à la conception des contenus. Certaines start-ups mieux dotées en compétences et en ressources financières, et/ou plus avancées dans leur développement, fournissent des plateformes modulables comprenant des contenus, des algorithmes et des interfaces. Elles perçoivent ainsi des taux de redevance plus avantageux que dans la simple sous-traitance de modules.

[107] Cet algorithme fonctionne selon l'*Item Response Theory*, enrichie de mesures d'apprentissage variées telles que l'engagement. Son fonctionnement est détaillé dans le livre blanc de la firme, daté de 2015 (Wilson et Nichols, 2015), et par nos soins au chapitre VIII.

C’est la voie empruntée par Evidence B, ayant conçu des ITS pour des éditeurs variés comme Hatier ou Bordas en France. La start-up a pris ce virage, avec des coûts importants à assumer, principalement en matière de recrutement. Devant les besoins en développeurs, en experts pédagogiques pour éditer des contenus pédagogiques en amont, et en *designers* pour concevoir une interface utilisateur appropriée, les levées de fonds sont apparues indispensables. Comme on l’a vu dans l’encadré 7 (p.167), la firme en a réalisé une première en 2019 auprès des fonds Finorpa, Nord-France Amorçage et IRD à hauteur de 2 millions d’euros, puis en novembre 2022 auprès d’Educapital, Finorpa et d'IRD Gestion pour 4 millions d’euros.

La seconde stratégie des éditeurs états-uniens, l’acquisition de start-ups du secteur, est bien plus rare en France. Mis à part le rachat d’Educlever – une firme périphérique des médias adaptatifs – par Editis en 2022, aucun éditeur scolaire français n’a opté pour cette voie. Plusieurs explications peuvent être avancées pour expliquer cette différence. Une première est la fréquence traditionnelle des fusions-acquisitions et autres mouvements de concentration dans les industries éducatives aux États-Unis, et leur relative rareté en France. Une autre explication est la taille largement supérieure du marché *K-12* des médias numériques éducatifs outre-Atlantique, et la propension des grands groupes états-uniens à mener depuis plusieurs années leur transition vers le numérique éducatif.

Les marchés publics, lieux de fédérations et d’éclatements

Les appels d'offres ministériels conduisent tantôt à des partenariats entre éditeurs scolaires et start-ups, lorsqu’ils s'unissent en consortiums pour y répondre, tantôt à des divisions lorsqu’ils présentent des projets concurrents en leur nom propre. Certains appels ont encouragé des fédérations via la constitution de projets groupés. Ainsi, en réponse au « Partenariat d’Innovation Intelligence Artificielle », deux alliances de ce type ont vu le jour :

- Hachette Education a collaboré avec Domoscio pour concevoir le système d’enseignement de la lecture et de l’écriture *Navi*. Selon le site « compagnon », Hachette a fourni les exercices de vocabulaire, quand Domoscio a apporté ses algorithmes d’apprentissage adaptatif et ses outils de *learning analytics* (Navi, 2024).
- Nathan a monté un consortium avec Evidence B pour concevoir *Adaptiv’Math*, pour l’enseignement des mathématiques. Selon le site de l’ITS, l’éditeur a apporté son « expertise pédagogique, ses équipes d’auteurs […], sa connaissance des référentiels et programmes en mathématiques Cycle 2, son expertise en design et en développement d’outils numériques à destination des enseignants et des jeunes enfants, et sa capacité à assurer un lien direct avec les pratiques de classe et à se faire l’écho du projet ». Evidence B a été plutôt chargé de la « coordination du projet, de l’expertise pédagogique pour l’apport en sciences cognitives », et a été « porteur de la compétence IA via les contrats avec le LIP6, l’INRIA et Isograd », une entreprise spécialisée dans l’évaluation algorithmique (Adaptiv’Math, 2021).

Les éditeurs ont l’occasion, en candidatant à ces appels à projets liés à l’innovation, d’entretenir leur position historiquement privilégiée auprès des pouvoirs publics, et de ne pas laisser les start-ups s’adjuger seules un ethos innovant. Néanmoins, on observe que la prise en charge de la réelle plus-value technologique de ces projets échoit aux start-ups ou aux laboratoires universitaires (dans le cas d’*Adaptiv’Math*), alors que les éditeurs ont surtout pour responsabilité d’apporter leurs contenus.

Dans le sens de l'éclatement, d'autres appels font l'objet de candidatures « mono-partenaires » dans lesquelles des structures de production informatique et des éditeurs se concurrencent frontalement – comme cela a été le cas pour le marché public portant sur le service de remédiation en français et en mathématique en seconde, remporté par Evidence B.

Des tentatives d'émancipation de part et d'autre

En dehors des appels d'offres, les éditeurs scolaires et les start-ups tendent également à mener des développements de médias éducatifs en autonomie. Sa collaboration avec Knewton n'ayant guère débouché sur un LMS robuste, Hachette Éducation a ainsi poursuivi la conception de médias éducatifs en interne, reprenant peu ou prou les promesses de l'adaptativité. Ainsi l'éditeur propose-t-il des « exercices interactifs » en complément de plusieurs de ses manuels. Il les présente sur son site comme « adapté[s] à tous les usages et idéaux pour la différenciation » (Hachette Éducation, 2024a). Il s'agit en fait de systèmes pratiquant un test de positionnement à l'issue duquel l'élève est dirigé « vers un parcours d'exercices de niveau 1 ou de niveau 2 » (Hachette Éducation, 2024b). Ce procédé d'adaptation est donc basique, sans collecte de données en continu qui permettrait une réelle adaptativité. En outre, l'arborescence ne présente que deux niveaux de difficulté, ce qui limite grandement le diagnostic du niveau de l'élève et les possibilités d'adaptation. Sur le plan des remédiations, l'éditeur explique qu'il y a « une règle par notion, accessible à tout moment, pour aider l'élève » (ibid.). Il n'y a donc pas d'adaptation du *feedback* selon le profil de l'élève, dans la lignée de l'*adaptive content*. Ce type d'outils statiques illustre le retard technologique des éditeurs, n'ayant pas le savoir-faire en interne pour développer ces algorithmes.

L'avantage pour l'éditeur est cependant de capitaliser seul sur son catalogue de contenus, sans contrepartie financière à concéder à un prestataire. Les éditeurs peuvent en outre capitaliser sur un ethos d'acteur innovant, non plus au service de la forme scolaire, mais d'un projet de différenciation compatible avec les attentes nouvelles du Ministère, notamment en termes de technologisation additive. Les exercices interactifs sont en effet des outils positionnés au service de l'enseignant. Sur son site destiné aux professeurs, l'éditeur met en avant la possibilité d'utiliser ces exercices « lors de séquences en classe entière ou petits groupes ainsi que de manière individuelle pour un entraînement différencié » (Hachette Éducation, 2024a). Ces dispositifs incluent en outre un tableau de bord permettant à l'enseignant de visualiser les résultats de chaque apprenant et de sa classe. Ils doivent donc lui donner des outils pour rendre sa pédagogie plus flexible, et lui offrir des retours informatifs plus nombreux sur ses élèves.

De nombreuses start-ups tentent également de s'émanciper des éditeurs scolaires, en éditant des systèmes de tutorat à destination des organisations éducatives. Elles entrent alors dans une concurrence frontale avec les éditeurs et leurs LMS. Cette diversification des métiers entraîne des coûts d'autant plus grands, en particulier sur la question des contenus. Les marchés publics portant sur la production de ressources (comme la Banque de Ressources Numériques Éducatives ou Édu'up) peuvent être des tremplins pour des firmes lauréates.

En règle générale, les start-ups françaises des ITS exploitent peu les ressources éducatives libres comme le font certains acteurs aux États-Unis, et privilégient surtout les deux autres solutions observées outre-Atlantique : le recrutement d'équipes pédagogiques et le recours aux LLM.

La première voie consiste à recruter des équipes en charge de la conception des contenus, grâce aux dispositifs fiscaux et aux apports de capitaux extérieurs. Cela a été l'option choisie par Lalilo par exemple, qui confie la rédaction d'histoires à lire à des instituteurs de l'Éducation nationale ayant une expérience dans la rédaction de manuels scolaires[108]. Evidence B a également fait ce choix pour se doter d'équipes opérationnelles concevant des exercices avec des chercheurs issus des sciences cognitives, comme Emmanuel Sander ou André Knops.

La seconde solution est représentée par l'essor des grands modèles de langage (*large language models*), permettant de générer des contenus. Ont émergé des services positionnés auprès des enseignants, opposant une concurrence frontale aux catalogues des éditeurs. Ainsi la start-up Neuronys créée en 2018 propose l'outil *Nolej AI* permettant à des enseignants des niveaux primaire, secondaire et supérieur de téléverser des documents écrits et de les voir automatiquement transformés, grâce à *ChatGPT*, en contenus interactifs (vidéos interactives, *flashcards*, glossaires, mots croisés notamment). Ceux-ci peuvent *in fine* être intégrés aux LMS de leurs organisations (Nolej, 2024).

Subsistent cependant les mêmes limitations de l'IA générative qu'outre Atlantique, notamment sur la qualité des contenus pédagogiques. Neuronys réhabilite ainsi le rôle de l'enseignant : « l'intelligence artificielle étant très prolixe, le travail du professeur est de recentrer les modules de formation sur les concepts les plus importants » selon son fondateur (Cormaty, 2023). La firme reconnaît par là que son IA ne peut seule produire des contenus conformes aux programmes ou aux attentes pédagogiques des institutions clientes, et se positionne en faveur d'une technologisation additive.

Comme aux États-Unis, ces technologies donnent un pouvoir accru à un groupe restreint de multinationales propriétaires de ces LLM, dont des *Big Tech*. Cet avantage les situe comme des prestataires de premier choix pour les start-ups désirant générer des contenus pédagogiques. Outre le partenariat d'OpenAI avec Neuronys, Google distribue également son agent Gemini dans son offre *cloud* à destination des start-ups.

La distribution des médias adaptatifs

En aval de la filière, la mise à disposition des médias adaptatifs auprès des publics éducatifs est un enjeu crucial à la fois en termes commerciaux, pour entretenir un contact avec les usagers, et en termes techniques, pour contrôler les modalités de collecte de données nécessaires à l'entraînement des algorithmes. Pour les firmes éditant elles-mêmes leurs ITS, un circuit de distribution est principalement à l'œuvre sur le marché français : les licences de services web en SaaS, par usager, auprès d'établissements scolaires, du supérieur ainsi que d'employeurs.

À l'inverse des États-Unis, aucune application adaptative française n'est vendue directement au consommateur à ce jour, selon la veille que nous avons effectuée. Reste l'application Duolingo, dont la France serait le neuvième plus gros marché en mars 2024 en nombre d'utilisateurs – celui-ci aurait crû de 134% depuis 2020 (Blanco, 2024). Cette firme semble donc représenter seule le marché « BtoC » des médias adaptatifs en France. On peut expliquer cela par la faiblesse structurelle de la vente au consommateur dans les modèles d'affaires de l'*EdTech* : seuls 25% des entreprises du secteur le pratiqueraient selon EY-Parthenon (2022, p.2). Il suppose en effet

[108] Témoin les profils, présentés sur la page d'accueil de Lalilo, de Pierre Puddu, instituteur de l'Éducation nationale ayant rédigé des ouvrages pour Belin, ou Mélanie Pouëssel, institutrice de CE2 et rédactrice pour les éditions Retz.

des apports de capitaux extérieurs suffisants pour grandir et constituer une masse critique d'usagers. Chose que Duolingo est parvenu par exemple à réaliser grâce à un environnement de capital-risque bien plus développé aux États-Unis, alors que les subventions publiques hexagonales visent davantage les éditeurs des niveaux éducatifs institutionnels.

Parmi les entreprises *EdTech* vendant leurs produits aux consommateurs, on compte le plus souvent des applications de formation à distance non-adaptatives selon la définition que nous avons retenue, comme *OpenClassRooms* ou *Le Projet Voltaire* (Encadré 9). Elles s'adressent aux apprenants achetant des prestations via leurs comptes personnels de formation. Si certains prétendent se positionner sur l'*adaptive learning*, leurs services reposent soit sur des arborescences opérant des adaptations seulement à la suite de tests de positionnement, soit sur des algorithmes de répétition espacée limités à la mémorisation.

Encadré 9. Le projet Voltaire, autre service *EdTech* aux marges de l'*adaptive learning*

Le Projet Voltaire est une application de remédiation en orthographe concurrente d'Orthodidacte, comptant plus de 7 millions d'utilisateurs en 2024. Elle fonctionne avec l'algorithme d'« ancrage mémoriel » de Woonoz, sa société éditrice. Celui-ci consiste en un dispositif de répétition espacée d'exercices. Woonoz prétend offrir ainsi de l'« *adaptive learning micro* » – en adaptant le contenu des exercices, par opposition à l'« *adaptive learning macro* » qui porterait sur l'agencement du programme (Woonoz, 2024). Sa technologie est néanmoins limitée à la mémorisation de règles orthographiques, selon une acception restrictive de l'apprentissage qui l'écarte de notre corpus. Par cette stratégie communicationnelle, Woonoz s'apparente à Educlever et Maskott, autres acteurs anciens du numérique éducatif en quête d'un ethos innovant.

Dans le circuit institutionnel prédomine le modèle du club, à travers des licences annuelles vendues par les firmes éditrices aux organisations éducatives. Les entreprises sont cependant soumises à une double difficulté : concilier d'une part la désirabilité de leurs services, en vue d'accroître le recrutement de clients, et d'autre part la satisfaction de ceux présents, pour maximiser leur rétention. Si s'adresser aux usagers et « développer une communauté » est relativement aisé dans la distribution directe au consommateur, l'objectif est difficile à atteindre dans le circuit institutionnel, où la distinction entre usagers, acheteurs et prescripteurs complexifie les décisions d'achat.

Comme aux États-Unis, des acteurs des technologies endossent le rôle d'intermédiaires entre éditeurs et organisations clientes. Si les *Big Tech* sont présents en France aussi dans l'enseignement supérieur ou la formation professionnelle, la situation diffère pour l'enseignement primaire et secondaire, où le marché n'est pas privatisé mais tenu par un acteur ministériel : le Gestionnaire d'Accès aux Ressources (GAR).

L'étude des stratégies de firmes, en référence aux modèles du club et du courtage, doit nous permettre d'observer des mutations dans leurs rapports de force. Ainsi des spécificités affleurent en France, d'une part pour les échelons primaires et secondaire, où le GAR tend à imposer un peu plus la mainmise ministérielle sur le marché éducatif ; et d'autre part pour les marchés de l'enseignement supérieur et du *mobile learning*, où l'incursion des *Big Tech* est jugulée par une politique de régulation menée par les législateurs français et européens.

La redynamisation des marchés éducatifs institutionnels par l'aval

Les secteurs éducatifs institutionnels, allant de l'enseignement primaire à la formation professionnelle, sont plus ou moins complexes à pénétrer pour les firmes éditrices. Aux niveaux

primaire et secondaire, la distinction entre le prescripteur (le Ministère et l'académie), le financeur (les collectivités ou le Ministère), l'acheteur (l'établissement) et l'usager (l'enseignant, l'apprenant voire le parent) crée une complexité aiguë pour l'éditeur. D'autre part, l'éclatement de la demande entre diverses collectivités conduit à une chaîne d'achat complexe pour les offreurs. Dans l'enseignement supérieur français, le marché est contraint par la faible marge financière des universités, ne permettant pas à une demande publique suffisante d'émerger. Les clients adressables sont surtout les écoles privées et de commerce qui investissent le plus dans ces technologies. Dans la formation professionnelle, en appoint de la distribution individualisée instituée par le CPF, les employeurs continuent d'acheter des licences pour des formations en lien avec leurs activités. Les configurations de ces marchés institutionnels, aussi diverses soient-elles, conditionnent la stratégie de distribution des firmes, qui doivent à la fois mobiliser des arguments commerciaux, éducatifs et politiques pour satisfaire ces pôles de réception.

Dans l'éducation primaire et secondaire, la vente traditionnelle de licences par élève ou enseignant constitue également la règle pour les médias adaptatifs. Elle permet aux firmes de sécuriser leurs revenus dans une chaîne d'achat difficile à pénétrer. Leur objectif est, en amont, de recevoir un agrément des pouvoirs publics pour que leurs produits puissent être achetés par des établissements. Convaincre le prescripteur ministériel est historiquement une barrière à l'entrée pour des petites structures ne disposant pas des connaissances du système français ni du « réseau » suffisants, et devant passer par un processus de démarchage long avant de voir leurs produits agréés.

Cependant une initiative ministérielle, le Gestionnaire d'Accès aux Ressources (GAR), constitue un nouveau filtre pour les offres numériques. Il affermit la mainmise ministérielle sur le marché du numérique éducatif en imposant des normes en matière de traitement des données. Il ne s'interpose cependant pas comme une « place de marché », contrairement aux plateformes états-uniennes d'intermédiation.

Le GAR est un dispositif conçu dans le cadre du plan « Numérique pour l'éducation » de 2013, testé à partir de 2017 puis généralisé à partir de 2018 dans les établissements des niveaux primaire et secondaire. Il s'inspire de l'initiative Corrélyce, lancée en mars 2007 à l'échelle de la Région Provence-Alpes-Côte d'Azur par le conseil régional et les Centres régionaux de documentation pédagogique d'Aix-Marseille et de Nice. Corrélyce (pour « Catalogue ouvert régional de ressources éditoriales pour les lycées ») était un portail agrégeant des contenus d'éditeurs variés dans une interface d'authentification unique. Il a été initié alors que les catalogues mutualisés des éditeurs scolaires, comme le Kiosque Numérique des Éditeurs (KNE) ou le Canal Numérique des Savoirs (CNS), manquaient de lisibilité et de diversité en termes de produits (Mœglin, 2010, p.98). Gérard Puimatto, directeur-adjoint du CRDP d'Aix-Marseille à l'époque puis associé au sein du cabinet de conseil Effios, chargé de conseiller le Ministère pour l'implémentation du GAR, justifie le lancement de ce dernier par la nécessité d'empêcher les éditeurs d'instaurer un rapport direct à l'usager, élève ou enseignant, et de capitaliser ainsi sur les données récoltées (Entretien 5) :

> « Si vous prenez Hachette par exemple, il a le portail KNE pour toutes les ressources numériques. KNE- Hachette suit un élève pendant toute sa scolarité du CP jusqu'à la terminale voire après. Et à partir du moment où il a un identifiant unique pour un élève, ça veut dire qu'il peut savoir tout ce que l'élève a fait pendant toute sa scolarité. Alors, je vous laisse imaginer tout l'intérêt que ce genre de données pourrait avoir dans des procédures marketing avec une utilisation du *Big Data* ou des choses de ce genre. Donc il fallait absolument borner ça […] » (ibid.).

Le but de Corrélyce, repris par le GAR, fut à ce titre d'affirmer la responsabilité et le contrôle de l'établissement et de l'Éducation nationale sur le type d'outils contractés et notamment la gestion des données personnelles par les éditeurs. Ce dernier argument prit une place encore plus grande à partir du vote du règlement général de protection des données (RGPD) par le Parlement européen le 24 mai 2016, puis son entrée en vigueur en France le 25 mai 2018. Celui-ci contraint toutes les structures traitant des données à caractère personnel sur le territoire européen à veiller, pêle-mêle, à sécuriser leur traitement, à ne collecter que les données nécessaires à leurs activités et à informer l'usager des finalités du traitement en lui réservant un droit de non-consentement (Commission Nationale de l'Informatique et des Libertés, 2019).

Dans le cadre du RGPD, l'établissement scolaire a été reconnu responsable du traitement des données des élèves. Toutefois, la mise en pratique de ces responsabilités, telle que la tenue d'un registre de traitement ou la réalisation d'études d'impact pour chaque éditeur, est apparue intenable pour un seul établissement. Dès lors, le Ministère a également justifié la généralisation du GAR par le besoin de prendre en charge nationalement cette responsabilité pour « soulager » les établissements.

Le GAR consiste donc en une plateforme « branchée » aux ENT des établissements, donnant accès, pour chaque utilisateur, aux licences que son établissement lui a achetées. Ce dispositif assure le partage de données personnelles entre l'usager et l'éditeur dans un cadre sécurisé, que schématise le Ministère en Figure 14.

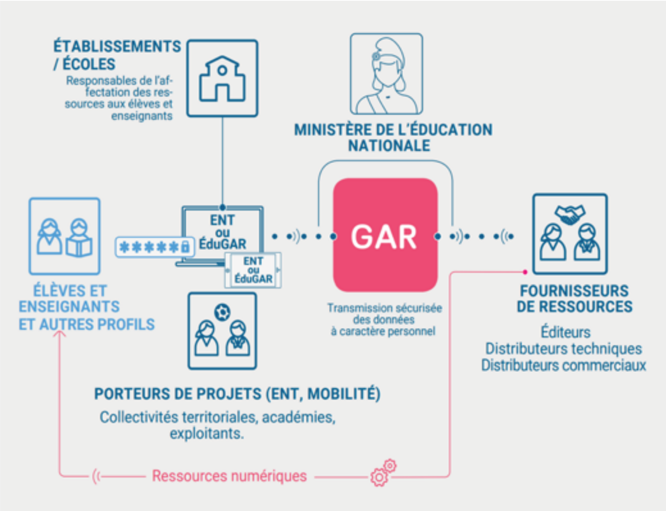


Figure 14. Schématisation du fonctionnement du GAR.
Source : Ministère de l'Éducation nationale et de la Jeunesse (2024b)

Il présente deux versants : un premier d'ordre technique consistant dans l'identification unique, et un autre d'ordre juridique, « sans doute le plus important » selon Gérard Puimatto, consistant à déporter la responsabilité de la gestion des données personnelles sur le Ministre. L'arrêté ministériel du 18 décembre 2017 annonçant la création du dispositif soutient que tout éditeur inscrit au GAR s'engage en tant que sous-traitant du Ministère pour le traitement des données personnelles.

Pour figurer dans le GAR, il n'y a guère de licence à payer au gestionnaire comme l'impose Clever aux États-Unis, mais plutôt un agenda de prérequis techniques et juridiques. Les services du GAR évaluent et suivent l'outil de l'éditeur pendant son développement pour assurer la conformité de son

traitement des données. L'offreur doit *in fine* transmettre une déclaration de conformité décrivant quelles données sont collectées, à quelles fins, et de quelles manières elles sont traitées et sécurisées.

Une fois distribué et utilisé, l'éditeur se voit livrer les données d'apprentissage de ses usagers via le GAR, selon le principe de proportionnalité du RGPD : seules les traces utiles au fonctionnement du service sont transmises. Comme nous l'explique Gérard Puimatto, pour des services de suivi et d'assignation de tâches comme les ITS, les données d'identification peuvent aller jusqu'aux noms et prénoms. Par ailleurs, les travaux réalisés par l'élève dans l'outil ont également le statut d'œuvres protégées par le Ministère, conformément à un second arrêt ministériel. Les éditeurs sous-traitants doivent donc également les sécuriser (Entretien 5).

Gérard Puimatto relate les négociations qu'il a conduites avec les start-ups *EdTech*. Il note l'incompréhension de la plupart des dirigeants de start-ups devant ces contraintes, *a fortiori* quand ils ne sont pas accompagnés par des incubateurs qui les sensibiliseraient au fonctionnement du marché scolaire :

> « Il y a toujours une phase difficile au début, ils [les start-ups] arrivent avec leur solution qui est extraordinaire et ils ne comprennent pas pourquoi le ministère ne saute pas dessus en fait. Donc au début, il peut y avoir des résistances, il y en a régulièrement pour beaucoup de start-ups qui ne sont pas passées par un canal un peu structuré pour comprendre un peu le marché. Aujourd'hui, il y a les *EdTech* par exemple qui discutent beaucoup entre eux, ils sont au courant, il n'y a pas de problème. C'est une association, ils sont très opérationnels. L'AFINEF c'est pareil. Ces gens-là, ils ont des associations, ils savent, ils n'ont pas besoin de beaucoup d'infos complémentaires. Par contre tous ceux qui passent à côté, par exemple pour ceux qui passent par les incubateurs qu'il y a dans quelques collectivités, les débuts sont un peu compliqués, parce qu'il faut qu'ils comprennent que l'éducation n'est pas un marché comme un autre » (ibid.).

Ainsi ces prérequis tendraient à discriminer des start-ups de petite taille et isolées, n'ayant pas l'expertise du marché scolaire et éprouvant plus de difficultés pour se conformer aux obligations légales en matière de traitement des données personnelles. En cela, cet agenda technico-juridique constituerait possiblement une nouvelle barrière à l'entrée pour des structures d'édition informatique.

Sur l'aspect technologique, le conseiller souligne l'aspect hautement « marketing » de l'intelligence artificielle, estimant qu'une proportion minime des start-ups prétendant faire de l'IA en font réellement. Parmi celles-ci, il cite l'exemple d'Evidence B. Un procédé de « pseudonymisation » des données a été négocié avec elle pour lui permettre de conduire des recommandations personnalisées : les données à caractère personnel sont pseudonymisées puis « dépseudonymisées » pour pouvoir rétablir le lien entre l'élève et ses réalisations dans l'ITS. Gérard Puimatto voit dans ce processus une étape critique en matière de sécurité informatique, devant faire l'objet d'une surveillance étroite de la part du Ministère :

> « En fait, c'est ce processus de pseudonymisation qui aujourd'hui est le plus critique, parce qu'évidemment il peut y avoir des failles de sécurité monstrueuses. Alors les gens sérieux, professionnels de l'IA, connaissent et maîtrisent bien, mais les bricoleurs qui arrivent, il y en a quand même un petit peu dans les *EdTech*, il faut beaucoup plus se méfier parce qu'ils n'ont pas toujours des processus vraiment très fiables dessus » (ibid.).

Sur les modèles économiques, nulle intervention du Ministère n'est prévue. Cela répond à un souhait cher des éditeurs de garder leur autonomie éditoriale, datant des débuts du manuel scolaire. La licence vendue de l'éditeur à l'établissement, conformément au modèle du club, reste ainsi hégémonique.

Gérard Puimatto relate une levée de boucliers à l'époque des premières négociations sur les modalités du GAR, allant jusqu'au refus qu'un catalogue des éditeurs présents soit publié par le Ministère :

> « C'était un refus massif de la profession. En 2015-2016 quand on a fait ses préfigurations, on a eu des quantités de réunions avec les éditeurs, évidemment leurs représentants, donc le Syndicat national de l'édition, l'association de Savoir lire, l'AFINEF, à l'époque il n'y avait pas encore les *Edtech*. Et leur souci était de dire « ok on veut bien comprendre qu'on est sous-traitants sur les données mais on veut préserver notre autonomie éditoriale donc il est hors de question que le ministère donne un avis sur la qualité pédagogique et il est hors de question que le ministère participe à quoi que ce soit sur le plan de la commercialisation et de la diffusion économique ». À tel point qu'ils nous ont demandé à l'époque de ne pas faire de catalogue. Le catalogue, il y en a un interne au GAR évidemment, mais il n'est pas publié » (ibid.).

Un dispositif de *monitoring,* le *Portail GAR*, permet à tous les publics autorisés (éditeurs, cadres de l'éducation au niveau national, académique, départemental, chefs d'établissement), d'avoir accès à des statistiques d'usage de chaque ressource. Ces statistiques sont utilisées par les collectivités et les directions d'établissement à des fins gestionnaires proches du *Social Efficiency Movement*, pour veiller au retour sur investissement lié à l'achat d'une ressource.

Gérard Puimatto insiste en définitive sur l'effet « modélisant » du GAR, ayant été adopté par les collectivités qui prennent en charge les équipements en ressources des établissements. Pour les ENT, qui constituent l'un des principaux marchés du numérique éducatif gérés par les collectivités, cela a ainsi incité l'ensemble des organisations éditrices à se conformer au GAR, afin de pouvoir répondre aux appels d'offres. Plusieurs collectivités ont également lancé des marchés publics pour l'acquisition de ressources, à l'image de l'initiative « Lycées 4.0 » de la région Grand Est en 2016, consacrée à l'équipement des lycées en manuels numériques. Ces appels ont imposé la conformité au GAR aux établissements (pour l'obtention de crédits) ainsi qu'aux éditeurs candidats.

Le dispositif est rapidement devenu incontournable dans l'enseignement secondaire, touchant, pour la rentrée 2023, plus de 9 000 collèges et lycées, soit 85% des établissements du secondaire, et 23 000 écoles, soit environ 48% des établissements du primaire – un rapport moindre étant donnés leurs plus faibles usages des ENT (GAR, 2024).

Ainsi le Ministère a su imposer sa place centrale dans la distribution des médias éducatifs, lui permettant d'encadrer le traitement des données par les éditeurs d'ITS. Ceci à la faveur d'un impératif juridique lié à la régulation européenne en matière de protection des données personnelles. Toutefois, la socio-économie du numérique éducatif reste dominée par le modèle du club, à la différence des États-Unis qui tendent à lui superposer le modèle du courtage. Dès lors, les éditeurs gardent leur position centrale, récoltant directement les revenus des licences, mais doivent dorénavant satisfaire aux obligations du GAR en matière de traitement des données personnelles pour entrer dans le marché.

Les tentatives d'incursion des Big Tech *dans la distribution des produits éducatifs*

Les suites bureautiques de Google et Microsoft, largement adoptées par les établissements scolaires et les universités en France dans les années 2010 et pendant le Covid-19, conduisent, comme aux États-Unis, un système de « place de marché » où sont appliquées des redevances sur les achats institutionnels, suivant le modèle du courtage.

Ces médias font cependant l'objet de restrictions de plus en plus fortes de la part des pouvoirs publics, à nouveau motivées par le traitement des données personnelles. La circulaire publiée par le Premier ministre en juillet 2021, « relative à la doctrine d'utilisation de l'informatique en nuage par l'État » (ou doctrine « *cloud* au centre ») invite les différents ministères à s'assurer que les organisations sous leur autorité ayant recours à des offres de *cloud* commercial soient protégées de toute réglementation extracommunautaire, et que ces services respectent la souveraineté et la sécurité des données (Secrétariat général du Gouvernement, 2021). Cette doctrine s'inscrit dans la continuité de la politique européenne, et notamment de l'arrêt « Schrems II » pris par la Cour de Justice européenne en 2020, invalidant les garanties fédérales de protection (« *Privacy Shield* ») sur les données transitant vers les États-Unis.

Une note du directeur interministériel du numérique datant du 15 septembre 2021 précise en outre que la suite collaborative *Microsoft Office 365* n'est pas conforme à la doctrine ministérielle. En cause, la loi « *Cloud Act* » votée en 2018 outre-Atlantique, qui autorise l'État fédéral à imposer à une entreprise opérant sur son sol de lui communiquer des données, où que celles-ci soient stockées dans le monde.

En matière d'enseignement supérieur, la Commission nationale de l'informatique et des libertés (CNIL) a également recommandé en octobre 2021 aux établissements de ne plus utiliser des services états-uniens de travail collaboratif comme *Microsoft Teams*. C'est ainsi qu'en 2023, le Ministère de l'enseignement supérieur français a mis à disposition gratuitement des établissements quatre plateformes « souveraines » et *open source* (Cormaty, 2023b).

Dans l'Éducation nationale, suite à une question parlementaire déposée par le député Modem Philippe Latombe en 2023, le Ministre Pap Ndiaye s'est prononcé contre l'utilisation de Microsoft et de Google dans les écoles, collèges et lycées (Assemblée Nationale, 2023). En se basant sur la doctrine ministérielle et européenne, il a annoncé avoir demandé aux recteurs d'arrêter tout déploiement ou extension des suites *Microsoft Office 365* et *Google Workspace for Education* (ibid.).

Ces régulations visent également à soutenir des outils « souverains ». Dans cette idée, la DNE et Inria prennent part à deux initiatives à l'échelle nationale et européenne :

- un projet de « plateforme de données de l'éducation » décidé à la suite des États généraux du numérique. Il vise à « répondre à quatre ambitions : soutenir la collecte du patrimoine de données éducatives, jouer un rôle de tiers de confiance vis-à-vis de la circulation des données, accompagner l'écosystème éducatif dans le développement des usages et acculturer à l'usage des données » (Ministère de l'Éducation nationale et de la Jeunesse, 2023a, p.31).
- le *Data Space Education and Skills* (DASES), dans le cadre du projet de *cloud* numérique européen GAIA-X. Il doit s'agir d'« une plateforme qui permettra d'échanger et de partager les données d'éducation pour développer des accompagnements pertinents pour les élèves et les étudiants (apprentissage, orientation, formation) » (Ministère de la Transformation et de la Fonction Publiques, 2024).

Ces initiatives, encore balbutiantes, montrent néanmoins la volonté politique française et européenne de ne pas laisser les « géants des technologies » détenir des gisements de traces numériques potentiellement sensibles. Pour les acteurs de l'*EdTech*, elles représentent l'opportunité d'entraîner leurs algorithmes sur un vaste répertoire de données ouvertes – une perspective susceptible d'atténuer les difficultés des jeunes structures aux bases d'utilisateurs restreintes.

Dans l'autoformation et la distribution directe au consommateur, nous avons vu que seul Duolingo était parvenu à importer ce circuit de distribution dans l'Hexagone, alors qu'aucun acteur national n'a su franchir les barrières à l'entrée pour l'adopter. En France, la prédominance des géants des technologies sur ce circuit est extrêmement forte. Comme aux États-Unis, Google et Apple se partagent le marché des systèmes d'exploitation, mais selon une répartition des parts de marché très marquée à l'avantage du premier : 73,8% de parts de marché pour *Android* contre 25,9% pour *iOS* en juin 2024, selon le cabinet d'études Kantar (Kantar, 2024).

La mainmise des deux firmes, à travers leurs magasins d'applications, sur les marges des éditeurs est de plus en plus combattue par la Commission européenne. Visant à dénouer les pratiques anticoncurrentielles des *Big Tech*, un texte important, le *Digital Market Act* a été voté le 5 juillet 2022 et est entré en vigueur en France le 6 mars 2024. Il contient une série d'obligations et d'interdictions à l'encontre de ceux qu'il définit comme des « contrôleurs d'accès ». Est ainsi qualifiée une entreprise qui :

- « occupe une position économique forte, a une incidence significative sur le marché intérieur et est active dans plusieurs pays de l'UE ;
- occupe une position d'intermédiation forte, ce qui signifie qu'elle relie une base d'utilisateurs importante à un grand nombre d'entreprises;
- et occupe (ou est sur le point d'occuper) une position solide et durable sur le marché » (Commission européenne, 2024).

L'un des objectifs de cette législation est de lever des barrières à l'entrée pour permettre aux « innovateurs et [aux] jeunes entreprises technologiques » de disposer « de nouvelles possibilités d'affronter la concurrence et d'innover dans l'environnement des plateformes en ligne, sans devoir se soumettre à des clauses et conditions abusives limitant leur développement » (ibid.). Parmi ces obligations, la Commission contraint les contrôleurs d'accès à « autoriser les entreprises utilisatrices à promouvoir leur offre et à conclure des contrats avec leurs clients en dehors de leur plateforme » (ibid).

Face à cette décision similaire à celle de la Cour Suprême états-unienne à l'encontre d'Apple, les *Big Tech* adoptent également des mesures analogues à celles qu'ils ont prises outre-Atlantique. Ainsi, Apple a annoncé en janvier 2024 se conformer à cette obligation en autorisant les places de marché tierces, dans un communiqué où la firme souligne également les « nouvelles voies pour les logiciels malveillants, la fraude et les escroqueries, les contenus illicites et nuisibles, et d'autres menaces à la vie privée et à la sécurité » que cette disposition représenterait (Apple, 2024). Comme aux États-Unis, la firme à la pomme a en fait assoupli les redevances des petits éditeurs tout en créant une nouvelle taxe pour les plus gros éditeurs contournant son système de paiement. Pour les applications *iOS* téléchargées depuis l'*App Store*, les commissions sont ramenées à 10% pour les petits développeurs contre 15% auparavant, et 17% sur les transactions de biens et services numériques, contre 30% jusqu'alors. Les développeurs contournant son système de paiement doivent quant à eux s'acquitter de cinquante centimes d'euros par nouvelle installation ou mise à jour à partir d'un million de téléchargements dans l'année. Par ce qu'il nomme des « frais de technologie de base », Apple entend préserver sa position dominante dans la distribution des applications les plus lucratives. L'annonce a engendré une levée de bouclier en France, à l'image d'une lettre de protestation envoyée à la Commission européenne par un groupe de 34 entreprises et associations professionnelles de l'économie numérique, dont Deezer ou Spotify (Debes, 2024).

Google a annoncé, en juillet 2022 suite au vote du DMA, un taux de redevance pour les développeurs choisissant un système de paiement alternatif inférieur de 3% à celui des éditeurs adhérents du Google Play Store. (Vitard, 2022). La firme de Mountain View a en outre mis en place en mars 2024 deux nouveaux types de redevances sur les transactions passées dans des systèmes de paiement tiers. Les premiers sont des « frais initiaux » s'élevant à 10 % pour les achats intégrés et à 5% pour les abonnements, pour une période limitée à deux ans. Les seconds sont des « frais de services continus » relatifs au développement des mises à jour, du contrôle parental ou encore des correctifs de sécurité, s'élevant à 17% pour les achats intégrés et à 7% pour les abonnements. Analogues aux « frais de technologie de base » d'Apple, ces redevances peuvent être abandonnées par le développeur au bout de deux ans, s'il obtient le consentement de l'utilisateur (Bazoge, 2024).

Les deux *Big Tech* ont ainsi créé de nouvelles taxes sur les achats passés via des systèmes de paiement alternatifs affectant en particulier les plus gros développeurs. Duolingo, seul représentant des médias adaptatifs dans ce circuit de distribution en France, est resté jusqu'ici fidèle aux systèmes de paiement d'Apple et Google, qui grèvent une grande part de ses revenus comme on l'a vu dans le chapitre précédent.

Jusqu'à présent, et sous réserve de nouvelles actions menées par la Commission européenne, ces *Big Tech* maintiennent leurs positions dominantes dans la distribution des applications. Certains acteurs français de l'*EdTech*, pour l'essentiel des petites entreprises éloignées des quantités de téléchargement engendrées par Duolingo, pourraient bénéficier des allègements opérés par les *Big Tech* sur les taux de redevance des petits développeurs. Les nouvelles taxes instaurées par ces « contrôleurs d'accès » sur les plus grosses entreprises épaississent toutefois un plafond de verre.

Les trois domaines générateurs de valeur dans la filière des médias éducatifs, à savoir le développement d'algorithmes, la production de contenus et la distribution des médias éducatifs, ont fait apparaître des rapports de force nouveaux, tantôt communs avec le cas états-unien, tantôt spécifiques. Dans le développement d'algorithmes, les principaux acteurs impliqués, les start-ups, sont soumises à des inégalités importantes, entre celles fragiles, vivant d'aides publiques pour investir à perte dans leur R&D, et les autres parvenant à la rentabilité en remportant des marchés publics et en effectuant des levées de fonds. Les acteurs natifs du numérique, loin d'être aussi concentrés qu'aux États-Unis, tentent d'adopter une stratégie de repositionnement vers l'*EdTech* et les marchés publics. En matière de production de contenus, les alliances entre éditeurs scolaires propriétaires de catalogues certifiés et les structures de production se sont développées mais les acteurs les mieux dotés de part et d'autre tendent à poursuivre des chemins séparés. Dans la distribution des médias adaptatifs auprès des établissements, la pénétration des LMS de *Big Tech* tels que Google ou Microsoft est entravée par les régulations européennes et françaises, de même que la constitution d'une centrale d'accès ministérielle dans le primaire et le secondaire, le GAR, qui assoit la mainmise ministérielle sur ces échelons. Prévaut donc le modèle du club dans le circuit institutionnel des médias adaptatifs, alors que la distribution directe au consommateur, témoin d'innovations en termes de modèles d'affaires aux États-Unis, n'a jusqu'ici pas permis le développement d'acteurs hexagonaux.

Les trois tendances socio-économique patentes aux États-Unis, à savoir la montée en puissance des start-ups en matière d'innovations, la dynamisation de la filière par l'aval, et les tentatives d'imposition des *Big Tech* dans l'ensemble de l'industrie, subissent quelques nuances en France. D'une part, la montée en puissance des start-ups est réelle mais encore contrainte, alors que

la demande institutionnelle est encore timide et les levées de fonds largement moins développées. D'autre part, la dynamisation par l'aval est réelle, mais s'exerce sous le contrôle ténu des pouvoirs publics, qui préservent le modèle du club dans le circuit institutionnel. Troisièmement, ceux-ci jugulent fortement la pénétration des *Big Tech* dans les marchés institutionnels. Ces tendances spécifiques sont conditionnées par la constitution de compétences et de solidarités professionnelles parmi les praticiens des diverses filières, qui, comme aux États-Unis, s'unissent de façon croissante sous l'« étendard » *EdTech*.

3. Des professionnalités dynamisées par l'*EdTech*

On a souligné plus haut qu'une clé de voûte des industries éducatives était la constitution de professionnalités transversales, c'est-à-dire de compétences et d'intérêts unissant des professionnels variés de l'éducation, à même de soutenir et d'influencer l'industrialisation des médias, tant dans les débats politiques que dans les stratégies des firmes. Il convient d'observer ceux-ci à la lumière des spécificités économiques et culturelles françaises.

Des réseaux diversifiés et dynamiques

En France l'existence de ces professionnalités est d'abord visible dans la création de communautés de deux types : des syndicats d'acteurs industriels formulant des préconisations auprès des pouvoirs publics, et des réseaux d'incubateurs initiés par des acteurs publics et privés.

Les syndicats professionnels du numérique éducatif

Le premier type de réseau est principalement représenté par deux organisations, l'Afinef et EdTech France. La première, créée en 2012, a pour mission de « rassembler, fédérer et structurer l'ensemble des entreprises de la filière afin d'assurer l'interface de la profession avec les autres acteurs présents sur ces secteurs pour développer une puissante industrie numérique pour l'éducation et la formation » (Afinef, 2024a). Sa création est liée au pôle de compétitivité Cap Digital, organisateur en 2010 des « Assises de l'Éducation et de la Formation Numériques », à l'issue desquelles un projet de fédération des acteurs du numérique éducatif aurait émergé (Afinef, 2024b).

En septembre 2024, le réseau affiche sur son site près de 120 membres, dont de nombreux acteurs anciens du numérique éducatif (Maskott, Educlever, Cabrilog), quelques start-ups de l'*adaptive learning* (Lalilo, Evidence B), des médias traitant d'éducation (Ludomag, Le Monde, Educ'Arte), des acteurs de la distribution issus d'éditeurs (le Canal Numérique des Savoirs) ou des *Big Tech* (Google for Education, Microsoft) (Afinef, 2024c). La direction du réseau donne à voir l'implantation forte d'acteurs anciens du numérique éducatif. En témoigne la présence ancienne dans les bureaux exécutifs de Pascal Bringer, directeur de Maskott, ou de Pierre Laborde, directeur de Cabrilog et actuel président de l'association.

Le syndicat produit en outre des enquêtes statistiques (comme le Baromètre du *Digital Learning*) et des recommandations à l'attention des pouvoirs publics et des professionnels de l'éducation.

EdTech France est un réseau plus récent, davantage axé sur les start-ups. Son action fondatrice a été la publication, en novembre 2017, d'un manifeste intitulé « faire de la France la *EdTech Nation* » par

ce qui était alors un collectif d'entrepreneurs (EdTech France, 2024a). C'était un clin d'œil au mantra de « *startup nation* », désignant Israël au mitan des années 2010, avant qu'Emmanuel Macron n'en fasse son aspiration pour la France durant la campagne présidentielle de 2017.

Constituée en association en mai 2018, EdTech France « rassemble dans un même espace de dialogue et de rencontre des entrepreneurs *EdTech* ainsi que des partenaires "engagés" » selon la page d'accueil de son site web (EdTech France, 2024b). Parmi les « entrepreneurs *EdTech* », le réseau fédère près de 400 acteurs, allant de jeunes start-ups (Neuronys), d'anciennes start-ups étant parvenues à s'internationaliser (Domoscio, Evidence B, Lalilo), d'acteurs anciens du numérique éducatif (OpenClassrooms, Educlever), ainsi que quelques associations (Agir pour l'école) (EdTech France, 2024c). S'y ajoutent les « partenaires engagés », organisations soutenant les événements et les actions du réseau. Le site recense des fonds européens de capital-investissement (Educapital, Bright Eye Ventures), des médias (L'Étudiant, CampusMatin), des acteurs des infrastructures technologiques (AWS, Samsung), des écoles de commerce (EM Lyon, Neoma Business School), des universités (CY Cergy Paris Université, IUT Villetaneuse), des organismes de formation publics (e-GRETA-CFA), des associations favorisant la recherche et le déploiement des technologies en éducation (l'Association des Technologies de l'Information pour l'Éducation et la Formation, les Acteurs de la Compétence) ou encore des réseaux *EdTech* régionaux (EdTech Lyon, EdTech Grand Ouest) (EdTech France, 2024b). Enfin le « Club d'experts EdTech France » est un programme animé par des entreprises spécialisées dans le conseil, faisant bénéficier des start-ups de leur expertise en matière de fonctionnement des marchés éducatifs ou de levées de fonds.

L'organisation diffuse en outre un podcast portant sur les développements du secteur *EdTech*.

Ces syndicats jouissent de liens forts avec les services de l'État. Témoin les conventions de partenariat signées en mai 2021 par EdTech France et en mai 2023 par l'Afinef avec la Banque des Territoires. La direction de la Caisse des dépôts et consignations créée en mai 2018 est engagée dans l'accompagnement et le financement de projets des acteurs territoriaux engagés dans des enjeux de société comme l'éducation. Ces partenariats visent d'une part à faciliter les échanges entre les acteurs territoriaux publics et privés et ces organisations, par la co-organisation d'événements. Ils consistent d'autre part à publier des études communes sur le secteur, à l'image du rapport d'EY-Parthenon cité plus haut, auquel ont participé la Banque des Territoires et EdTech France.

L'Afinef et EdTech France sont en outre liés au Ministère de l'Éducation Nationale, auprès duquel ils mènent des activités de lobbying. Ainsi l'idée de « compte ressources » pour les enseignants, devant permettre à ceux-ci de commander des licences dans les outils de leur choix, est une recommandation ancienne des deux syndicats, qui a finalement été reprise par le Ministère dans sa Stratégie « Numérique pour l'éducation 2023-2027 » (en tant qu'expérimentation pour l'année 2024). Autre signe de l'influence de ces syndicats sur le Ministère, la directrice générale d'EdTech France de 2020 à 2023, Anne-Charlotte Monneret, est devenue en août 2023 la conseillère en charge du numérique et de l'innovation au sein du cabinet de Gabriel Attal, alors Ministre de l'éducation et de la jeunesse.

Tous deux sont en outre partenaires du principal salon français consacré aux acteurs des technologies éducatives, Educ@tech Expo (anciennement Educatech-Educatice). Selon le site de l'événement, pour l'édition 2023, les visiteurs étaient issus de l'Éducation Nationale à hauteur de 56%, des collectivités à 33%, puis de l'industrie à 11% (Educ@tech Expo, 2024). Parmi les professionnels

de l'Éducation Nationale, 46% étaient des personnels enseignants/formateurs/documentalistes, 25% des cadres du Ministère de l'éducation nationale et des académies, 13% des responsables de services informatiques et numériques et 10% des chefs d'établissements. Parmi les représentants et professionnels issus des collectivités, 53% venaient des mairies, 37% des départements et 10% des régions (ibid.).

Les exposants sont également variés. Ils sont issus de différentes filières du numérique éducatif (des matériels et équipements informatiques à l'édition scolaire, en passant par l'informatique pédagogique, la formation à distance et les systèmes numérique de gestion et d'information), du Ministère de l'Éducation Nationale ainsi que de structures publiques accompagnant le développement du secteur (Réseau Canopé, la Banque des Territoires).

Les principaux éditeurs scolaires (Hachette Education, Editis) prennent part aux événements du secteur. Néanmoins, ils disposent de leur propre syndicat distinct d'EdTech France ou de l'Afinef : Les Éditeurs d'éducation. Fondé en 1985 par Belin, Bordas, Hachette, Hatier, Magnard et Nathan, ce syndicat regroupe une trentaine de maisons d'édition scolaires. Il porte les intérêts collectifs des éditeurs auprès des pouvoirs publics dans les enjeux liés aux ressources pédagogiques et aux médias éducatifs imprimés et numériques.

Les *Big Tech* ont quant à eux entrepris une stratégie d'implantation dans ces organisations, comme aux États-Unis. AWS figure parmi les « partenaires engagés » d'EdTech France ainsi que les « membres partenaires » de l'Afinef aux côtés de Google Education et de Microsoft. Ils exposent également lors de salons comme Educ@tech Expo.

Certaines de ces multinationales ouvrent en parallèle des programmes d'« accélération », non pas de façon spécifique à l'*EdTech* comme aux États-Unis, mais pour les start-ups françaises en général. Ainsi *Cloud Google for Startups* propose-t-il de bénéficier de services commerciaux, d'aides techniques et de crédits vers l'environnement *cloud* du groupe (Google Cloud, 2024).

Les incubateurs publics et privés

Le second type de communautés interprofessionnelles correspond ainsi aux incubateurs soutenant le développement de start-ups. Outre les programmes de *Big Tech*, ceux-ci sont initiés par des organisations aux profils variés :

- Les structures ministérielles comme l'« *accélérateur d'innovations pédagogiques* » de Réseau Canopé, dont l'accompagnement passe par des conseils personnalisés pour les start-ups, des tests utilisateurs, ou encore une aide à la conception de contenus (Réseau Canopé, 2024b).
- Les incubateurs associatifs comme l'accélérateur « *Passerelles* », créé en 2021 en période de pandémie par la Banque des Territoires et l'association Makesense. Celui-ci cherche à accompagner des start-ups « à impact » visant l'inclusion scolaire, l'égalité des chances, l'attractivité des territoires ou encore la prévention des risques psychosociaux (Makesense, 2024).
- On peut aussi citer les incubateurs d'établissements d'enseignement et de recherche, accompagnant des entreprises extérieures ou des projets entrepreneuriaux d'étudiants et de chercheurs. Du côté des universités, « *CY edtech labs* » a été lancé en 2023 par CY Transfer, la structure en charge de la valorisation de CY Cergy Paris Université, en partenariat

avec la *University College* de Londres (CY Cergy Paris Université, 2024). Centré sur la recherche de « preuves d'impact », il contribue à organiser des expérimentations dans des établissements scolaires du Val-d'Oise, et offre un accès à l'expertise de chercheurs en sciences de l'éducation de quatre laboratoires de CY Cergy Paris Université. Des écoles de management disposent également de leurs programmes, à commencer par NEOMA Business School et son « *NEOMA Edtech Accelerator »*. Celui-ci accompagne pendant cinq mois des start-ups dans leur développement (NEOMA Business School, 2024). De son côté, l'EM Lyon a créé en 2017 l'incubateur « *Ed.Job.Tech »*, misant sur un accompagnement intensif de six mois comprenant l'accès à un réseau de professionnels et de chercheurs de l'école, ainsi que la mise à disposition de ressources numériques et de crédits pour des services *cloud* partenaires (EM Lyon Business School, 2024).

- Les incubateurs privés sont lancés par des entreprises des technologies en vue de constituer autour d'eux des réseaux d'acteurs partenaires. Un exemple est « *L'Accélérateur »*, initié par Unowhy. Fondée en 2007, cette entreprise qui s'auto-proclame « première *EdTech* française » s'est spécialisée dans la conception de terminaux numériques et de logiciels éducatifs avec son offre SQOOL. Cet incubateur offre des espaces de travail ainsi qu'un accompagnement pour 6 à 12 mois (Unowhy, 2024).

En définitive, les réseaux professionnels français de l'*EdTech* se sont fortement dynamisés à partir de la seconde moitié des années 2010, avec l'apparition d'un deuxième syndicat et la multiplication des programmes d'incubation initiés par des structures variées. Dans les salons, la diversité des visiteurs et des exposants, issus à la fois d'organisations privées et publiques, confirme ce mouvement de fédération des professions éducatives autour de l'*EdTech.* Au-delà d'espaces de dialogue communs, les publics partagent-ils de nouvelles compétences voire des idéaux pédagogiques ?

La construction de professionnalités

Le livre blanc cité précédemment, intitulé « Les technologies au service de l'éducation : une chance pour la France » (Livingstone-Wallace et Welter, 2024) a formulé en mars 2024, au nom d'un collectif emmené par l'Afinef et EdTech France, douze recommandations à destination des décideurs politiques. Cette initiative commune part du constat dressé par ces divers acteurs de la « nécessité d'accélérer la transition numérique » :

> « Alors que l'État semble bel et bien conscient de la nécessité d'accélérer la transition numérique dans la formation et l'éducation, l'accélération n'est pas à l'œuvre. Les mesures peinent à se déployer, les bonnes intentions restent (encore) trop souvent à l'état de « pilote ». Autrement dit, il semble urgent de passer des mots aux actes... et de rendre l'accélération effective » (ibid., p.106).

Dans cette direction sont formulées des propositions diverses. Néanmoins l'idée des syndicats n'est guère de bouleverser la forme scolaire en place dans les organisations éducatives :

> « Les *EdTech* sont un levier indispensable – mais insuffisant – pour répondre aux enjeux de montée en compétences et d'inclusion des apprenants. Elles sont porteuses de promesses, mais ne seront efficaces que couplées à un enseignement classique, où les professeurs et les formateurs gardent le rôle central » (p.106).

Un discours à la fois réformateur et conciliant avec les organisations traditionnelles est donc de mise. La première des recommandations est l'accélération de la formation des professeurs et formateurs, qui constituent selon le consortium « le "socle" sans lequel la formation par le numérique ne pourra

jamais décoller » (ibid., p.107). Cette accélération interviendrait dès la formation initiale des professeurs et de façon transversale : « Le digital n'est plus un « module » à part, mais une « brique » qui vient s'insérer dans tous les modules, pour toutes les matières » (ibid.). Elle consisterait d'autre part dans des enseignements liés au numérique et à l'ingénierie pédagogique par le numérique dans la formation continue des professeurs :

> « Étant donnée l'évolution rapide des technologies – le déferlement de ChatGPT en témoigne –, cette formation doit se faire tout au long de la carrière des professeurs. Ces actions vont de pair avec la nécessaire montée en puissance des référents numériques, en renforçant les moyens liés aux opérateurs de formation (Canopé…) » (ibid.).

Selon le collectif, les enseignants ne devraient plus seulement recevoir des formations à l'utilisation des outils numériques, mais pouvoir les intégrer pleinement à leurs pratiques par des compétences d'ingénierie pédagogique. Le rôle d'assistance assumé par le *coach* dans les préconisations de l'ISTE outre-Atlantique serait ici assuré par les référents numériques, notamment au sein de Réseau Canopé.

La proposition suivante est la généralisation du « compte ressources » ou « chèque *EdTech* » à l'ensemble des instituteurs et enseignants du système éducatif et universitaire. Cette mesure s'accorderait avec leur formation préalable :

> « Nous estimons qu'il est urgent de donner le pouvoir d'acheter aux professeurs et de déployer rapidement cette action, à grande échelle. Si les professeurs formés n'ont pas les moyens d'acquérir les solutions une fois convaincus de leur utilité… la formation aura été un coup d'épée dans l'eau » (ibid., p.108).

Partageant avec les acteurs industriels états-uniens l'idéal d'une technologisation additive, la vision commune aux deux syndicats dépasse cependant l'idée de faire des enseignants de simples analystes de données et des facilitateurs au service de l'autonomie des élèves. Il s'agirait de responsabiliser les enseignants en faisant d'eux des acheteurs. Cela fait écho à la marge de manœuvre plus large laissée aux enseignants dans la culture éducative française. Cependant, le dispositif devrait, selon ces syndicats, faire l'objet d'une évaluation étroite, par des « indicateurs d'utilisation effective des plateformes » et des « indicateurs sur l'efficacité en termes d'appropriation des connaissances des élèves ». Si ces prescriptions restent vagues en termes de mise en œuvre, elles reprennent les préceptes du *Social Efficiency Movement* et de l'*Evidence-Based Education* communes aux politiques françaises et états-uniennes.

Les associations n'oublient donc pas la nécessité de rencontrer la vision des Ministères, en soutenant à la fois une technologisation additive et un objectif d'efficience. Une stratégie consensuelle qui semble être payante, étant donnée le statut d'interlocuteurs privilégiés qu'ont acquis ces associations auprès des Ministères.

Mis à part les enseignants, les communautés *EdTech* tournées vers les entrepreneurs diffusent des pratiques professionnelles empreintes de flexibilité, de connaissances des institutions éducatives, ainsi qu'une volonté de réformer l'éducation.

Google, Microsoft ou AWS ont par ailleurs des relais d'influence importants dans l'*EdTech* française, *a fortiori* chez des firmes qu'ils contribuent à outiller. Ainsi Benoît Praly, co-fondateur de Domoscio, a-t-il mis en avant dans le média spécialisé Campus Matin, parmi les avantages d'AWS, la contribution du prestataire à la formation de ses équipes : « AWS assure gratuitement la formation

de nos développeurs et met à disposition des équipes pour nous aider dans nos choix d'architecture » (Campus Matin, 2023).

Cependant, les divers réseaux du numérique éducatif se fédèrent dans des initiatives dirigées en partie contre les stratégies hégémoniques des *Big Tech*. Témoin leurs contributions au projet de base de données européenne « *Data Space Education Skills* ». Celui-ci est motivé, comme on l'a vu, par une volonté européenne de ne pas laisser les multinationales des technologies détenir des gisements de traces potentiellement sensibles, tout en permettant à des petites structures de l'*EdTech* d'entraîner leurs algorithmes sur un vaste répertoire de données ouvertes. Ce projet compte ainsi parmi ses membres fondateurs les deux syndicats du numérique éducatif EdTech France et l'Afinef, de même que celui de l'édition scolaire Les Éditeurs d'éducation (Belkadi, 2022).

Comme aux États-Unis, les communautés du numérique éducatif apparaissent très structurées et dynamiques. Les tendances observées précédemment dans les stratégies des firmes, à savoir l'émergence contrainte des start-ups et un contrôle de plus en plus ténu de la mainmise des *Big Tech*, se retrouvent dans les modalités de constitution de professionnalités. L'essor tumultueux des start-ups est validé par la constitution d'un premier réseau, l'Afinef, puis d'un second, EdTech France, axé sur leur accompagnement et la défense de leurs intérêts dans l'espace public. Les tentatives d'imposition des *Big Tech* dans l'ensemble de l'industrie sont visibles, comme aux États-Unis, dans leurs stratégies de sponsorat et de participation active à ces réseaux, ainsi que la création de leurs propres communautés d'acteurs. Leurs stratégies hégémoniques suscitent toutefois des oppositions parmi les réseaux de l'*EdTech*, défendant la souveraineté numérique et le partage de données entre les firmes. Si l'idée d'assistance du formateur par la technologie est, comme aux États-Unis, consensuelle parmi ces acteurs divers, la responsabilisation de l'enseignant est défendue jusqu'à une fonction d'acheteur.

Conclusion provisoire

Il est temps de revenir sur notre hypothèse esquissée au début de cette partie, selon laquelle les médias adaptatifs feraient advenir une socio-économie dominée par les acteurs des technologies, fédérée autour d'un projet éducatif transversal. Le cas français, décliné sur les trois niveaux interdépendants que sont les politiques publiques, les stratégies des firmes et les professionnalités, a offert une vision plus nuancée des implications socio-économiques de l'*adaptive learning.*

Au niveau des politiques publiques, une volonté de réforme de la forme scolaire s'est construite à l'échelle ministérielle en réponse à un agenda culturel et politique lié à l'*Evidence-Based Education*. Par accumulation d'expérimentations, un projet techno-éducatif nous est apparu de façon latente, et non stabilisée comme dans les discours fédéraux. Celui-ci tend vers une conception socio-constructiviste de l'apprentissage et une conception additive de la technologie. La mise en place de ce projet par l'administration et les établissements éducatifs publics apparaît toutefois erratique, en raison des budgets contraints des établissements et des dissensions autour de l'efficacité des technologies.

Les stratégies des firmes engagées dans l'adaptativité donnent à voir une reconfiguration des positionnements de chaque acteur selon les principaux domaines générateurs de valeur dans l'industrie. Ainsi, le développement d'algorithmes est d'abord assumé par des start-ups jouissant de guichets de financements importants de la part des pouvoirs publics, mais dont la croissance reste

grevée par la faiblesse des marchés et la variabilité des levées de fonds. Les acteurs plus anciens et peu concentrés du numérique éducatif tentent de repositionner leurs recherches dans une acception minimaliste de l'intelligence artificielle, pour répondre aux appels à projets ministériels et capitaliser ainsi sur un ethos innovant. En matière de production de contenus, des alliances entre éditeurs scolaires propriétaires de catalogues certifiés et les structures de production se sont développées, mais les acteurs les plus puissants et innovants tendent à explorer des chemins séparés, notamment via les LLM pour les start-ups. Dans la distribution des médias adaptatifs, l'hégémonie des *Big Tech* tels que Google ou Microsoft est jugulée par les régulations européennes et françaises, de même que la constitution d'une centrale d'accès ministérielle dans le primaire et le secondaire, le GAR, qui assoit la mainmise ministérielle sur ces échelons. La distribution inexistante de médias adaptatifs hexagonaux auprès du consommateur, ainsi que la vigueur de l'intervention étatique dans les marchés institutionnels contribue à la prévalence du modèle du club.

Les professionnalités offrent une nouvelle lumière sur ce décor, où apparaissent des communautés du numérique éducatif structurées et dynamiques, dont témoignent les réseaux comme l'Afinef et EdTech France. Leurs préconisations axées sur la promotion des marchés du numérique éducatif et la formation des enseignants à l'usage des technologies résonnent avec la politique ministérielle, quand leurs postures vis-à-vis des *Big Tech* restent variables.

Nous avons tenté, au cours de ce chapitre, de faire apparaître les similarités et les spécificités qui caractérisaient le déploiement de médias adaptatifs dans les appareils éducatifs et de formation français au regard du cas états-unien. Il est temps, en conclusion de cette partie, de procéder à une synthèse des permanences socio-économiques qui accompagnent l'industrialisation des médias adaptatifs.

Chapitre VII. Des permanences et des spécificités dans les implications socio-économiques de l'adaptativité

Nous synthétisons dans le tableau ci-dessous les divers enseignements tirés de notre analyse contrastive des cas français et états-uniens (Figure 15). Outre les nombreuses différences que nous avons mises au jour entre les deux pays ainsi qu'entre les secteurs éducatifs, il permet de dégager des permanences précieuses au moment de penser les mutations paradigmatiques induites par les médias adaptatifs.

Sur le plan des politiques publiques, ces technologies s'inscrivent dans des programmes volontaristes des pouvoirs publics, consistant à placer le numérique au centre des pratiques pédagogiques. Leur perspective est celle du *Social Efficiency Movement*, visant à transformer de façon efficiente les systèmes éducatifs, en prônant un idéal de performance individuelle. À travers la technologie, ils encouragent d'une part une différenciation de l'apprentissage au service des besoins d'apprenants de tout âge et de tout parcours. Ils prônent d'autre part une ouverture des frontières spatio-temporelles des organisations éducatives pour impulser des pratiques collaboratives, autonomes et créatives chez les apprenants, évalués de façon formative dans une veine socio-constructiviste. Ils défendent la figure de l'enseignant, qui devrait voir ses pratiques pédagogiques « augmentées » par l'analyse et la visualisation automatiques des apprentissages numériques. L'administration des établissements doit selon les pouvoirs publics échoir aux acteurs éducatifs « de terrain », dont des professions spécialisées liées à l'ingénierie pédagogique, et s'ouvrir aux firmes innovantes. La vocation démocratique de ces technologies est palpable dans le soutien des acteurs publics aux ressources éducatives libres, de même que leur attention au respect des libertés individuelles et notamment la confidentialité des données numériques traitées.

Le projet techno-éducatif soutenu par les pouvoirs publics n'est donc pas un paradigme concurrent de la forme scolaire, mais une tentative d'amendement et de flexibilisation des modes pédagogiques existants. La marge de manœuvre budgétaire contrainte d'un grand nombre d'établissements, de même que les doutes subsistant sur l'utilité pédagogique des technologies numériques tendent cependant à rendre l'intervention des acteurs privés cruciale dans la réalisation de ce projet.

			États-Unis	France
Politiques publiques	Degré d'intervention dans les industries éducatives		Rôles prescriptifs et financiers des échelons fédéraux et fédérés, en soutien de la recherche et de l'équipement des établissements ; autonomie laissée aux établissements sur le plan opérationnel	Centralisation forte du système éducatif et de formation autour des Ministères en termes financiers, prescriptifs et logistiques. Autonomie laissée aux universités et employeurs sur le plan opérationnel.
	Projet éducatif entourant l'*adaptive learning*	Degré de formalisation	Avancé	Modéré
		Nom	*Personalized learning*	N/D
		Apprentissage	Vision socio-constructiviste autour de l'"expérience" individuelle	Focale culturelle et utilitariste d'autonomie des individus, par des apprentissages démultipliés et collaboratifs
		Enseignement	Technologisation additive augmentant les capacités analytiques et créatives de l'enseignant	Technologisation additive augmentant les capacités analytiques de l'enseignant et préservant sa liberté pédagogique
		Évaluation	Évaluation formative bénéficiant à la performance des divers acteurs éducatifs	Évaluation formative et potentialités métacognitives, évaluation sommative des compétences plutôt que des connaissances aux niveaux supérieurs
		Gouvernance	Territorialisation forte et contrôle par les personnels administratifs et éducatifs (*IA centrée sur l'humain*)	Territoralisation, collaboration des publics par projet et ouverture vers les acteurs socio-économiques, dans une perspective de performance. Professionnalisation des ingénieurs pédagogiques.
		Infrastructure	Éclatement des clôtures spatio-temporelles par le numérique et des inégalités d'accès par les REL ; précautions éthiques quant à la confidentialité des données et techniques sur la pertinence pédagogique de l'IA	Conception avec les acteurs des REL et contrôle éthique des dispositifs (confidentialité des données, souveraineté, interopérabilité des outils et « éco-responsabilité »)
	Intensité de la mise en place par les établissements		Faibles investissements financiers et déficit d'informations des établissements	Faible marge de manœuvre financière des établissements publics et doutes suscités par l'instrumentalisation de la recherche scientifique
Stratégie des firmes	Conception des algorithmes	Stratégies des start-ups	Investissement majoritaire des ressources dans la R&D grâce aux financements de *venture capital,* dons de fondations et aides fédérales, sous-traitance de *cloud computing* voire d'algorithmes adaptatifs auprès des *Big Tech*	Investissement majoritaire des ressources dans la R&D grâce aux crédits d'impôt recherche, aux financements et marchés publics ainsi qu'aux fonds de capital-risque ; transferts de technologies, recrutement et sous-traitance de *cloud computing* auprès des *Big Tech*
		Stratégies des acteurs anciens du numérique éducatif	Acquisition de start-ups par les groupes concentrés	Repositionnement à la faveur de marchés publics et de partenariats avec des laboratoires sur des IA symboliques, et capitalisation sur un ethos innovant pour entamer une financiarisation, par fusions-acquisitions ou levées de fonds. Fragilisation financière
	Conception des contenus de remédiation	Tendances fédératives	Sous-traitance ou rachat de start-ups par les éditeurs scolaires pour leurs ITS	Sous-traitance de module ou d'ITS entiers par les éditeurs scolaires. Consortiums entre éditeurs et start-ups lors d'appels d'offres ministériels
		Tendances à l'éclatement	Recrutement d'auteurs, exploitation de REL ou capitalisation sur les LLM par les start-ups	Concurrence entre éditeurs et start-ups lors d'appels d'offres ministériels. Edition d'ITS, recrutement d'auteurs et recours à l'IA générative par des start-ups, conception de médias éducatifs simplistes par les éditeurs scolaires
	Distribution des médias adaptatifs…	…auprès des institutions et employeurs	Modèle du club via licences perçues par les éditeurs ; modèle du courtage via des intermédiations croissantes d'acteurs privés	Modèle du club majoritaire via des licences perçues par les éditeurs ; intervention forte des pouvoirs publics et intermédiation du GAR
		auprès du consommateur	Modèle du courtage via les magasins d'applications d'Apple et Google percevant des redevances sur les abonnements ou achats intégrés	Modèle du courtage via les magasins d'applications d'Apple et Google percevant des redevances sur les abonnements ou achats intégrés
Profession-nalités	Typologie des réseaux d'acteurs		Associations prescriptrices de standards, incubateurs de start-ups privés ou liés aux *Big Tech*	Syndicats d'acteurs du numérique éducatif et incubateurs publics, associatifs, ou privés
	Compétences prescrites		Compétences pédagogiques liées au *personalized learning,* tentatives de codification des pratiques communicationnelles par les *Big Tech*	Compétences liées à la *digital literacy* pour les professionnels éducatifs et à l'ingénierie pédagogique pour les enseignants, allant jusqu'à leur donner une responsabilité dans l'achat d'outils

Figure 15. Tableau récapitulatif de l'analyse contrastive des socio-économies états-unienne et française des médias adaptatifs.

Il apparaît sur ce plan que le développement de médias adaptatifs forme une industrie autour de trois domaines générateurs de valeur : la conception d'algorithmes d'apprentissage adaptatif, la production de contenus pédagogiques et la distribution des médias adaptatifs. Nous avons ci-après synthétisé sous forme de chaînes de valeur la morphologie de ces filières aux États-Unis et en France, et les rapports de force entretenus par les divers acteurs (Figure 16). Les spécificités apparaissent rapidement à différents endroits :

- D'abord, en amont, le dynamisme entrepreneurial et capitalistique des États-Unis engendre des relations plus denses entre les parties prenantes (start-ups, acteurs anciens du numérique éducatif, éditeurs scolaires et organismes de formation) qu'en France.
- En aval, à la privatisation de la fonction d'intermédiation et à la montée du modèle du courtage dans le circuit institutionnel états-unien s'opposent une présence plus forte des pouvoirs publics et un maintien du modèle du club du côté français.

Néanmoins des permanences apparaissent tant sur ces frises que dans le tableau contrastif. La typologie des acteurs présents et leur positionnement général est identique : des start-ups prennent en charge la majeure partie des recherches algorithmiques, avec le concours régulier de laboratoires, et bénéficient des infrastructures *cloud* des *Big Tech,* voire de leurs modèles de langage et autres briques algorithmiques. Les acteurs anciens du numérique éducatif adoptent des positions volontaristes en se repositionnant sur ce segment, de façon plus ou moins ambitieuse selon leurs ressources. De même, des éditeurs scolaires et organismes de formation se positionnent en éditant leurs LMS, capitalisant sur leurs catalogues de contenus, avec le concours plus ou moins prolongé de start-ups, et des ambitions adaptatives variables. Du côté de la distribution, des plateformes d'intermédiation sont devenues des acteurs incontournables de l'accès aux publics éducatifs – étape cruciale dans cette filière, tant du point de vue technique que commercial. Les *Big Tech* sont quant à eux présents sur les principaux domaines générateurs de valeur, éditant leurs propres plateformes en aval et sous-traitant leurs technologies en amont, sous le contrôle accru des pouvoirs publics inquiets de leurs pratiques hégémoniques. Trois tendances transparaissent dans l'évolution des rapports de force industriels*,* permettant à certains acteurs d'influer plus que d'autres sur la définition d'un paradigme : l'importance des start-ups *EdTech* en tant que moteurs d'innovations pour les offres adaptatives, l'importance croissante des plateformes de distribution, qu'elles soient privées ou publiques, et les stratégies hégémoniques mais de plus en plus régulées des *Big Tech*. Se joue en définitive une lutte entre les éditeurs concentrés et les acteurs de l'intermédiation pour s'installer au centre d'une nouvelle socio-économie – pour les uns à travers le modèle du club, et pour les autres via le modèle du courtage.

Enfin, la constitution de professionnalités transverses, c'est-à-dire de compétences et d'intérêts partagés par des professionnels variés, est pleinement à l'œuvre autour du mouvement de l'*EdTech*, auquel s'identifient les parties prenantes de l'*adaptive learning.* Des associations à but non-lucratif défendant les intérêts de la filière et des incubateurs soutenant les start-ups constituent tous ensemble des agents très importants pour la rencontre de professionnels variés intéressés par ces technologies nouvelles, des professionnels de l'éducation aux acteurs industriels. Ces réseaux confortent voire influencent les projets technopédagogiques des pouvoirs publics, en se focalisant sur la technologisation additive et l'assistance à l'enseignant. Des compétences semblables de compréhension du fonctionnement des outils, en particulier d'intelligence artificielle, et de maniement de leurs potentialités pédagogiques sont au cœur de leurs préconisations. S'y jouent également des enjeux d'influence et de codification des professionnalités par certains acteurs

– à l'image des *Big Tech*, très engagés dans ces communautés d'acteurs –, qui donnent lieu à des dissensions.

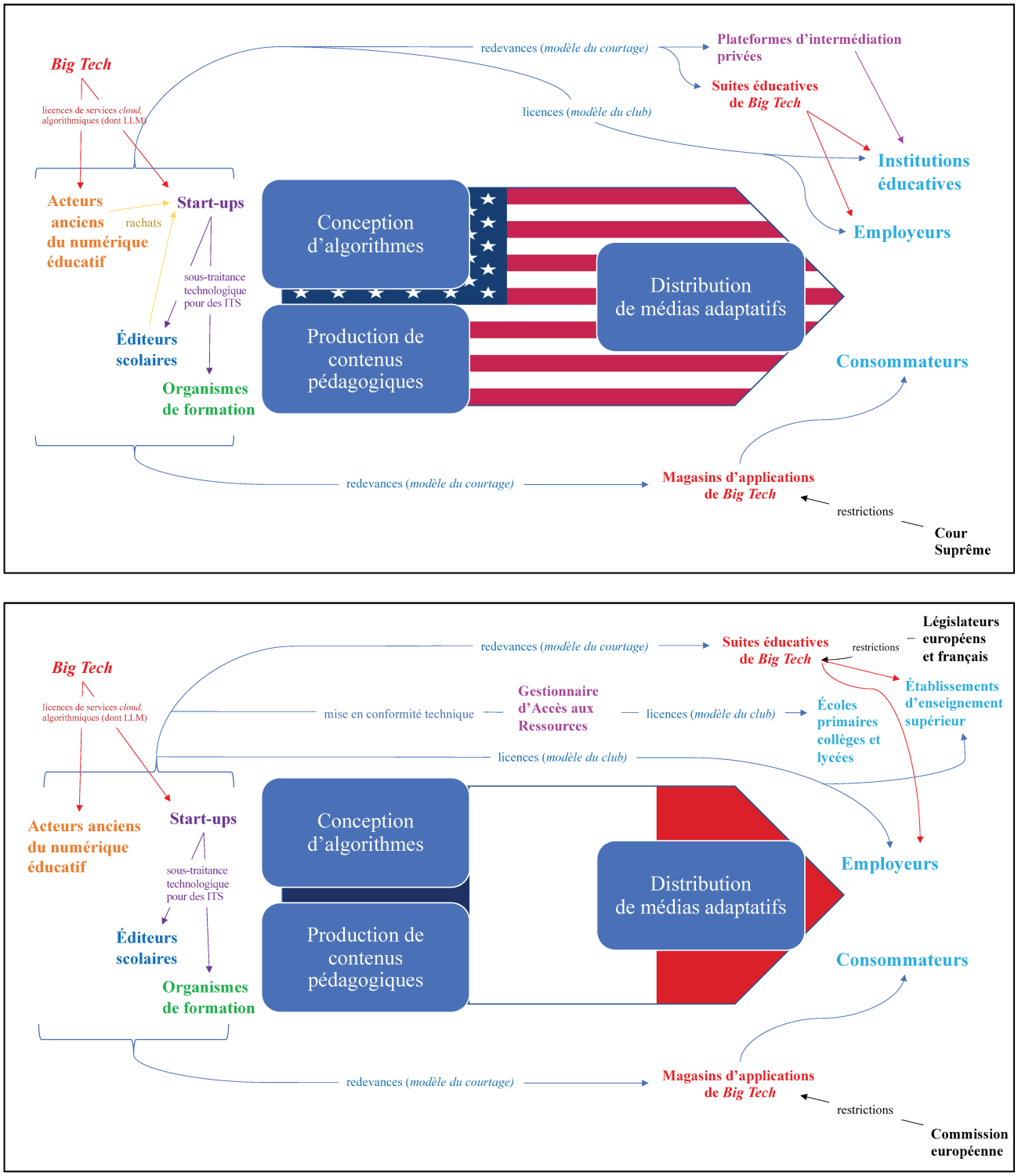


Figure 16. Visualisation des filières états-unienne et française des médias adaptatifs.

Si l'on reprend notre problématique de départ, qui est de déterminer la contribution des médias adaptatifs à un renouvellement paradigmatique au sein des systèmes éducatifs et de formation, les acquis glanés jusqu'ici sont nombreux. Nous avons observé en première partie la difficulté des chercheurs états-uniens et français à impulser une vision scientifique cohérente de l'adaptativité. Plutôt que de les rallier derrière un projet éducatif commun, ce chantier technologique a suscité des

aspirations diverses – de la tendance behavioriste issue des « machines à enseigner », focalisées sur l'enseignement, aux inclinations constructivistes des premiers programmes informatiques, basés sur le modèle du tutorat, aux formes d'apprentissage socio-constructivistes et connectivistes propres aux environnements informatiques d'apprentissage humain. Le bouleversement scientifique entraîné par l'IA connexionniste et la fouille de données a finalement relancé le débat central entre techno et anthropocentrisme.

Cette deuxième partie a quant à elle mis en lumière, sur les trois niveaux de décision que sont les pouvoirs publics, les firmes et les professionnalités, des stratégies globalement volontaristes en matière d'adaptativité. Ces trois focales donnent à voir des soutiens divers au développement d'une industrie des médias adaptatifs, animée par un projet techno-pédagogique consistant à rendre, par la technologie, les appareils d'éducation et de formation plus flexibles et collaboratifs, et les publics éducatifs plus autonomes et mieux outillés dans leurs pratiques. Ces discours s'apparentent à l'un des marqueurs de l'industrialisation cernés par Pierre Mœglin, l'idéologisation, définie en ces termes par le chercheur :

> « La notion d'idéologisation désigne […] le processus par lequel, se parant de justifications du type de celles que leur fournissent la religion du progrès et de la modernité, la croyance en l'ordre irréversible de la technique et d'autres arguments semblables, des acteurs rationalisent la rationalité de leurs stratégies. Autrement dit, ils justifient le volontarisme de leurs stratégies modernisatrices et se convainquent de la légitimité de la priorité qu'ils accordent à l'optimisation des moyens par rapport aux fins, au détriment de toute autre considération » (Mœglin, 2016, p.45).

L'on a observé dans les discours des acteurs socio-économiques cette croyance dans le progrès et la modernité, augmentée d'une aura démocratique, dans la mesure où les médias adaptatifs devaient permettre à tous d'accéder à des « expériences » d'apprentissage et d'enseignement plus proches de leurs besoins et aspirations.

L'idéologisation résonne avec un autre marqueur de l'industrialisation, la rationalisation. Celle-ci désigne l'application d'une « rationalité en finalité » weberienne dans l'éducation, soutenant une organisation rationnelle du travail éducatif conforme à des idéaux d'efficacité et de rendement économique. On retrouve ainsi la quête d'optimisation des moyens par rapport aux fins que l'idéolologisation se charge de rendre désirable. Dans notre partie, cette rationalisation est apparue dans les discours politiques et marchands emprunts des préceptes du *Social Efficiency Movement* et de l'*Evidence-Based Education*, consistant à rendre l'éducation plus efficiente notamment en termes de coûts d'équipements.

Idéologisation et rationalisation résonnent enfin avec un troisième marqueur, la technologisation, qui indique à la fois « la présence et l'utilisation de dispositifs techniques » et prend en compte la « dimension des usages qu'ils prescrivent et des pratiques qu'ils modèlent » (Mœglin, 2016, p.41). La technologisation additive prônée par les acteurs publics, les communautés *EdTech* ou les *Big Tech* ont donné à voir, à travers les technologies numériques, des habitus d'apprenants entrepreneurs de leurs apprentissages, et à travers l'analytique, des pratiques gestionnaires chez les élèves et les enseignants.

Toutefois, l'analyse des discours programmatiques des acteurs socio-économiques apparaît insuffisante pour caractériser un processus d'industrialisation en lien avec les médias adaptatifs. D'une part, on a vu que de nombreux projets éducatifs, notamment formulés par les pouvoirs publics,

restaient superficiels et vagues en termes d'acceptions technologiques, ce qui affecte en l'état notre quête d'un projet spécifique à l'adaptativité. D'autre part, l'*adaptive learning,* à travers les promesses marchandes qui l'entourent, s'apparente à ce que Gaston Bachelard nommait une « expression obstacle » (Bachelard, 1938) : il absorbe le récepteur dans un imaginaire « marketing » lié à la personnalisation, empêchant de comprendre le fonctionnement des médias adaptatifs et leurs implications concrètes sur la pédagogie.

Ainsi, l'industrialisation doit être observée non seulement dans les discours mais également au niveau pragmatique des médias éducatifs et des normes d'usage qu'ils conduisent. L'étude d'un renouvellement paradigmatique nécessite donc de se focaliser sur un corpus de médias adaptatifs issus des deux pays et visant divers niveaux éducatifs. Il s'agit d'étudier à la fois le fonctionnement de leurs algorithmes, leurs modes d'existence « à l'écran », et les discours promotionnels qui les accompagnent, pour tirer de notre méthode contrastive de nouvelles permanences. Nous nous efforçons dans la partie suivante d'étudier ces trois formes de concrétisation des stratégies d'éditeurs. Ceci à la lumière de l'histoire scientifique tumultueuse de l'adaptativité telle qu'on l'a observée en première partie, et des positions et calculs socio-économiques qui animent les acteurs étudiés dans la deuxième partie.

TROISIÈME PARTIE. Les modalités d'industrialisation de l'éducation à l'œuvre dans les médias adaptatifs : une convergence vers les industries de la communication ?

Les contours d'un corpus de médias adaptatifs

Nous étudions la réalisation de ces stratégies à partir d'un corpus de sept EIAH visibles dans le tableau ci-dessous (Figure 17). Ceux-ci ont en commun les trois critères retenus jusqu'ici pour l'adaptativité : la médiatisation, la captation des besoins individuels en continu, et la flexibilité de la pédagogie.

	Maternelle, école primaire : apprentissage de la lecture	Ecole primaire, collège, lycée : apprentissage des mathématiques	Enseignement supérieur	Formation Professionnelle	Auto-formation : apprentissage des langues
France	*Lalilo*	*Adaptiv'Math*		*Domoscio Hub*	
Etats-Unis	*Scholatic FIRST, Lalilo*	*MATHia*	*Knewton Alta*		*Duolingo*

Figure 17. Acteurs du corpus rangés par modèle d'affaires : en rouge, par abonnement ; en orange, en freemium *incluant de la publicité personnalisée.*

Ils se distinguent en revanche sur de très nombreux aspects. Le premier d'entre eux est la nationalité des firmes éditrices : françaises pour Lalilo, Evidence B (*Adaptiv'Math*) et Domoscio ; états-uniennes pour Scholastic FIRST, Carnegie Learning (*MATHia*), Knewton, Duolingo (et en partie Lalilo depuis son rachat par Renaissance Learning).

Le deuxième facteur différenciant est le profil économique, d'une firme ayant ouvert son capital à des investisseurs privés (Evidence B), à des filiales d'éditeurs (Scholastic FIRST, Knewton) ou de groupes du *e-learning* (Lalilo, Domoscio), en passant par une firme cotée en bourse (Duolingo).

Un troisième aspect distinctif est le niveau d'enseignement visé, de l'alphabétisation de jeunes enfants à l'auto-formation en langues, certaines visant plusieurs échelons – à l'image de Duolingo, qui a lancé une application scolaire (*Duolingo for Schools*) en plus de son service d'apprentissage des langues.

Un quatrième facteur est le modèle d'affaires, figuré par un code couleur dans notre tableau : toutes les firmes optent exclusivement pour un modèle par abonnement ou par licence (en rouge), à l'exception de Duolingo (en orange), qui opte pour un mode *freemium* incluant des publicités personnalisées.

Un cinquième et dernier facteur distinctif est la filière industrielle. Celle-ci tient aux usages pédagogiques prévus par les firmes pour leurs EIAH. Certaines projettent des usages hybrides, en classe et à la maison, les affiliant au *e-learning* (*Lalilo, Domoscio Hub*). Pour d'autres prévalent des usages en classe inspirés de l'informatique pédagogique (*Scholastic FIRST, Adaptiv'Math, MATHia*). D'autres encore promeuvent d'abord des usages à la maison, ce qui les fait pencher vers la formation à distance (*Knewton Alta*). Duolingo vise quant à lui des usages ubiquitaires propres aux

produits pour l'éducation non-formelle. Étant donnés tous ces facteurs contrastifs, nous comptons analyser les stratégies industrielles adoptées par chaque acteur, pour tenter de dégager des permanences significatives.

Une méthodologie partagée entre deux théories

Pour étudier les stratégies embarquées dans ces médias, nous nous tenons au croisement de la théorie de l'industrialisation de la formation explorée par Pierre Mœglin et les chercheurs du SIF puis du GIS2IF, et de celle de la *trivialité*. Ce concept directeur de l'œuvre d'Yves Jeanneret désigne « le caractère transformateur et créatif de la transmission et de la réécriture des êtres culturels à travers différents espaces sociaux » (Jeanneret, 2014).

Dans la quatrième partie de *Critique de la trivialité,* intitulée « Le capitalisme médiatique. Sémiologie des médias et économie politique de la communication », Yves Jeanneret opère un dialogue avec les théories des industries culturelles, lui permettant de penser « la relation dialectique entre conditions de possibilité de la communication et modèles d'économie politique » (Jeanneret, 2014, p.50). Pierre Mœglin appelle également à ce dialogue, considérant que « les questions d'industries culturelles ne se posent pas seulement au niveau des grands modèles socio-économiques et des concurrences que se font mutuellement les acteurs qui les pratiquent. Ces questions sont aussi celles des formes concrètes de la communication, car elles intéressent le stade infra-ordinaire des signes naturalisés et des médiations quotidiennes » (Mœglin, 2015).

En l'occurrence, le lien projeté par des acteurs industriels entre une doctrine pédagogique visant l'efficience et la conception d'un média est symptomatique de ce qu'Aude Seurrat nomme une « conception médiagénique de la médiation ». Empruntant à Philippe Marion le concept de *médiagénie*[109], elle désigne ainsi « une certaine conception de la médiation qui verrait dans l'usage de telle ou telle forme de médiation […] les garanties du succès » en matière communicationnelle. Ce concept sert ainsi de « catégorie pour penser des conceptions de l'opérativité sociale (présentée comme des effets garantis) de dispositifs de médiation des savoirs » (Seurrat, 2018, p.35). Ce positionnement théorique poursuit le dialogue opéré par Yves Jeanneret entre *promesse* et *implication* au sein des médias : « La promesse consiste à rendre publique une définition de l'acte de communication et de sa visée ; l'implication consiste à créer un dispositif d'objets et de discours qui incarne une telle prétention et surtout convoque les sujets de communication dans une posture, une situation, un contexte d'expression » (Jeanneret, 2014, p.66).

Cette théorie rencontre sur ce point celle de l'industrialisation de la formation et ses trois marqueurs (technologisation, rationalisation, idéologisation), dans la mesure où ceux-ci portent également sur les modalités de codification de l'éducation et d'organisation des rôles de chaque partie prenante, ainsi que sur la circulation d'utopies technologiques parmi les acteurs.

Nous appliquons donc aux médias adaptatifs composant notre corpus une grille de lecture alliant des outils conceptuels issus des deux théories. Pour commencer, l'industrialisation de l'éducation doit nous permettre de cerner les dimensions sur lesquelles le projet industriel est formalisé, et celles sur lesquelles il est, au contraire, moins patent et laisse des espaces de liberté aux usagers.

[109] Philippe Marion désigne ainsi « l'évaluation d'une "amplitude" : celle de la réaction manifestant la fusion plus ou moins réussie d'une narration avec sa médiatisation, et ce dans le contexte – interagissant lui aussi – des horizons d'attente d'un genre donné ». (Marion, 1997)

Saisir les marqueurs d'industrialisation

La caractérisation de chacun de ses marqueurs doit s'opérer avec une méthodologie propre. La technologisation est d'une part jaugée dans la nature des algorithmes utilisés et le type d'adaptativité qu'ils introduisent dans les EIAH. Pour identifier ces aspects, nous mobilisons la triade de l'adaptativité proposée par Edsurge (2016) et présentée en introduction. Pour rappel :

- Un premier mode d'intervention, l'*adaptive content,* ayant émergé avec les premiers « tuteurs intelligents » des années 1970, concerne les algorithmes opérant à l'intérieur de l'exercice, à la suite d'une erreur commise par l'élève, par le choix d'un support de remédiation adapté à ses besoins : découpage de l'exercice en micro-actions, affichage d'une vidéo, d'une précision grammaticale par exemple.
- Au sein d'une évaluation, des algorithmes peuvent également adapter la difficulté d'une question en fonction du résultat obtenu aux précédentes. C'est l'*adaptive assessment*, apparu avec les développements du « *computer adaptive testing »* (CAT) dans les années 1970.
- Le troisième mode d'intervention, l'*adaptive sequencing,* se situe au niveau plus vaste du programme. Ces algorithmes ayant émergé plus récemment, à partir des années 1990, ré-agencent le parcours de l'apprenant en modifiant la chronologie des notions et formats, pour s'adapter au niveau de l'apprenant – là où les deux modes précédents conservent un programme unique pour tous.

Pour jauger ces niveaux d'adaptativité, nous mobilisons les publications scientifiques des ingénieurs de ces firmes. En complément, ou lorsque de telles publications n'étaient pas disponibles, nous avons conduit des entretiens semi-directifs avec ces ingénieurs.

Toujours sous le sceau de la technologisation, la capacité de ces dispositifs à modeler des habitus, ou « systèmes de disposition à la pratique » (Bourdieu, 1987, p.94) chez les utilisateurs doit d'autre part être jugée à travers une analyse techno-sémiotique des « expériences-utilisateur ». Nous empruntons cette méthode à Yves Jeanneret et Emmanuël Souchier qui ont ainsi étudié le rôle structurant de la technique sur l'énonciation au sein des « médias informatisés ». Ceci à travers ce qu'ils nommèrent des « architextes » (Jeanneret et Souchier, 1999), c'est-à-dire des « outils singuliers » qui « permettent l'effectuation de toute activité médiée par un dispositif informatisé », de type logiciel, selon les mots d'Emmanuël Souchier (Souchier, 2012, p.90). Le chercheur souligne trois dimensions sémiotiques des architextes que l'on peut appliquer aux médias adaptatifs :

> « La première est liée au fait que la pratique d'écriture est à proprement parler *mise en abîme* dans le dispositif industriel d'écriture qu'est le logiciel : *j'écris dans un outil d'écriture*. La deuxième souligne l'*interdépendance* entre le dispositif et le processus d'écriture : *j'écris dans un outil d'écriture* qui est lui-même *écrit* ; autrement dit, *j'écris dans un outil d'écriture écrit*. En troisième lieu, […] le dispositif comprend nécessairement une représentation de l'activité à laquelle il est dédié en ce qu'il est une écriture de l'écriture. En d'autres termes, qu'un *architexte* comprend et se constitue nécessairement à partir des représentations que ses concepteurs se font du processus de communication, que ces représentations soient d'ordre imaginaire, idéologique ou théorique, par exemple. Un *architexte* est donc fatalement une mise en pratique industrielle d'une théorie communicationnelle, fût-elle implicite » (ibid.).

Les médias adaptatifs sont en effet, au premier niveau, des environnements d'écriture pour un apprenant, dans le cadre d'exercices. Au deuxième niveau, ils sont préconçus par des professionnels à la fois en *front-end,* via l'agencement d'éléments et de formes à l'écran, et en *back-end*, par l'écriture du code informatique et l'élaboration d'algorithmes chargés d'agencer

les éléments pédagogiques à l'écran. Au troisième plan, le média véhicule de façon industrielle la manière qu'a la firme éditrice, prise collectivement, de représenter l'apprentissage, en plaçant l'apprenant ou le formateur dans certains rôles – que cette représentation soit explicitée ou non, et imaginaire, idéologique ou théorique. La méthode techno-sémiotique vise à expliciter ces dimensions signifiantes que recèlent les « médias informatisés ».

En *back-end*, les algorithmes d'apprentissage adaptatif structurent la pédagogie en ligne mais interviennent de façon sous-jacente et personnalisée. Impossible, par une analyse techno-sémiotique portant sur les activités à l'écran d'un usager, de relever leurs interventions. Ils agissent davantage comme des « computextes », notion qu'a proposée Alexandra Saemmer pour les « processus de calcul qui ne visent plus seulement à formater l'apparence du texte, mais le contenu lui-même » (Saemmer, 2012, p.102), tels que les algorithmes de génération automatique de texte :

> « J'inclus dans les computextes à la fois les algorithmes qui effectuent des calculs et pondérations dans les traces laissées par les utilisateur.e.s, et les traces elles-mêmes, organisées en bases de données. Comme les architextes, les computextes cadrent et guident le processus d'écriture ; cependant, l'utilisateur.e n'écrit plus seulement *dans* ces outils, elle ou il écrit littéralement *avec* eux. Mieux encore, ces outils écrivent *à la place de l'utilisateur.e* au moment même où elle ou il écrit, sur la base de calculs effectués en direct à partir des traces passées. Les computextes se caractérisent ainsi par le fait de ne plus faire les mêmes propositions à tout le monde (comme c'était le cas pour les architextes), mais de faire émerger des propositions personnalisées » (ibid.).

Les algorithmes d'apprentissage adaptatif se distinguent des systèmes d'auto-complétion des *Big Tech* investiguées par la chercheuse, car ils ne visent pas à générer du texte. Toutefois, ils les rejoignent sur une acception large de l'écriture, dans la mesure où ils conçoivent des parcours pédagogiques à des niveaux de détail différents (de l'indice au programme entier). Si l'on prend le point de vue de l'enseignant qui doit *écrire* (au sens de *concevoir*) des activités personnalisées pour ses élèves, ces algorithmes sont effectivement susceptibles d'intervenir à sa place en opérant des calculs selon les données d'apprentissage de l'élève.

Pour saisir les conceptions pédagogiques véhiculées par les médias adaptatifs, en particulier dans l'articulation entre leurs architextes et leurs computextes, il convient donc de lier l'analyse des publications et verbatim d'ingénieurs, permettant de comprendre le fonctionnement des algorithmes d'apprentissage adaptatif, à l'analyse techno-sémiotique des interfaces. L'« observation participante médiatisée » (Bullich, 2018) semble être la démarche adéquate pour réaliser cette seconde analyse. Néanmoins, l'inscription sur les environnements adaptatifs institutionnels de notre corpus présuppose l'appartenance à une communauté éducative fermée (une classe, une promotion universitaire ou une équipe d'employés) à laquelle nous ne pouvions prétendre. Nous avons donc adopté cette démarche pour le seul outil d'auto-formation de notre corpus, *Duolingo* (dans ses versions gratuite et payante). Pour les autres dispositifs, nous avons contourné cette difficulté en procédant à une étude croisée de diverses ressources documentaires publiées par les firmes ou par des usagers, donnant à voir des utilisations des outils du point des élèves et des formateurs.

La rationalisation, autre marqueur d'industrialisation indiquant la tendance à organiser les EIAH selon une quête d'efficience et de rentabilité, est en partie relayée par la technologisation. Elle doit dès lors affleurer dans l'étude des algorithmes en *back-end* et l'analyse techno-sémiotique des interfaces décrites ci-dessus – par exemple dans les variables prises en compte par les algorithmes, ou dans les métriques affichées dans les tableaux de bord des utilisateurs. Elle doit également

transparaître dans l'analyse de discours explicatifs dispensés auprès du grand public sur les sites web et les blogs des firmes.

L'idéologisation, marquant la tendance des acteurs à admettre et rationaliser la rationalisation opérée dans les EIAH, doit être jaugée par la déconstruction sémio-discursive des ressources publiées par les entreprises à destination des divers publics d'usagers. Il s'agit de rechercher les argumentaires et représentations sociales sur lesquels ces firmes s'appuient, et notamment les « promesses socio-techniques » qu'elles formulent. Ce concept de la sociologie des techniques proposé par Pierre-Benoît Joly désigne les promesses faites en réponse à des problèmes sociétaux, constituant des modalités de « construction collective du futur » et introduisant des objets techniques et des pratiques normatives (Joly, 2010). Cette notion nous semble particulièrement pertinente pour appréhender les représentations sociales entourant l'*adaptive learning*, liées à des problématiques comme le décrochage scolaire.

Vers un nouveau mode d'organisation industriel

Cette méthodologie va nous permettre de mettre au jour, dans un premier chapitre, un *modus operandi* industriel perceptible à des degrés différents dans l'ensemble du corpus. Cet idéal-type, que nous nommons « industrialisation *par l'engagement* », consiste dans un mode d'organisation commun aux médias éducatifs présentant des algorithmes d'apprentissage adaptatif. En partie esquissé lors de notre analyse des emplois de l'IA chez Duolingo (Boët, 2023), ce mode est ici détaillé selon les trois marqueurs de l'industrialisation. Premièrement, sur le plan de la rationalisation, il se distingue par la poursuite de l'*engagement* de l'utilisateur comme indicateur de l'efficacité pédagogique et économique. Deuxièmement, il se concrétise dans les *learning experiences* en *front-end* et les algorithmes d'apprentissage adaptatif en *back-end*. Enfin, il s'appuie, en guise d'idéologisation, sur des récits faisant de la personnalisation le vecteur d'une éducation efficiente.

Nous interprétons ce *modus operandi* comme une énième étape de la convergence entre industries éducatives et culturelles. Nous avons évoqué précédemment les nombreux travaux de chercheurs renseignant des tendances anciennes au rapprochement des deux industries. Dans un article de 2018, Vincent Bullich a cependant démontré que « le mouvement […] s'amplifi[ait] sous l'effet de la « pression numérique ». Dans son étude des plateformes de MOOC, il a d'une part mis au jour la plateforme comme une « [forme] de production "distribuée", conjuguant intermédiation en ligne et activités de production hors ligne, à même de se substituer progressivement à des formes plus intégrées d'organisations dont les parangons sont la firme et […] l'université (ou tout établissement d'enseignement) ». Il a d'autre part souligné l'application de la plateformisation aux médias à la fois éducatifs et culturels, selon des modalités et finalités communes[110]. En conclusion, l'auteur gageait « que cette pression dev[ait] s'accentuer à l'avenir, et que de nouveaux modèles d'organisation ne manquer[aient] de s'adjoindre à celui des plateformes » (Bullich, 2018).

Les médias adaptatifs semblent suivre un mode d'organisation similaire à la plateformisation sur plusieurs points : mobilisation des individus par des métriques, promesse de personnalisation de la formation, mise en activité des usagers pour multiplier les traces récoltées… Cependant des spécificités fortes apparaissent : ces EIAH ne sont pas des « places de marché » visant à mettre

[110] Parmi les finalités, il y a l'obtention de gains de productivité, et parmi ces modalités, des activités « sur mesure » grâce aux algorithmes de recommandation, l'inclusion de l'usager dans une « communauté », des opérateurs de mise en activité, ou la « réquisition » par les métriques (Bullich, 2018).

en relation des internautes, mais des systèmes conservant l'apprenant dans un parcours où se conjuguent travail humain et adaptativité algorithmique. En découlent une structure de coûts et des stratégies économiques très différentes. Vincent Bullich distingue en outre la *plateforme* d'un média traditionnel, dans la mesure où son gestionnaire « ne paie pas pour l'obtention des contenus/produits/services qu'il propose » (Bullich, 2018). Les EIAH adaptatifs sont quant à eux bel et bien des médias, qui tirent la majeure partie de leurs revenus des contenus qu'ils éditent, via des abonnements ou licences. Ces spécificités nous poussent donc à nous intéresser à l'industrialisation *par l'engagement* en tant que nouveau mode d'organisation.

Nous observons dans un second chapitre les facteurs socio-économiques qui s'opposent à la diffusion de ce *modus operandi* parmi les acteurs économiques. Nous observons la présence de modèles alternatifs favorisant, selon les cultures et les échelons éducatifs, d'autres formes industrielles, sinon des espaces d'artisanat et de créativité pour les publics éducatifs.

Chapitre VIII. L'industrialisation *par l'engagement*, un idéal-type pour la conception de médias hybrides

L'industrialisation *par l'engagement* peut être dégagée de l'analyse de notre corpus, au-delà des spécificités des filières éducatives, des niveaux d'enseignement et des cultures éducatives françaises et états-uniennes. Il s'agit dans ce chapitre de la définir sur les trois marqueurs d'industrialisation, grâce à des exemples empruntés aux divers médias composant notre corpus.

Comme son nom l'indique, ce modèle consiste dans une poursuite quantitative de l'*engagement* de l'utilisateur. Celui-ci est traduit quantitativement par des indicateurs que les firmes cherchent à maximiser dans une tendance à la rationalisation. Il porte dans son étymologie et ses évolutions sémantiques une ambivalence pleinement exploitée par les firmes du secteur.

1. L'engagement de l'utilisateur, indicateur de l'efficience des médias adaptatifs

Selon le Dictionnaire historique de la langue française d'Alain Rey (2016), *engagement* prend sa racine dans *gage*, ce qui lui a valu dès le XII^e^ siècle une acception juridique, en tant qu'« action de mettre (quelque chose) en gage », lier par un contrat, puis de se lier par une promesse ». Il a pris au XVI^e^ siècle un sens figuré, topologique, en tant que « fait d'entrer dans un passage étroit ». Ces deux valeurs sont unies par cet état final du sujet, pris dans « une situation qui ne laisse pas libre » de ses choix ou de ses mouvements.

Maude Bonenfant et Thibault Philippette constatent dans un article de 2018 que la motivation de l'individu à s'engager « sous-entend une forme d'investissement matériel et affectif ». La circulation du terme au XX^e^ siècle l'a ainsi amené vers le domaine du marketing : « Dans [ce] domaine par exemple, l'engagement est une notion commune qui réfère à la loyauté du consommateur (investissement affectif) envers une marque ou un produit pour s'assurer de l'acte d'achat (investissement matériel) » (Bonenfant et Philippette, 2018). Concrétisé par des mesures diverses portant sur l'attraction, la rétention ou la satisfaction d'une audience, l'engagement est devenu un symptôme de la *quantophrénie* – tendance moderne des sociétés occidentales décrite par Alain Supiot (2014), consistant à mettre en chiffres le social et l'individu.

L'engagement de l'individu est ainsi porteur d'une forte dimension symbolique, susceptible d'en faire un objet tant culturel qu'économique. Il a gagné les secteurs culturels connaissant une industrialisation poussée à travers les dispositifs numériques, dans une veine qualifiée par certains de « postfordiste » (Bates, 2000), où la participation de l'individu à la production du service est attendue. Il a été repris par les plateformes audiovisuelles dans les industries culturelles. À titre d'exemple, Netflix publie chaque semestre des « *engagement reports* » sur les programmes les plus vus sur sa plateforme, quand Spotify met en avant auprès des annonceurs une étude de neurosciences mesurant l'« engagement émotionnel » de ses usagers, qui favoriserait leurs intentions d'achat (Spotify, 2024).

En éducation, il est investi à la fois d'enjeux pédagogiques, relatifs à la concentration et à la participation de l'apprenant dans le service, d'une acception affective liée à la satisfaction

procurée par l'usage du produit, et d'enjeux économiques relatifs à la disposition du consommateur ou de l'institution à continuer à payer pour utiliser le service. De cette ambivalence, les EIAH tirent un *modus operandi* industriel, que l'adaptativité doit optimiser sur ses différentes acceptions, en figurant plusieurs profils pour un même récepteur (apprenant, usager et consommateur).

L'*engagement* de l'apprenant, indicateur de l'efficacité pédagogique

L'engagement est d'abord entendu par les firmes éditrices d'EIAH sur un plan pédagogique : il indiquerait l'implication de l'individu dans son apprentissage, et sa propension à interagir avec le produit pour apprendre. Il est en effet mobilisé dans la manière même dont les firmes décrivent leurs activités. Dans l'enseignement de la lecture aux États-Unis, Scholastic (Encadré 4, pp.112-113) vante sur le site de son ITS une « expérience d'apprentissage engageante » (« *Engaging Student Experience* », Scholastic FIRST, 2024). Dans son livre blanc, Duolingo présente l'un des piliers de sa méthodologie « Rester motivé » (« *Stay motivated »)* avec cette assertion : « L'engagement est l'un des facteurs clés dans l'acquisition de compétences »[111] (Freeman et al., 2023, p.6). Enfin, dans la formation professionnelle, Domoscio met en avant, sur son blog, la contribution de l'adaptativité à cet objectif : « Les modules adaptatifs sont conçus pour engager activement les apprenants » (Domoscio, 2024). L'engagement traverse en définitive les cultures et les niveaux éducatifs comme un objectif cardinal des médias adaptatifs.

Cette rationalisation de l'apprentissage, transversale aux secteurs éducatifs, se base sur une histoire longue de la notion dans la recherche académique, *a fortiori* constructiviste, empruntant à la psychologie, aux sciences cognitives et plus récemment aux neurosciences.

Une histoire scientifique disputée entre constructivisme et behaviorisme

L'engagement a connu, outre-Atlantique, une fortune intellectuelle dès les débuts du constructivisme en sciences de l'éducation, comme on l'a vu en première partie. Seymour Papert, le créateur de Logo, le voyait ainsi comme la clé de l'éducation, plutôt que l'explication. Trouver la manière de maximiser la participation active des apprenants a ainsi été un chantier prioritaire de la recherche dans les micromondes puis les EIAH. Dans un article de 1992 que nous avons évoqué dans le premier chapitre, un collectif d'auteurs emmené par l'informaticienne Marlene Jones opposait l'engagement de l'apprenant, artisan de son apprentissage, à la transmission verticale de la connaissance.

En 2000, deux chercheurs importants du courant, John Guthrie et Allen Wigfield synthétisent les points de concorde des chercheurs dans le domaine de l'apprentissage de la lecture, dans un article intitulé « *Engagement and Motivation in Reading* » (Guthrie et Wigfield, 2000). Ils notent d'une part la corrélation entre l'engagement et l'efficacité de l'apprentissage : « La lecture engagée [« *Engaged reading* »] est fortement associée à la réussite en lecture. Les élèves qui lisent activement et fréquemment améliorent en conséquence leur compréhension écrite »[112] (ibid., p.404). Ils perçoivent ainsi l'engagement à travers des comportements quantifiables comme la fréquence de la pratique. Les deux auteurs relèvent d'autre part des travaux ayant conclu que les effets du contexte de la classe sur les résultats des élèves étaient médiés par l'engagement. Celui-ci serait donc le médium par lequel

[111] « *Engagement is one of the key factors that lead to skill development. The more learners interact with content, the more opportunities they have to learn* ».
[112] « *Engaged reading is strongly associated with reading achievement. Students who read actively and frequently improve their comprehension of text as a consequence* ».

l'élève bénéficierait d'externalités positives venues de son environnement. Guthrie et Wigfield ont enfin défini l'engagement par quatre facteurs explorés par les chercheurs en *educational psychology* et en sciences cognitives : la motivation, la stratégie cognitive, le savoir conceptuel et les interactions sociales.

En termes de motivation, Richard Ryan et Edward Deci ont proposé en 1985 la théorie de l'« autodétermination », centrale dans les recherches sur l'engagement au sein des médias numériques (Deci et Ryan, 1985). Ils distinguent deux formes de motivations :

- celle extrinsèque, découlant de facteurs extérieurs comme les récompenses ou les punitions, et étrangère au libre-arbitre de l'individu, et pouvant disparaître après l'apparition de ces facteurs.
- et la motivation intrinsèque, accompagnant la réalisation par l'individu d'une activité, pour sa satisfaction inhérente et son accomplissement personnel.

Entre les deux, ils établissent un continuum de facteurs extérieurs qui motiveraient l'individu à s'engager dans une activité : la régulation externe, proche de la motivation extrinsèque, consistant à obéir aux ordres d'une autorité ou à vouloir profiter d'une récompense ; la régulation introjectée, une forme de rationalisation des coûts/bénéfices d'un comportement particulier ; la régulation identifiée, une prise de conscience de l'importance de certains actes ou de certaines activités ; et enfin, à la lisière de la motivation intrinsèque, la régulation intégrée, qui caractérise une identification de l'individu à ses comportements (Ryan et Deci, 2000).

Une autre théorie faisant référence parmi les chercheurs pour tenter de maximiser la motivation intrinsèque est celle du *flow* (ou flux en français), développée par Mihaly Csikszentmihalyi (1975). Cet état mental est, selon le psychologue, un sentiment de concentration totale et de plaisir dans le processus d'apprentissage, face à des activités qui présentent un niveau de stimulation optimal. Pour atteindre cet état, le niveau de difficulté de l'activité devrait correspondre aux capacités de l'individu. Ainsi la motivation serait maximale lorsque la personne penserait qu'elle a des chances de réussir.

Ces deux théories ont grandement influencé les chercheurs d'obédience constructiviste, s'intéressant aux mécanismes pour faire accéder les apprenants à une motivation intrinsèque jugée plus opérante pédagogiquement que la motivation extrinsèque, qui ferait quant à elle écho aux procédés de renforcement behavioristes. Un foisonnement d'expérimentations sur les EIAH a dès lors eu lieu, jusqu'aux *Learning Analytics* et *Educational Data Mining,* pour y embarquer des dispositifs de motivation intrinsèque, à l'instar des travaux, en France, de Pierre-Yves Oudeyer et de ses collègues d'Inria Flowers (Oudeyer et al., 2016).

Les trois autres facteurs de l'engagement cernés par Guthrie et Wigfield donnent *a priori* à voir les mêmes préoccupations constructivistes – la *stratégie* sous-tendant les capacités métacognitives de l'apprenant, le *savoir conceptuel* ses acquis théoriques – voire socio-constructivistes, à travers les *interactions sociales*.

Néanmoins, le behaviorisme n'est pas loin lorsque, réemployés par des acteurs des ITS, ces outils théoriques visent à faire apparaître des comportements prédéfinis chez l'usager. C'est dans une perspective comportementale que l'engagement a été exploré par les sciences cognitives en particulier. Il a émergé ces dernières années comme un mantra pour favoriser l'acquisition et

la rétention de connaissances. Ceci en complémentarité avec les neurosciences, qui s'attachent depuis plusieurs années à rechercher les circuits neuronaux reliant la détection de la nouveauté (soit une forme d'apprentissage) à la sécrétion de dopamine, hormone de la récompense (Kang et al. 2009, Blanchard et al., 2015).

Le neuropsychologue et président du CSEN Stanislas Dehaene distingue notamment l'*engagement actif* parmi ce qu'il nomme les quatre « piliers de l'apprentissage ». Ainsi, dans *Apprendre ! Les talents du cerveau, le défi des machines* (Dehaene, 2018), il explique à partir de force exemples tirés des neurosciences que la posture active de l'élève est nécessaire à son apprentissage. Il récuse néanmoins l'idée d'autonomie de l'élève et pointe du doigt l'échec du constructivisme sur ce plan, laissant selon lui l'apprenant démuni face au milieu qui l'entoure, en prenant pour exemple l'échec de *Logo* :

> « [Le] principe d'*engagement actif* énonce qu'il est crucial que l'étudiant soit motivé, actif, engagé dans l'apprentissage – mais cela ne signifie en rien qu'il doive être laissé à lui-même. L'échec du constructivisme le montre bien : il est tout aussi crucial que l'enseignant réponde à sa demande en lui fournissant un environnement d'apprentissage progressif, structuré, explicite, conçu pour le guider le plus vite possible vers les sommets » (Dehaene, 2018, p.397)

Le chercheur tranche ainsi le débat scientifique observé en première partie, en justifiant un guidage personnalisé de l'apprenant par une autorité sachante.

Il développe enfin trois autres piliers. Le premier est l'*attention*, consistant pour l'enseignant à chercher à capter l'attention de ses élèves, la focaliser et la maintenir. Le second est le *retour d'information*, devant permettre à l'élève d'expérimenter et de commettre des erreurs pour s'ajuster. Dans la retranscription de son cours au Collège de France, il appelle à des « motivations positives et des récompenses », tout en écartant un « behaviorisme naïf » : « chez notre espèce, éminemment sociale, le regard des autres et la conscience de progresser constituent des récompenses en soi » (Collège de France, 2015b). Le dernier consiste dans la *consolidation,* ou l'automatisation progressive des circuits qui sous-tendent un apprentissage : « L'automatisation transfère les connaissances acquises du compartiment conscient vers des circuits spécialisés et non-conscients, libérant ainsi les ressources mentales pour de nouvelles tâches » (Collège de France, 2015c).

Ainsi Stanislas Dehaene, rendant compte des avancées des sciences cognitives et de la neurodidactique, récuse le constructivisme et rejette toute parenté avec un « behaviorisme naïf ». Pourtant, force est de constater que ces piliers font écho à certains principes du behaviorisme skinnerien, comme le principe du retour sur erreur et de conditionnement opérant. Dans son cours sur l'*engagement actif* au Collège de France, il mettait ainsi en avant le besoin de préserver la motivation de l'élève dans ce guidage :

> « Pour maximiser la curiosité, il faut veiller à présenter à l'enfant des situations d'apprentissage qui ne soient ni trop faciles, ni trop difficiles : c'est le principe d'adaptation de l'enseignement au niveau de l'enfant. Afin de préserver l'engagement et la curiosité, l'enseignant doit éviter d'asséner un long cours magistral, mais faire intervenir les enfants, les tester fréquemment, les guider tout en leur laissant découvrir certains aspects par eux-mêmes, et récompenser systématiquement leur curiosité plutôt que de la décourager » (Collège de France, 2015a).

Par ce principe d'adaptation de la difficulté de l'enseignement, le neuropsychologue fait écho aux préceptes du *flow* et de zone proximale de développement de Lev Vygotsky. Plaçant l'enseignant

au cœur de ses thèses, il offre néanmoins des principes pour des guidages automatisés au sein des ITS.

Cette histoire scientifique de l'engagement, mêlant *educational psychology*, sciences cognitives et neurosciences, est mobilisée par les entreprises du secteur, notamment dans leurs livres blancs et autres parutions expliquant leurs méthodologies. Leurs argumentaires sont empreints de préceptes behavioristes, à l'image de ceux évoqués indirectement par Stanislas Dehaene.

L'idéologisation de l'engagement pédagogique

Les éditeurs édulcorent cette quête d'engagement avec deux types de « promesses socio-techniques » :

- Du côté de l'apprenant, face au problème du décrochage des élèves dans la forme scolaire traditionnelle – que ce soient les plus en difficulté, qui seraient frustrés face à des contenus trop complexes, ou les plus avancés, qui s'ennuieraient en classe face à des contenus trop faciles –, la promesse est de fournir à chacun un tuteur efficient, capable de maintenir son intérêt et sa motivation.
- Du côté de l'enseignant, face aux problèmes du manque de temps pour différencier ses pratiques et des classes trop nombreuses, les analytiques offriraient des possibilités de visualisations, et les recommandations algorithmiques des aides à la décision.

Pour y parvenir, se dessine un schéma centré sur deux relations de cause à effet. Premièrement, la mise en place d'une pédagogie « positive », faite notamment de récompenses, doit mener à une maximisation du temps passé et des interactions de l'usager dans le produit, qui équivaudraient à son engagement. Deuxièmement, l'engagement conduirait à des gains d'apprentissage pour l'élève et à l'efficacité pédagogique du produit. Ramenée, par relation transitive, à des mesures du temps passé et de la quantité d'interactions avec le produit, l'efficacité en deviendrait plus simple à attester pour les firmes éditrices.

La citation du livre blanc de Duolingo exposée plus haut, faisant de l'engagement un ingrédient-clé de l'apprentissage, était suivie de celle-ci : « Plus les apprenants interagissent avec le contenu, plus ils ont d'occasions d'apprendre » (Freeman et al., 2023, p.6). La firme simplifie en effet l'engagement en lui associant des indicateurs nombreux : durée de chaque session, fréquence de l'utilisation, variétés et densité des zones de clic… Elle rationalise la maximisation de ces métriques en vertu d'un enjeu d'efficacité pédagogique.

Pour contribuer à la première relation de cause à effet, les entreprises cherchent à créer les conditions d'un bien-être dans l'apprentissage, suivant une acception psychologique de l'engagement. Comme le notent en effet Bonenfant et Philippette (2018), « du côté de la psychologie, la notion d'engagement est associée positivement à une forme de bien-être (Brault-Labbé et Dubé, 2012) ». Un autre pilier de la théorie de Duolingo est, à cet égard, intitulé « S'épanouir » *(« Feel the delight »)* :

> « Duolingo exploite le principe de conception émotionnelle qui stipule que les gens apprennent mieux lorsque la conception d'une expérience d'apprentissage induit un état émotionnel propice à l'apprentissage (ex. : curiosité, confiance, faible anxiété, etc) sans augmenter le besoin de traiter des informations non pertinentes (Plass et Hovey, 2021). Des expériences joyeuses, qui impliquent souvent des émotions positives, sont liées à un apprentissage accru (Zosh et al., 2018), et une méta-analyse des émotions dans l'apprentissage basé sur

> la technologie a démontré une corrélation légère et positive entre le plaisir et les résultats d'apprentissage (Loderer et al., 2020) » (Freeman et al., 2023, p.7)[113].

Sur le site « compagnon » de son ITS développé dans le cadre du P2IA, Lalilo présente une rubrique explicative de sa « Pédagogie », qui inclut la section « L'engagement des élèves » (Lalilo, 2024) :

> « Nous croyons dans la pédagogie de la réussite pour faire progresser les élèves efficacement. Avec Lalilo, ils apprennent à lire dans un univers ludique avec des encouragements réguliers. Au fil de leur parcours d'apprentissage, ils explorent différents mondes (la forêt, la montagne, la banquise…) et reçoivent des récompenses qui leur permettent de rester motivés » (ibid.).

L'engagement éducatif est ainsi inséparable d'un « *mindset* » positif, devant être transmis par les formats ludiques des produits, pour renforcer la confiance et l'autonomie de l'élève dans l'apprentissage. Toutes ces méthodes font écho aux mantras « *feel good* », ésotériques, des utopies techno-industrielles californiennes caractérisées par Fred Turner dans son ouvrage *Aux sources de l'utopie numérique* (Turner, 2012). Cette dimension contribue à la convergence de ces offres vers les industries culturelles et du divertissement : il s'agit de maximiser le bien-être et de minimiser les efforts de l'usager (dans les proportions permises par la zone proximale de développement).

S'ensuit la seconde relation de cause à effet entre l'engagement, entendu comme une quantité, et l'efficacité pédagogique. Celle-ci tend elle-même à être présentée en termes quantitatifs, liés à la productivité et l'efficience de la pédagogie. Dans l'enseignement supérieur, le livre blanc publié par Knewton en 2015 – à l'époque où la firme licenciait ses algorithmes auprès d'éditeurs – décrivait une relation de cause à effet entre engagement et productivité : « Lorsque l'engagement d'un étudiant diminue, sa productivité a également tendance à baisser, parfois au point de mettre fin entièrement à la séance »[114] (Wilson et Nichols, 2015, p.13). Une rhétorique productiviste encore présente quatre ans plus tard, alors que la firme vend désormais son propre ITS *Alta* et qu'elle valorise les retombées de l'engagement en termes quantitatifs sur son site : « Les étudiants restent engagés avec *Alta* tout au long du semestre. Le taux d'utilisation hebdomadaire d'*Alta* par les étudiants a augmenté de 25 % à l'automne 2018. En moyenne, les étudiants passent 2,8 heures par semaine sur *Alta* »[115] (Wiley, 2019).

Dans un autre domaine, la formation professionnelle, le français Domoscio promet également en conséquence de l'engagement accru des apprenants une efficience quantitative de la formation : « 37 % de temps gagné pour acquérir les compétences cibles et 46 % de rétention supplémentaire des notions 3 mois après la formation » (Domoscio, 2024). Au-delà de la focale entrepreneuriale propre au marché de la formation professionnelle, la promesse donne à voir l'argument commercial que les acteurs des EIAH peuvent tirer de l'engagement.

Au coeur du schéma des *flywheels* observé plus haut chez Duolingo, le duo « efficacité-engagement » est présenté aux investisseurs comme le secret d'un effet de réseau : « Plus nos produits

---

[113] *« Duolingo leverages the emotional design principle which states that people learn better when the design of a learning experience induces an emotional state conducive to learning (e.g., curiosity, confidence, low anxiety, etc.) without increasing the need to process irrelevant information (Plass & Hovey, 2021). Joyful experiences, which often involve positive emotions, are linked to increased learning (Zosh et al., 2018), and a metaanalysis of emotions in technology-based learning demonstrated a small, positive correlation between enjoyment and learning outcomes (Loderer et al., 2020) ».*

[114] « *When a student's engagement drops, her productivity also tends to drop, some times to the point of ending the session entirely* ».

[115] « *Students are staying engaged with alta throughout the semester. The rate of weekly student usage of alta increased by 25% in Fall 2018. On average, students spend 2.8 hours per week in Alta* ».

sont engageants et plus nous enseignons efficacement, plus nos apprenants parlent de Duolingo à leurs amis et plus nous continuons à développer notre base d'apprenants »[116] (Duolingo, 2021, p.3).

Par relation transitive, l'efficacité du produit est simplifiée par les éditeurs : elle doit découler de la simple maximisation des interactions des utilisateurs avec leurs produits, via des métriques comme la fréquence de connexion ou le nombre d'exercices terminés quotidiennement. L'efficacité est donc subjective, dérivant de l'engagement de l'apprenant. Celui-ci apparaît comme le principal objectif opérationnel des firmes.

La confusion entre données d'apprentissage et données d'usage

De ce principe découle une certaine confusion, dans les politiques de confidentialité des firmes, entre données d'usage et données d'apprentissage. À l'étude de ces documents chez les entreprises éditant elles-mêmes leurs ITS – ce qui exclut celles spécialisées dans la formation professionnelle (comme Domoscio en France) qui licencient leurs technologies à des organismes de formation –, il apparaît que l'acception de l'apprentissage est en réalité très large, conditionnée par des aspects liés à la facilité ou la durée de navigation des apprenants, voire à leurs comportements de consommation.

Ce constat est particulièrement palpable dans la formation pour adultes. Sur le marché de l'apprentissage des langues en ligne, Duolingo explique que « le type principal de données que nous collectons est directement lié à l'enseignement et l'apprentissage », mais reconnaît que « ceci inclut également des données sur le moment où les utilisateurs utilisent Duolingo, le nombre de leçons qu'ils terminent et les achats qu'ils font dans l'appli » (Duolingo, 2024).

Dans l'enseignement supérieur, Knewton décrit de façon indistincte la collecte d'« informations personnelles » (« *personal information* »), « toute information relative à une personne identifiée ou identifiable […] » dont les différentes types sont « déterminés par la manière dont les individus interagissent avec nous et par les produits et services particuliers que nous fournissons »[117] (Wiley, 2024b). Parmi les types d'informations personnelles traitées, l'entreprise liste pêle-mêle « les informations académiques ou professionnelles, telles que les publications, antécédents professionnels, antécédents scolaires, expertise », jusqu'aux « informations commerciales sur la manière dont les individus interagissent avec nous, telles que l'historique des achats et les informations d'abonnement ». Entre ces deux pôles figurent les « informations d'engagement », telles que « les commentaires, retours d'expérience, publications, demandes de renseignements et communications »[118] (ibid.). La firme liste enfin quatre finalités, autant liées aux prestations éducatives (« Administration de produits et services »/ » *Administration of products and services* ») qu'à des aspects commerciaux (« Marketing et communication »/« *Marketing and communication* », « Développement de produits et services »/ *« Development of products and services* », « Gestion et administration d'entreprise /« *Business management and* administration » ; ibid.), sans préciser quels types de données sont assignés à chacune.

[116] « *The more engaging our products are, and the more effectively we teach, the more our learners tell their friends about Duolingo and the more we continue to grow our learner base* ».
[117] « *Personal information means any information relating to an identified or identifiable individual. The types of personal information that we process are determined by how individuals interact with us and the particular products and services that we provide* ».
[118] « *Commercial information about how individuals engage with us, such as purchase history and subscription information; […] Academic or Professional information, such as publications, work history, education history, expertise; Engagement information, such as comments, feedback, posts, enquiries, and communications* »

Cette confusion est également visible dans l'éducation primaire et secondaire, bien que le traitement des données des mineurs fasse davantage l'objet de régulations en matière de traitement *marketing* dans les deux pays. Sur le marché de l'enseignement secondaire en France, Evidence B aborde les « données personnelles collectées sur l'usager », qu'elles soient éducatives ou d'usage, et mentionne une finalité centrée sur l'expérience-utilisateur et la personnalisation des services, dans un objectif très large de suivi de la motivation :

> « Lors de l'utilisation du site, Evidence B peut collecter des données personnelles pour améliorer l'expérience utilisateur et personnaliser ses services. Ces données sont utilisées par Evidence B pour l'analyse et l'affichage. Ces données sont utilisées également pour recueillir l'expérience utilisateur et suivre la progression de l'utilisateur, mettre en place un suivi et des statistiques en fonction de la motivation » (Evidence B, 2024).

Transparaît, à travers cette acception très large de l'apprentissage, allant jusqu'à des comportements de navigation, de communication voire de consommation, une acception davantage économique de l'engagement, investi d'enjeux cette fois commerciaux.

L'engagement de l'utilisateur, indicateur de l'efficacité économique

L'engagement est investi d'enjeux commerciaux cruciaux dans les marchés institutionnels et de l'auto-formation. Les métriques qui le quantifient doivent refléter la capacité des firmes à constituer, fidéliser et rentabiliser une communauté éducative. Elles doivent permettre aux entreprises de trouver une prise sur les incertitudes liées aux contraintes d'accès de chacun de ces circuits.

Un indicateur à forts enjeux communicationnels dans les marchés institutionnels

Dans le circuit institutionnel, de l'enseignement primaire à la formation professionnelle, on a observé précédemment que le contact avec les organisations et les usagers était complexe. D'une part, la distinction entre l'acheteur (l'établissement ou le réseau d'établissements), le financeur (l'établissement, la collectivité territoriale ou l'État), l'usager (l'enseignant, l'apprenant voire le parent) et le prescripteur (les autorités éducatives étatiques) crée une complexité aiguë pour les stratégies commerciales des éditeurs.

S'y jouent deux formes d'organisation socio-économiques concurrentes : d'une part, celle obéissant au modèle du club, où les éditeurs collectent directement les revenus des licences par apprenant ou par formateur. Les contraintes d'accès liées à l'éclatement des acheteurs au sein de nombreuses chaînes d'achat institutionnelles sont une barrière pour les plus petites entreprises. D'autre part, une configuration conforme au modèle du courtage, des plateformes d'intermédiation prélevant des redevances sur les licences commandées en leur sein. Dans ce cas, les éditeurs ne perçoivent plus directement les revenus, mais économisent des frais de démarchage particulièrement coûteux pour les plus petits acteurs.

Ces deux types de configurations contraignent les firmes à mobiliser à la fois des arguments commerciaux, éducatifs et politiques pour satisfaire ces divers pôles de réception. L'engagement des usagers est dès lors visé au sein des EIAH, d'abord pour qu'ils jouent un rôle de prescripteurs auprès de leurs hiérarchies (en particulier les enseignants). Il est également visé, peut-être à plus forte raison, comme un enjeu communicationnel à travers des analytiques à destination des échelons prescripteurs et acheteurs, qui peuvent vérifier ainsi l'intensité des usages des produits achetés.

Les plateformes d'intermédiation permettent ces retours informatifs sur chacun des produits licenciés par les districts, tant en France, à travers le *Portail GAR*, que dans les plateformes privées états-uniennes. Dans Clever, un tableau de bord pour administrateurs de district, visible à gauche de la Figure 18, permet de visualiser les usages pour chaque application *EdTech* achetée, via des métriques comme le pourcentage d'élèves et de professeurs actifs et le temps qu'ils y passent en moyenne.

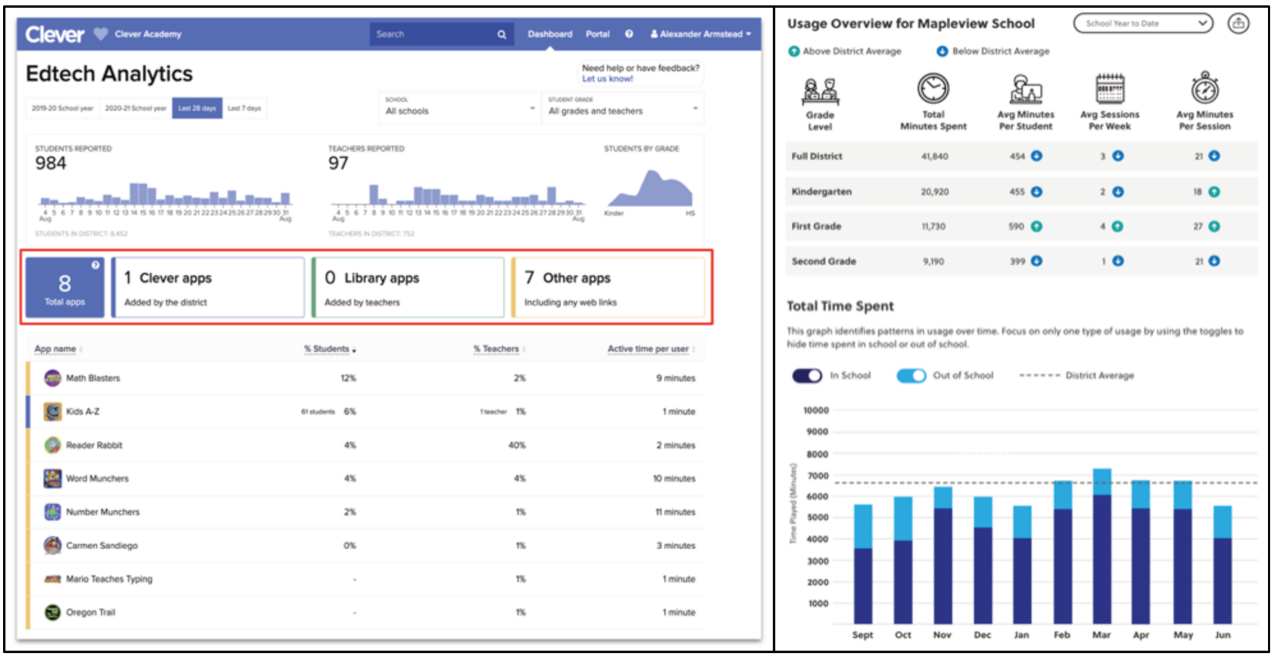


Figure 18. À gauche : Tableau de bord d'un administrateur de district sur Clever. *À droite : tableau de bord d'un directeur d'école sur* Scholastic FIRST. *Sources : Clever, 2024 ; Scholastic First, 2019, p.18*

Du côté des produits distribués directement par des éditeurs, des analytiques sont également embarquées dans les interfaces. *Scholastic FIRST* propose des tableaux de bord pour administrateurs, leur donnant accès à des visualisations de l'engagement à l'échelle de leurs écoles (à droite de la Figure 18). Ceux-ci proposent de comparer les performances des élèves de leurs écoles à la moyenne du district, dans une logique de compétition entre établissements (liée à l'*accountability*). Ce tableau de bord met également en évidence, pour chaque classe allant de *Pre-K* (grande section) à *K-2* (CE1), des informations telles que le nombre total de minutes passées à utiliser le programme, ou la durée moyenne des utilisations par étudiant ou par session. L'administrateur doit en retirer des informations sur l'intensité des usages, grâce à une comparaison avec les moyennes du district pour chaque item (une flèche montante pour une quantité supérieure, une flèche descendante pour une quantité inférieure). Le directeur suit également l'évolution des usages au fil des mois, en comparaison avec la moyenne annuelle du district – un référentiel simpliste n'indiquant pas les variations mensuelles à l'échelle du district. Ce graphique offre également un aperçu de la répartition des usages entre ceux à la maison et à l'école.

Toutes ces informations délivrées sur les plateformes d'intermédiation ou au sein des EIAH sont censées offrir aux directeurs des visualisations hétéroclites et croisées sur l'intensité des usages à l'échelle des écoles, pour juger, outre les types d'usages à renforcer, la pertinence de reconduire des licences sur ces outils.

Un indicateur de l'effet de réseau dans l'auto-formation

Dans l'auto-formation, représentée dans notre corpus par Duolingo, le contact à l'usager-acheteur est plus direct. On a vu que le modèle du courtage était dominant, consacrant l'hégémonie des magasins d'applications des *Big Tech,* qui prélèvent des redevances substantielles sur les revenus des éditeurs. Ces derniers sont donc tenus de diversifier et de « sécuriser » leurs revenus. L'abonnement est dès lors une source de revenus dominante en raison de son aspect récurrent. S'y ajoutent la publicité personnalisée et la vente de bonus à l'acte dans des versions gratuites. Duolingo les opère par exemple en complément de ses formules par abonnement, suivant un modèle d'affaires *freemium*.

Les enjeux commerciaux liés à l'engagement sont dès lors cruciaux : dans le modèle de la formule gratuite avec publicité, l'engagement doit renseigner, à court et moyen terme, la satisfaction de l'usager quant aux contenus personnalisés, éducatifs comme publicitaires. Doit en découler la rétention des usagers « gratuits » comme prospects pour les annonceurs, acheteurs de bonus, et potentiellement futurs abonnés. Dans le modèle de l'abonnement, il prend le sens de *fidélité*, indiquant la propension des abonnés à le rester.

La circulation du terme *engagement* au XX^{e} siècle l'a en effet amené vers la sphère économique et industrielle. Devant la multiplication des contenus culturels, « l'économie de l'attention » a souvent été annoncée à tort et à travers comme le nouvel âge du capitalisme (Citton, 2014). Ce procès a connu un nouvel élan au XXI^{e} siècle, à l'heure de l'essor du « web social ». L'engagement est devenu un horizon à enjeux économiques très forts pour des sites et plateformes cherchant à « capter » l'attention de consommateurs. Il est renseigné par une cohorte d'« indicateurs clés de performance » (« *key performance indicators* »), comme le nombre d'utilisateurs quotidiennement actifs, la durée des sessions, ou le nombre de clics par session.

La poursuite industrielle de l'engagement s'inscrit dans les tendances à la personnalisation qui ont gagné de nombreux secteurs socio-économiques grâce au numérique, faisant converger un peu plus les industries éducatives et culturelles. Le recours aux techniques de fouille de données massives a refaçonné nombre d'offres dans les industries culturelles, en bonne partie devant l'essor des plateformes. Dans ces « service[s] d'intermédiation entre, d'une part, des acteurs qui proposent un bien ou un service et, d'autre part, des acteurs qui demandent à en bénéficier » (Bullich, 2018), l'idée de personnalisation du contenu, notamment par des algorithmes de filtrage collaboratif, est mobilisée pour attirer les consommateurs. Les plateformes ont ainsi gagné nombre d'industries culturelles dès les années 2000 : le cinéma ou la musique enregistrée avec les plateformes de streaming. La personnalisation algorithmique a en définitive une histoire longue en matière d'usages chez les consommateurs. Ceux-là sont habitués à la fourniture de services adaptés, dans des interfaces façonnées pour être flexibles, selon la doctrine ATAWAD (« *Any Time, AnyWhere, Any Device* »), sous l'influence du secteur stratégique de l'« *UX design ».*

L'engagement de l'utilisateur est surveillé par Duolingo pour ses retombées commerciales. Centré sur le nombre d'usagers quotidiennement actifs, il indiquerait d'une part la satisfaction des usagers quant à la formule gratuite, dont découle leur éventuelle propension à s'abonner, sinon leur rétention en tant que prospects pour des annonceurs. Pour maximiser la conversion, un chantier consiste à rendre l'abonnement plus désirable que la formule gratuite aux yeux du grand public. Dans le cas de Duolingo, on a observé que l'abonnement, via les formules *Super Duolingo* et

Duolingo Max, représentait la majeure partie de ses revenus – 76,2% en 2023 (Duolingo, 2024a, p.74) – quand la formule gratuite ne pesait « que » 15,9% du chiffre d'affaires – avec 9,4 points pour la publicité ciblée et 6,5 points pour les achats intégrés liés à la gamification (ibid.).

Cependant, on a observé que la conversion était particulièrement difficile, le ratio du nombre d'abonnés sur les utilisateurs totaux étant d'environ 7,5% (Duolingo, 2024a, p.65). En cause, la forte volatilité des utilisateurs sur le marché des applications d'apprentissage des langues, étant donnés des coûts de sortie quasi-nuls. Ainsi la firme est tenue d'investir des sommes très importantes en marketing et, en particulier, en recherche et développement. Celles-ci ont représenté 36,6% de son chiffre d'affaires en 2023. Elles visent à développer des algorithmes, mais aussi à tester sur des millions d'apprenants des innovations en *A/B testing*.

Dans un article tiré du blog de Duolingo, Lavanya Aprameya, ingénieure du service d'expérimentation, donne à voir l'importance de l'engagement jusque dans les critères retenus en interne pour généraliser ou non ces expérimentations (Aprameya, 2020). Elle explique la méthode suivie : à partir des projets d'expérimentation des ingénieurs est automatisé un protocole de constitution des groupes « test » (par *A/B Testing*) parmi les millions d'utilisateurs de Duolingo dans le monde.

Vient ensuite la phase d'analyse des résultats. Pour expliquer quels indicateurs priment dans la décision de généraliser ou non une fonctionnalité, l'ingénieure prend l'exemple des tests organisés autour d'une nouvelle annonce pour l'abonnement, apparaissant à chaque fois qu'un utilisateur de la version gratuite tentait de réaliser une leçon en étant hors ligne. Devant l'impossibilité d'accéder à l'exercice était alors soulignée la possibilité donnée aux abonnés de télécharger les leçons pour des usages hors ligne. L'argument d'appel vers l'abonnement était opérant : le revenu moyen collecté par utilisateur dans le groupe « test » était supérieur à celui dans le groupe « témoin » (Figure 19. haut). En revanche, cette publicité a nui à l'engagement moyen dans la formule gratuite : le nombre d'usagers quotidiennement actifs a baissé (Figure 19. bas). Dès lors, il a été décidé d'arrêter l'expérimentation : « Nous ne voulons jamais lancer de nouvelles fonctionnalités qui auraient un impact négatif sur les habitudes et les comportements d'apprentissage. Ainsi, même si cette expérience a été un succès du point de vue des revenus, nous avons décidé de l'arrêter et de la réitérer »[119] (ibid.). L'engagement des utilisateurs « gratuits » prime donc sur la rentabilité à court-terme, Duolingo privilégiant le verrouillage d'un effet de réseau à long terme.

[119] « *We never want to launch new features that negatively impact learning habits and behavior, so even though this experiment was successful from a revenue standpoint, we decided to shut it down and iterate on it* ».



Figure 19. Évolution du revenu moyen récolté par utilisateur dans le groupe « test » (courbe verte ») et le groupe « témoin » (courbe bleue) en haut ; Évolution du nombre d'usager quotidiennement actifs dans le groupe « test » en bas. Source : Aprameya, 2020.

Cette prévalence vaut également pour des acteurs des industries culturelles comme Netflix. La firme californienne a ainsi annoncé à partir de 2025 ne plus publier l'évolution trimestrielle de sa base d'abonnés mais plutôt des rapports d'engagement de ses usagers, relatifs au temps passé sur ses programmes. Elle a fait de l'engagement son indicateur central. Une dépêche de l'AFP du 18 juillet 2024 rapporte les justifications de la firme auprès des investisseurs :

> « C'est le meilleur indicateur que nous ayons pour mesurer le plaisir de nos utilisateurs. […] Lorsque les gens regardent plus, ils restent plus longtemps (fidélisation), parlent plus souvent de Netflix (ce qui favorise l'acquisition de nouveaux abonnés) et accordent une plus grande valeur à notre service. » (Challenges, 2024).

En révélant ses dix programmes les plus visionnés chaque semaine, la firme, qui perçoit comme Duolingo des revenus publicitaires en prime de ses abonnements, peut également faire miroiter des espaces publicitaires aux annonceurs. Ainsi ce modèle de développement, consacrant l'engagement comme « indicateur clé de performance », est un autre terrain de convergence entre les industries culturelles et éducatives.

Pour les EIAH distribués auprès du grand public comme pour les plateformes audiovisuelles, les coûts importants en termes de recherche et développement mais également de marketing, induits par la poursuite industrielle de l'engagement, rendent les investissements extérieurs indispensables. Ces dépenses en recherche et développement, visant à soutenir la quête d'engagement, induisent

une tendance à la financiarisation, c'est-à-dire au « développement de l'actionnariat financier » (Bouquillion, 2012).

Pour attirer les investisseurs, les firmes formulent des promesses de rentabilité à court et long terme à destination des investisseurs. La rhétorique est usuelle chez les acteurs du web en phase de croissance. Il s'agit ainsi de lever des fonds pour investir, souvent à perte, dans le perfectionnement de leurs produits, avec l'espoir final d'un effet de réseau. L'engagement prend alors une dimension stratégique, des indicateurs liés comme l'évolution du nombre d'utilisateurs actifs journaliers ou mensuels (« *Daily Active Users* », « *Monthly Active Users* ») devenant des garanties de succès auprès des investisseurs.

Sur le marché de l'apprentissage des langues en ligne, cette promesse prend la forme stéréotypique de roues. Duolingo et Babbel ont recours à cette métaphore (Figure 20), qui a l'avantage communicationnel de faire apparaître des relations de cause à effet de manière mécanique.

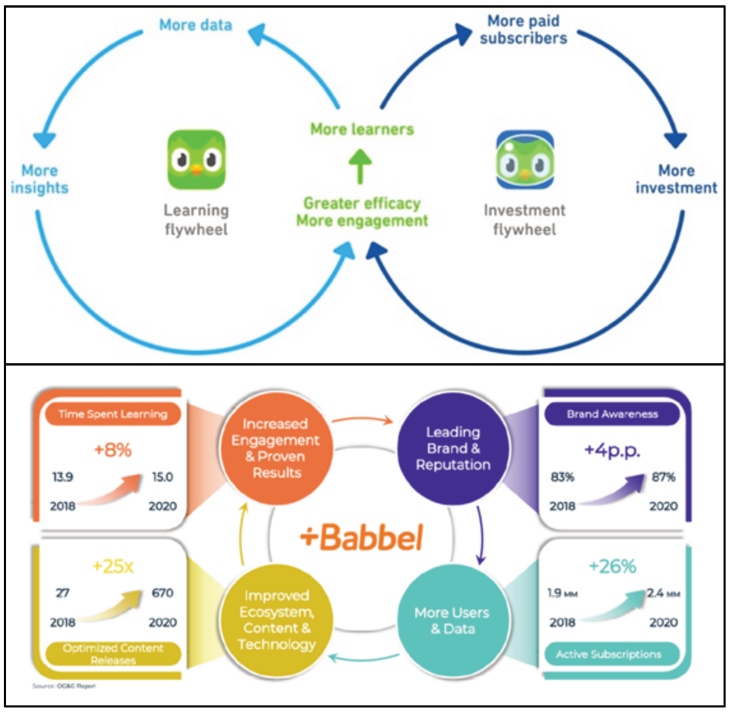


Figure 20. La stratégie de développement de Duolingo en haut ; la flywheel de Babbel en bas. Sources : Duolingo, 2021, p.113 ; Babbel, 2021, p.79

On avait déjà évoqué dans la partie précédente les deux « *flywheels* » de Duolingo, qui postulent, par l'accroissement du duo « efficacité-engagement », l'entraînement des roues pédagogiques et de l'investissement vers l'augmentation toujours plus forte de la base d'utilisateurs, avec cette promesse centrale : « Plus nos produits sont engageants et plus nous enseignons efficacement, plus nos apprenants parlent de Duolingo à leurs amis et plus nous continuons à développer notre base d'apprenants »[120] (Duolingo, 2021, p.3). On retrouve la même rhétorique que celle avancée plus haut par Netflix, faisant miroiter le bouche-à-oreille comme levier d'un effet de réseau.

[120] « *The more engaging our products are, and the more effectively we teach, the more our learners tell their friends about Duolingo and the more we continue to grow our learner base* ».

Babbel, a repris cette dialectique à l'occasion de sa propre tentative d'introduction en bourse quelques mois plus tard[121] (Figure 20. bas). La firme allemande a condensé ces deux *flywheels* en une roue faite de quatre engrenages. Un premier donne à voir le même duo composé de l'engagement (« *increased engagement* ») et de l'efficacité (« *proven results* »). En augmentant, celui-ci conduirait à une hausse de la notoriété de la marque et un accroissement de la base d'utilisateurs abonnés – Babbel ne propose pas de *freemium* mais seulement une offre par abonnement. Les hausses conjuguées des revenus et des données engrangées conduiraient à l'amélioration du produit, rehaussant à son tour le couple « engagement-efficacité » pour initier un nouveau cycle. La rationalisation de l'éducation est visible à travers la quantification de chaque rouage par un graphique censé l'appuyer. Ainsi pour l'accroissement de l'engagement et de l'efficacité, la métrique proposée est le « temps passé à apprendre » (« *time spent learning* »), en hausse de 8% entre 2018 et 2020.

Par ces codes marchands et financiers, l'industrialisation *par l'engagement* apparaît comme un idéal de développement pour les firmes des médias adaptatifs. Nous avons mis au jour les démarches de rationalisation et d'idéologisation de l'éducation pratiquées autour de la notion d'*engagement*. Maintenant que le principe directeur de ce *modus operandi* est explicité, il s'agit de voir en quoi les environnements le concrétisent, en projetant des rôles et comportements pour les publics éducatifs.

2. Les algorithmes d'apprentissage adaptatif comme gouvernail de la *learning experience*

Nous appréhendons les EIAH comme des *dispositifs*, selon l'un des sens donnés par Yves Jeanneret : dont « les normes et les formats qu'ils comportent influencent les formes de la communication » (Jeanneret, 2005, p.51). Pour saisir les conceptions de l'apprentissage et les habitus véhiculés, il s'agit donc d'opérer sur deux niveaux interdépendants dans leur quête de l'engagement : celui des algorithmes en *back-end* et celui de l'expérience utilisateur en *front-end*. Nous appliquons cette méthode dans une analyse croisée des sept médias composant notre corpus, afin de faire émerger des traits communs et saillants.

Nous commençons par le premier niveau, celui des algorithmes d'apprentissage adaptatif. Nous introduisons leurs contributions techniques à l'aune des deux « promesses socio-techniques » évoquées plus haut. (Joly, 2010). Pour expliciter leurs *implications*, nous menons une étude croisée de diverses publications et verbatim d'ingénieurs, afin de saisir le fonctionnement général des algorithmes d'apprentissage adaptatif. Comme expliqué plus haut, nous appréhendons ceux-là en tant que *computextes*, c'est-à-dire comme « processus de calcul qui ne visent [pas] seulement à formater l'apparence du texte, mais le contenu lui-même » (Saemmer, 2012, p.102). Nous interpréterons ceux-ci sur les trois niveaux correspondant à la trilogie d'Edsurge : celui des contenus de remédiation (*adaptive feedback*), celui des questions (*adaptive assessment*), et celui, plus large, de l'ordre des notions (*adaptive sequencing).*

[121] Deux mois après l'apparition du récit mécanique dans le *Form S-1* de Duolingo, Babbel a repris l'idée dans le prospectus préparatoire à son entrée à la Bourse de Francfort (Babbel, 2021, p.79).

La sélection algorithmique de contenus au service de la motivation des apprenants

Les éditeurs font comme on l'a vu la promesse, par leurs algorithmes, de fournir aux apprenants un tuteur personnel efficient, capable de maintenir l'intérêt et la motivation de l'apprenant. Le problème sociétal identifié est le décrochage des élèves dans la forme scolaire standardisée, que ce soient les plus en difficulté scolairement, qui seraient confrontés à des contenus trop difficiles, ou les plus avancés, qui s'ennuieraient en classe. Carnegie Learning prétend ainsi : « L'IA de *MATHia* ne fournit jamais de soutien arbitraire : elle est toujours ciblée et individualisée afin que les étudiants ne soient pas bloqués comme dans *Un jour sans fin*, apprenant sans cesse les mêmes compétences »[122] (Carnegie Learning, 2024).

S'y réfère le pilier de Duolingo, « Apprendre de façon personnalisée » (« *Learn in a personalized way* ») :

> « Duolingo rencontre les apprenants là où ils se trouvent et exploite leurs réponses et comportements dans l'application pour les aider à s'améliorer. De la fourniture de plus d'exercices là où les apprenants en ont besoin, à l'adaptation de l'ordre et de la difficulté des exercices, en passant par la fourniture de commentaires basés sur les réponses des apprenants, nous personnalisons l'environnement d'apprentissage pour stimuler la croissance des connaissances et leur maintien à long terme chez les apprenants. Nous savons que donner aux apprenants le bon niveau de difficulté conduit à un plus grand engagement et plus d'apprentissage »[123]. (Freeman et al., 2023, p.4)

Ainsi l'idéologisation de ces algorithmes et de leur quête d'engagement est-elle exposée. Pour tenir cette promesse, divers algorithmes de *machine learning* sont mobilisés.

Des algorithmes d'apprentissage-machine conduisant l'épistémologie de la donnée

Ces dispositifs sont d'abord chargés de modéliser les acquis de l'apprenant en temps réel, en fonction de quantités plus ou moins massives de données. Ils sont empreints de l'épistémologie de la donnée décrite en première partie, en référence à Bruno Bachimont. Celui-ci expose par ce concept « la compréhension nouvelle promise des activités humaines » du fait des « mégadonnées » (Bachimont, 2015, p.395). Dans les EIAH, le recours à des protocoles inductifs de *data mining* équivaut à une tentative d'objectivation de l'éducation, au moyen des corrélations révélées par les traces numériques. Avec cette démarche comportementale, le courant pédagogique du behaviorisme retrouve un nouveau souffle.

Le relai est ensuite confié à d'autres algorithmes qui vont adapter trois aspects différents : à l'intérieur d'un exercice, les contenus de remédiation en cas d'erreur (*adaptive content*) ou la sélection des questions (*adaptive assessment*) ; et à l'échelle du programme, l'agencement des notions à travailler (*adaptive sequencing*). Visant, explicitement ou non, à trouver la ZPD ou l'état de *flow* de l'élève, ces algorithmes cherchent un affichage optimal permettant à l'élève d'être motivé intrinsèquement par l'activité. Ces techniques de sélection sont employées également par des acteurs du web très variés, notamment issus des industries culturelles, pour personnaliser l'affichage

[122] « *MATHia's AI never provides arbitrary support—it's always targeted and individualized so students don't flounder in a Groundhog Day of learning the same skills over and over* ».

[123] « *Duolingo meets learners where they are and leverages their own responses and in-app behaviors to help them improve. From providing more practice where learners need it, to adjusting exercise order and difficulty, to providing feedback based on learner responses, we personalize the learning environment to boost learners' growth and long-term knowledge. We know that giving learners just the right level of difficulty leads to greater engagement and more learning* ».

de contenus sur leurs plateformes, ce qui renforce un peu plus la convergence entre les industries éducatives et celles-ci.

Le premier type d'algorithmes modélisateurs adopté par les éditeurs est issu de l'« *Item Response Theory* » (IRT). Cette méthode informatique ancienne a été reprise par Knewton, l'ex-start-up emblématique des débuts de l'*adaptive learning*, passée depuis sous le pavillon de Wiley. Pour rappel, cette théorie est apparue avec les débuts du *computer adaptive testing* dans les années 1970. Elle permet, à partir des réponses d'une population d'apprenants aux questions d'un test, de calculer la probabilité qu'un élève réponde correctement à l'une de ces questions, de même que la difficulté relative de ces dernières. Cette théorie a l'inconvénient de n'être valable qu'à l'instant où l'évaluation a été réalisée, ne prenant pas en compte l'oubli potentiel de l'apprenant à terme. Elle a cependant un fort pouvoir prédictif, notamment à l'heure du *Big Data,* ouvrant des perspectives pour situer le niveau d'un apprenant dans une population diverse en fonction de nombreux paramètres, comme l'inattention.

Dans le livre blanc publié par la firme en 2015 (alors qu'elle n'a pas encore été rachetée par Wiley), deux de ses ingénieurs, Kevin Wilson et Zack Nichols, expliquent que son moteur de recommandation repose sur l'IRT. En parallèle, une ontologie, le *knowledge graph*, modélise le domaine en représentant les interactions entre les concepts prérequis dans la maîtrise d'une connaissance (comme une relation compensatoire ou non dans le cas de la non-maîtrise d'un concept). Ainsi, le modèle IRT doit saisir le niveau de l'apprenant de manière précise : la maîtrise d'un élève n'est pas évaluée de façon vague pour une question, mais se décompose en compétences sur les prérequis propres à celle-ci. Ces compétences sont actualisées lors de chaque exercice effectué.

Cette architecture algorithmique prend également en compte des paramètres liés à l'engagement des apprenants. Les ingénieurs notent que « l'engagement d'un élève dans une leçon ou un morceau de contenu peut influencer la façon dont il apprend de cette expérience » (Wilson et Nichols, 2015, p.13). Les ingénieurs ont ainsi développé un modèle de *machine learning* nourri par les données de ses apprenants, devant déduire, à partir du temps passé sur un exercice, le « temps actif » (« *active time* ») ou contraire son risque de décrochage.

Des recommandations de contenus sont enfin opérées à travers des moteurs d'inférence reposant sur le filtrage collaboratif. Développée dans les années 1990, cette famille d'algorithmes a connu une fortune importante avec l'essor des plateformes sur le web. Elle a été appliquée pour la recommandation de contenus culturels selon les préférences déclarées des utilisateurs (Netflix) ou déduites implicitement de leurs données (Amazon ou Facebook par exemple), en vue de maximiser leurs consommations et leur rétention sur la plateforme. Chez Knewton, ce type de technique a été adapté pour formuler des recommandations personnalisées auprès de chaque utilisateur, selon les diverses modélisations évoquées ci-dessus.

Un deuxième type d'algorithme est dit de « bandit manchot multi-bras ». Il est représenté dans notre corpus par l'algorithme ZPDES (*Zone of Proximal Development and Empirical Succes),* chargé de sélectionner les exercices les plus adéquats pour chaque apprenant au sein d'*Adaptiv'Math*. Il a été développé dans le laboratoire de l'Inria Flowers (Clément, 2018). L'un de ses membres, Didier Roy, conseiller scientifique d'Evidence B chargé d'en assurer le transfert dans *Adaptiv'Math*, en présente le fonctionnement dans son article conjoint avec François Bouchet (Bouchet er Roy, 2021).

Ce type d'algorithme tient son nom d'une métaphore liée aux machines à sous de type « bandits-manchots ». Dans cette parabole, un joueur face à plusieurs machines doit décider sur laquelle il est le plus rentable de jouer. En fonction des tirages successifs, le joueur infère quelle est la machine qui a la plus grande espérance de gain en l'utilisant plus que les autres – c'est la « phase d'exploitation » –, tout en continuant de dépenser de l'argent sur les autres machines afin de vérifier qu'elles sont moins rentables (phase d'« exploration »). C'est un compromis « exploration-exploitation » couramment rencontré dans le *machine learning* et en particulier dans les algorithmes dits « de renforcement », qui apprennent par essais-erreurs. Dans le cadre des EIAH, le joueur est remplacé par l'algorithme, la machine à sous par un exercice, et le gain en argent par le progrès pédagogique de l'élève. Ainsi, l'algorithme choisit les exercices en fonction du progrès qu'ils permettent de faire à l'apprenant, tout en explorant un « graphe d'activités » pour trouver d'autres exercices pertinents. Une des particularités est que, si l'espérance de gain d'une machine à sous reste identique à tout moment, l'espérance de progrès d'apprentissage sur un exercice varie avec le temps et la maîtrise de l'apprenant (Clément, 2018).

Dans le cas d'*Adaptiv'Math*, les deux théories de la « zone proximale de développement » et du « *flow* » sont mobilisées afin de guider l'algorithme dans son exploration (Bouchet et Roy, 2021). Celui-ci vise donc l'objectif d'engagement. Développé dans le cadre du projet *Kidlearn*, jeu sérieux de numératie à destination d'enfants de six à huit ans, ZPDES vise fondamentalement à « optimiser l'engagement des apprenants en s'appuyant sur le modèle des progrès d'apprentissage, en lien avec la théorie des motivations intrinsèques »[124] selon les chercheurs impliqués dans le projet (Sauzéon, et al, 2021, p.21).

Il existe de nombreuses applications pratiques du « bandit manchot multi-bras ». Il a été appliqué dès les années 1960 dans l'industrie pharmaceutique, pour réaliser des essais cliniques en attribuant des traitements expérimentaux à des patients de façon optimale. Cette famille d'algorithmes a depuis été utilisée pour l'allocation de ressources entre divers projets de recherche au sein des organisations. Il peut également être utilisé par les analystes financiers en vue de concevoir un portefeuille financier à partir des valeurs boursières les plus rentables, ou dans l'industrie de la publicité en ligne, pour afficher des publicités de manière optimale en fonction de paramètres connus sur l'internaute.

Un troisième type d'algorithmes est le système de classement *Elo.* Il est utilisé par plusieurs entreprises comptant parmi les plus récentes de notre corpus, créées dans les années 2010, comme Lalilo dans l'alphabétisation, ou Duolingo dans l'enseignement des langues en ligne. De fait, ce type d'algorithme est récent en éducation, concomitant de l'arrivée du *Big Data*. On a vu en première partie qu'il était inspiré des travaux du mathématicien Arpad Elo pour classer les joueurs d'échecs. Il a ensuite gagné de nombreux secteurs culturels, et notamment le jeu vidéo compétitif en ligne, pour attribuer des classements aux joueurs et les apparier de façon optimale (Overwatch, World of Warcraft). Il a également été adopté par l'application de rencontre Tinder, son algorithme d'appariement associant des personnes selon leurs classements, qui augmentent ou diminuent en fonction de celui des personnes qu'elles « *matchent »* (Pidoux, 2019).

Autre contribution à la convergence entre industries culturelles et éducatives, cette technologie a été importée en éducation comme on l'a vu au début des années 2010. Ce sont la maîtrise de l'apprenant

[124] Le modèle des progrès d'apprentissage (Oudeyer et Gottlieb, 2016) est un modèle par récompense selon lequel la motivation intrinsèque et l'apprentissage sont liés par une boucle vertueuse : les tâches adaptées aux capacités de l'enfant seraient plus motivantes et généreraient des progrès, devant constituer une récompense stimulante pour de nouveaux apprentissages.

et la difficulté du contenu qui sont cette fois estimées dynamiquement en fonction des interactions des apprenants avec les questions. Grâce au *machine learning,* plus la base d'utilisateurs actifs croît, plus la maîtrise des différents apprenants et les niveaux de difficulté des contenus sont évalués de façon précise. Ce type d'algorithme est donc fortement lié à l'enjeu de l'engagement et à la quête d'un effet de réseau.

Dans le livre blanc publié pour sa version en anglais, Lalilo explique employer un algorithme Elo pour modéliser la cognition de l'élève, à deux moments :

- D'abord lors du « test de positionnement » que l'élève passe lors de sa première connexion, afin de situer son niveau dans le programme conçu par les experts pédagogiques (Renaissance Learning, 2021, p.8). À partir de celui-ci, une première leçon lui est proposée parmi les 650 qui composent le programme.
- Ensuite à l'intérieur des leçons, tout au long des exercices assignés. Chaque leçon est constituée de trois à huit « *Learning objectives* » (LO) que l'élève doit valider pour pouvoir passer à la leçon suivante.

Le niveau de maîtrise de l'apprenant et l'indice de difficulté de chaque exercice sont ainsi calculés et transmis à un autre algorithme, *Catboost*. Développé en *open-source* en 2017 (Prokhorenkova et al., 2017), il consiste dans la « la construction d'un arbre de décision pour identifier et attribuer un exercice avec la bonne difficulté »[125] (Renaissance Learning, 2021, p.9). Son usage par Lalilo vise explicitement la théorie de la zone proximale de développement (ibid.). Le seuil de maîtrise retenu par Lalilo dans un LO est un taux de réussite de 80% sur une séquence d'exercices, permettant à l'élève de passer à la leçon suivante.

Les adaptations qui en ressortent s'apparentent à de l'*adaptive assessment* : dans une séquence prédéfinie d'exercices, l'algorithme permet, pour une même leçon, de sélectionner des choix de réponses plus ou moins fournis selon l'aisance de l'apprenant : dans l'exemple proposé en Figure 21, l'élève débutant (représenté par un poussin) a ainsi le choix entre deux lettres pour compléter le mot, alors que l'élève plus avancé doit arbitrer entre quatre lettres.

[125] « *In combination with the Elo algorithm, it uses Catboost, an open-source algorithm that supports decision-tree building to identify and assign an exercise with the right difficulty (Prokhorenkova et al., 2017).* »

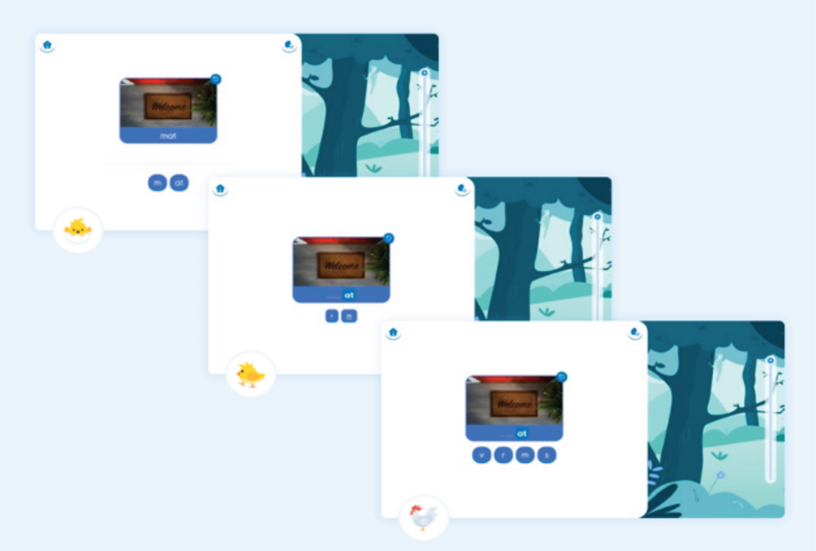


Figure 21. Aperçu de l'adaptive assessment pratiqué au sein d'un même exercice sur Lalilo. Source : Renaissance Lalilo, 2024, p.5.

Sur le marché de l'apprentissage des langues en ligne, Duolingo implémente la méthode Elo à travers son algorithme *Birdbrain*. Il assigne à chacun des exercices un indice de difficulté à partir des performances quotidiennes de millions d'utilisateurs. Il analyse ensuite les traces d'apprentissage passées de l'utilisateur pour prédire ses chances d'y répondre correctement. Comme chez Lalilo, il passe le relai à un autre algorithme, *Session Generator*, qui sélectionne le type d'exercice où l'élève aura une probabilité de succès jugée optimale suivant sa zone proximale de développement (Freeman et al., 2023).

Comme l'expliquait Lavanya Aprameya pour les procès d'*A/B Testing* au sein de la firme, le développement de *Birdbrain* a fait l'objet d'attention en termes d'engagement. Ainsi que l'expliquent d'autres ingénieurs de Duolingo, Klinton Bicknell, Claire Brust et Burr Settles en 2023, *Birdbrain* a d'abord été entraîné « dans l'ombre » en ajustant ses paramètres selon l'écart entre ses prédictions et le succès réels des apprenants, avant d'être testé auprès d'un groupe « test » :

> « Nous voulions voir comment *Birdbrain* affecterait l'engagement des apprenants (mesuré par le temps passé sur les tâches dans l'application) ainsi que l'apprentissage, mesuré par la rapidité avec laquelle les apprenants progressent vers un matériau plus difficile. […] Pour notre plus grand plaisir, *Birdbrain* a constamment entraîné une augmentation des mesures d'engagement et d'apprentissage »[126] (Bicknell et al., 2023).

Par ailleurs, à l'intérieur de chaque exercice, le modèle *Half-Life Regression* doit, après avoir modélisé l'oubli, permettre de personnaliser les mots affichés. Il obéit au principe de la « répétition espacée », qui vise à réafficher celui-ci de façon optimale pour sa mémorisation à long-terme, c'est-à-dire de manière espacée, à chaque optimum de la courbe d'oubli figurant le moment où l'apprenant serait sur le point de l'oublier. Dans le cas d'HLR, cet optimum se situe à la « moitié de durée de vie » (« *half-life* ») du mot dans la mémoire de l'individu, soit au moment où la probabilité de bonne réponse avoisine les 50%. Comme le montre la Figure 22, il en résulte un affichage

[126] *« We wanted to see how Birdbrain would affect learner engagement—measured by time spent on tasks in the app—as well as learning, measured by how quickly learners advanced to more difficult material. […] To our delight, Birdbrain consistently caused both engagement and learning measures to increase ».*

plus répétitif quand l'élève commet des erreurs sur le mot, et plus espacé quand il s'en souvient (Settles, 2016).

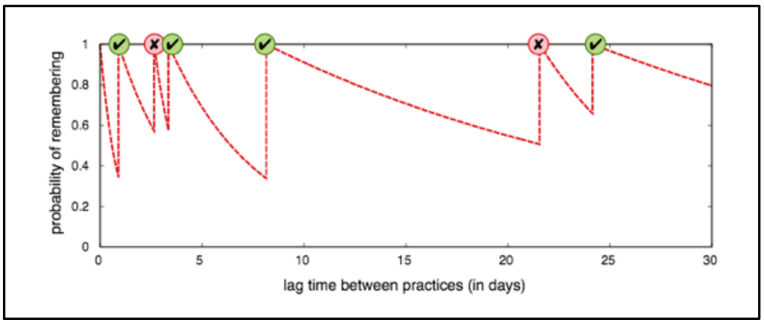


Figure 22. Courbe de mémorisation d'un apprenant concernant un mot, à partir de ses réponses dans le temps (vert pour une bonne réponse, rouge pour une mauvaise). Source : Settles, 2016.

Cependant la personnalisation des exercices obéit également à un impératif de diversification des contenus : « …tout cela est combiné avec le désir de garder les apprenants motivés en leur montrant un mélange de différents types d'exercices et une variété de phrases et de matériel linguistique »[127] (Pajak et Bicknell, 2022). Idem chez Lalilo, qui a fait de la diversité des contenus un critère de sélection pour son algorithme *Catboost* : « Pour garantir la diversité des exercices, l'algorithme tente de réduire la répétition des lignes de contenu »[128] (Renaissance Learning, 2021, p.9).

Des applications du ranking en éducation

Si l'on revient à la promesse de ces firmes de fournir aux apprenants un tuteur personnel capable de maintenir leur intérêt et leur motivation, on observe que ceux-ci sont considérés, quelles que soient les méthodes de modélisation, comme des comportements quantifiables en termes d'interactions et surtout de temps passé sur l'application. Dès lors, les modes d'adaptation poursuivent tous la maximisation de la présence des élèves, en sélectionnant des contenus jugés adéquats par rapport aux comportements d'apprentissage de l'ensemble des apprenants. Ces méthodes s'inscrivent dans la même logique de *ranking* (ou *classement*) que Matteo Pasquinelli (2014) a décrite dans les valorisations opérées par l'algorithme du moteur de recherche de Google, *PageRank* :

> « Par *classement*, j'entends une position dans une certaine hiérarchie, attribuée selon une procédure *objective*, une méthode, un algorithme (tel que cela a lieu lors de l'évaluation de revues académiques, dans les résultats du moteur de recherche de Google ou dans le calcul du nombre d'utilisateurs de Facebook ou Twitter) » (Pasquinelli, 2014).

Le *ranking* s'oppose à la logique du *rating* (ou *notation*), opérée traditionnellement par des humains, suivant des évaluations subjectives basées sur la reconnaissance ou la confiance. De la même manière que Matteo Pasquinelli a décrit la tendance du *ranking* à se substituer au *rating* dans différents modèles économiques de valorisation (des références bibliographiques dans le monde universitaire à l'économie de l'attention sur internet), les algorithmes d'apprentissage adaptatif substituent le *classement* à une évaluation subjective des contenus pédagogiques, telle qu'opérée

[127] « *…all of this is combined with a desire to keep learners motivated by showing them a mix of different exercise types, and a variety of sentences and language material* »
[128] « *To ensure diversity among exercises, the algorithm attempts to reduce repetition of content lines* ».

par un enseignant par exemple. De cette valorisation automatique ressort un habitus d'apprenant en quête d'un apprentissage efficient, divertissant et n'exigeant pas un niveau trop élevé d'efforts, qui le conduirait à quitter l'application.

Des algorithmes de *data visualization* pour permettre aux enseignants de différencier

Avant d'aborder l'autre type d'algorithmes engagés dans les médias adaptatifs, un constat s'impose sur les usages prescrits par les acteurs de notre corpus. Qu'ils privilégient des usages en classe ou à la maison, tous laissent la possibilité d'utiliser leurs médias dans les deux contextes. Pour Duolingo, le seul média initialement non-institutionnel, la firme a lancé, respectivement en 2016 et en 2020, une version pour les écoles de son application linguistique, *Duolingo for Schools,* et une application d'alphabétisation, *Duolingo ABC,* auprès des parents et des enseignants.

Ainsi la filière hybride du *e-learning* semble être un point de convergence des médias adaptatifs, attirés par l'acquisition d'un ethos institutionnel et éducatif – comme nous l'avons montré pour Duolingo sur le marché états-unien de la *children literacy* (Boët, 2023).

Une constante qui en découle est la délivrance d'analytiques à destination des formateurs. Interviennent en amont des algorithmes de *data visualization*. Ils sont chargés de filtrer les données d'apprentissage récoltées pour un élève, de les agréger selon des catégories prédéfinies (comme l'évolution du temps passé sur l'outil, ou les indices de maîtrise sur une compétence) ou encore de comparer les traces de plusieurs élèves pour opérer des regroupements (comme une catégorisation des élèves en difficulté), afin de les restituer sous la forme graphique choisie par la firme éditrice.

Des promesses de visualisations fractales pour l'enseignement

Les entreprises du secteur font la promesse de fournir aux enseignants, par des tableaux de bord, des diagnostics des forces et faiblesses de chaque élève, et des recommandations pour eux-mêmes différencier leurs pratiques pédagogiques. Le problème sociétal identifié est celui d'une surcharge de travail des formateurs face à des effectifs d'élèves trop importants et une pénurie d'enseignants, les empêchant de différencier correctement leurs cours. Carnegie Learning met l'accent sur le temps précieux des enseignants pour promouvoir la technologie de *MATHia* :

> « Le temps des enseignants est limité et, aujourd'hui plus que jamais, ils ont plus à accomplir chaque jour. Ainsi, nous avons mobilisé l'IA de *MATHia* pour développer des analytiques qui enlèvent une partie du fardeau de la collecte de données, très chronophage, des épaules des enseignants »[129] (Carnegie Learning, 2024).

Lalilo adresse pour sa part plusieurs promesses dans son support de formation aux enseignants (Lalilo, 2024). Les deux premières sont des promesses de simplicité : « Voir les informations importantes de la semaine en un coup d'œil » et « suivre la progression des élèves avec un code couleur très simple ». Le « coup d'œil » et « le code couleur très simple » projettent à nouveau l'habitus d'un enseignant pressé, ayant besoin d'une aide facile à comprendre.

La troisième promesse, « consulter le rapport détaillé de chaque élève », est un engagement sur l'aspect fractal des analytiques. Celui-ci constitue une constante des offres adaptatives. La Figure 23

[129] « *Teachers' time is limited, and now, more than ever, they have more to accomplish every day. That's why we've leveraged MATHia's AI to develop reports that take some of the burden of time-consuming data collection off teachers' shoulders* ».

donne à voir les visualisations d'abord collectives que *Scholastic FIRST, Lalilo, Adaptiv'Math* et *MATHia* offrent aux enseignants. Le stéréotype visuel en vigueur est celui d'un « mur » donnant à voir, par des « briques » de couleur, la maîtrise ou le degré de complétion des exercices pour chaque élève.

La perspective est comparative et doit d'abord permettre à l'enseignant de repérer les élèves en difficulté, du fait des couleurs rouge (*Lalilo, Adaptiv'Math* et *MATHia*) ou foncées (*Scholastic FIRST*) qui sont les plus voyantes. Cette dimension comparative est renforcée par la manière séquentielle d'afficher les exercices, ce qui conduit, comme dans *MATHia*, à faire ressortir les élèves en avance ou en retard sur les autres.

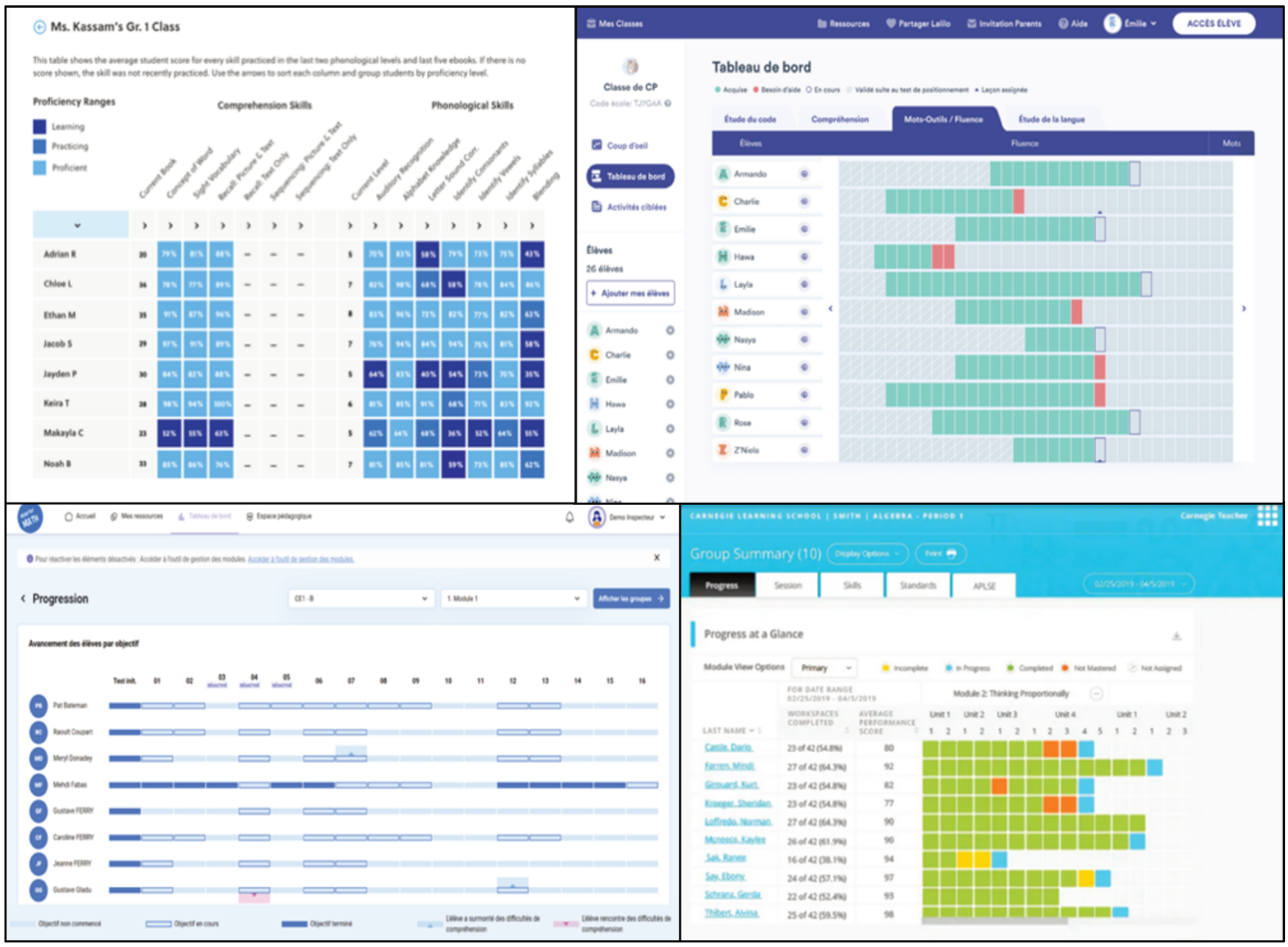


Figure 23. Tableaux de bord collectifs proposés aux enseignants sur Scholastic FIRST *(en haut à gauche),* Lalilo ; *(en haut à droite),* Adaptiv'Math *(en bas à gauche) et* MATHia *(en bas à droite). Sources : Scholastic Research & Validation, 2019, p.15 ; Lalilo, 2024.*

Ces analytiques sont déclinées au niveau individuel pour chaque élève. Elles donnent à voir diverses métriques comme le taux moyen de bonnes réponses, décomposable par compétences pédagogiques. *Scholastic FIRST* propose ainsi une modélisation dynamique de l'évolution des compétences au fil des exercices (Figure 24. gauche). Les analyses montrent d'autre part, quand le séquençage des leçons est individualisé comme chez *Lalilo* (Figure 24. droite), les activités effectuées par l'élève, le taux de réussite pour chacune, le nombre de questions effectuées dans la leçon, le temps passé, et la possibilité de voir les réponses de l'apprenant via un bouton figurant un œil. Une analyse du nombre de jours actifs sur le mois et de leur répartition entre des usages à l'école et à la maison est également disponible.

L'habitus d'un enseignant-surveillant

Ces dispositifs véhiculent une surveillance de l'élève par l'enseignant, qui serait à même de veiller à son assiduité. Enfin des outils plus positifs, devant conduire à la récompense de l'élève par le professeur sont également disponibles, comme la possibilité de voir le nombre de badges collectés par les élèves à l'issue des leçons, ou de « télécharger les diplômes ». On retrouve là une instrumentation du behaviorisme skinnerien et du principe de conditionnement opérant. Pris collectivement, ces dispositifs peuvent être interprétés à la lumière de la thèse de Michel Foucault dans *Surveiller et Punir*, décrivant l'école comme une « machine à apprendre », faite de récompenses et de sanctions normalisatrices auprès des individus. Dans ce sillage, les tableaux de bord font figure de *panopticon* permettant à l'enseignant de surveiller les comportements de chaque élève pour repérer les écarts à une double norme : statistique, relative à la progression moyenne de la classe, et productiviste, devant viser la maximisation des interactions personnelles avec le produit.

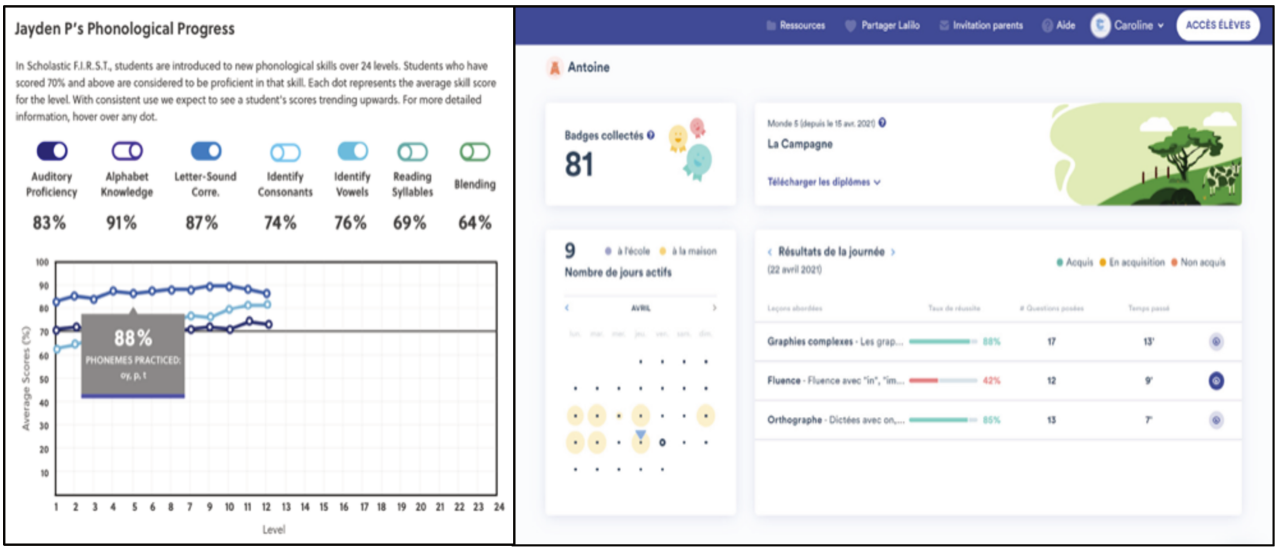


Figure 24. Rapports individuels proposés par Scholastic FIRST *(à gauche) et* Lalilo *(à droite). Sources : Scholastic Research & Validation, 2019, p.15 ; Lalilo, 2024.*

Dans *Scholastic FIRST*, le chapeau explicatif du tableau de bord en offre une illustration : « Avec une utilisation régulière, nous nous attendons à voir les résultats d'un élève augmenter »[130]. L'enseignant peut alors les sanctionner positivement, par des récompenses, ou de manière prophylactique par une différenciation en classe. Ceci d'autant plus que, dans le cas de *Scholastic FIRST*, les enseignants ont accès, comme les directeurs d'écoles, à la situation de leurs classes par rapport à la moyenne du district en termes d'indicateurs d'engagement (le nombre total de minutes passées à utiliser le programme, ou la durée moyenne des utilisations par étudiant ou par session) (Figure 18.droite).

Ces tableaux de bord reprennent les codes du *quantified self*. Ce courant industriel aurait émergé dès les années 1970 (Riphagen et al., 2013) avant de se massifier avec les smartphones, montres et bracelets connectés, permettant des mesures en temps réel d'aspects physiologiques (battements cardiaques ou nombre de pas). La vogue des outils de « quantification de soi » émane le plus souvent d'acteurs états-uniens des technologies, qui ont développé le secteur dans des directions nouvelles (du sport et de la santé à la mesure du temps passé sur les écrans). Les normes d'usage qui

[130] « *With consistent use we expect to see a student's scores trending upwards* »

en découlent sont en partie réinvesties par ces tableaux de bord qui servent aux enseignants pour jauger non pas leur propre activité, mais celle de leurs élèves au jour le jour.

L'engagement de l'apprenant s'y trouve ainsi restitué par les algorithmes de *data visualization*. La promesse de décharger les enseignants d'un travail chronophage en leur fournissant des diagnostics et des recommandations d'intervention pour chaque élève, se concrétise dans des dispositifs non-neutres, qui structurent leurs pratiques. En fonction d'une double norme statistique et productiviste, ils sont censés surveiller les performances de chaque élève pour attribuer des sanctions ou des récompenses, suivant un agenda behavioriste.

En définitive, les algorithmes d'apprentissage adaptatif et de *data visualization* agissent comme des *computextes* en mesure de formater le contenu pédagogique. Pour les uns, il s'agit d'automatiser des parcours pédagogiques pour maximiser les interactions de l'apprenant avec le système, en lui présentant des contenus de difficulté adéquate à des niveaux de détail différents. Pour les autres, il s'agit d'engager l'enseignant dans des visualisations panoptiques, faisant de lui le maillon d'une pédagogie axée sur la surveillance et le conditionnement de l'élève.

Ainsi apprenants et formateurs se trouvent embarqués dans des médias qui créent ce que nous nommons des « boucles d'engagement » : à l'intérieur de celles-ci, les algorithmes d'apprentissage adaptatif sont chargés d'assurer des rétroactions pédagogiques de façon autonome pour les élèves, et les visualisations de données et les recommandations qui en découlent doivent offrir des aides à la décision pour les formateurs. La pédagogique behavioriste qu'ils impulsent s'exprime concrètement dans l'interface, où les dispositifs d'engagement sont divers.

3. Un agencement de dispositifs multimédias dans une perspective behavioriste renouvelée

Pour appréhender le niveau *back-end*, nous menons ci-dessous une analyse techno-sémiotique des médias adaptatifs en tant qu'*architextes*, c'est-à-dire comme environnements « écrits » d'écriture, véhiculant de façon industrielle la représentation de l'apprentissage de la firme éditrice prise collectivement. Nous avons procédé par « observation participante médiatisée » (Bullich, 2018) dans le cas de Duolingo, et, dans celui des firmes issues des marchés institutionnels, par une étude croisée de diverses ressources documentaires publiées par elles ou par leurs usagers, donnant à voir des utilisations du point de vue des élèves ou des formateurs.

Les interfaces consistent en un bricolage de trois principaux dispositifs stéréotypiques, prolongeant la quête de l'engagement impulsée par les algorithmes.

L'objectif de diversification des contenus

Le premier ingrédient est la diversification des contenus pédagogiques affichés. Comme on l'a vu chez Duolingo ou Lalilo, la diversité des contenus est un critère d'adaptativité auquel obéissent les algorithmes d'apprentissage adaptatif. Il s'agit d'éviter ainsi la répétition et la lassitude des apprenants.

Sur le blog de Duolingo, les ingénieurs Bozena Pajak et Klinton Bicknell expliquent qu'après l'écriture, pour chaque leçon, de phrases ludiques par les équipes pédagogiques de la firme, l'IA intervient pour composer des types d'exercices variés (Pajak et Bicknell, 2022). La Figure 25 donne à voir, dans la partie haute, trois types d'exercices d'espagnol composés par l'IA à partir de la phrase « Mon cheval aime la télé », pour travailler la conjugaison du verbe *gustar* : une phrase à trous à remplir avec le verbe conjugué, une deuxième de recomposition de la phrase à partir des mots dans le désordre, et une dernière de reconnaissance orale du verbe conjugué. C'est parmi ces divers formats que l'algorithme *Session Generator* choisit ensuite l'exercice adéquat.

Le fil d'activités intègre en outre depuis quelques années des exercices audiovisuels : des questions de compréhension écrite (« *stories* ») ou des exercices de compréhension orale (« *Duoradio* »). Deux formules par abonnement offrent des fonctionnalités supplémentaires : celle intitulée *Super Duolingo* propose un algorithme interrogeant l'apprenant sur les questions où il a commis des erreurs par le passé, et *Duolingo Max* comprend deux agents conversationnels développés avec *ChatGPT*, lui permettant de s'entraîner à converser.

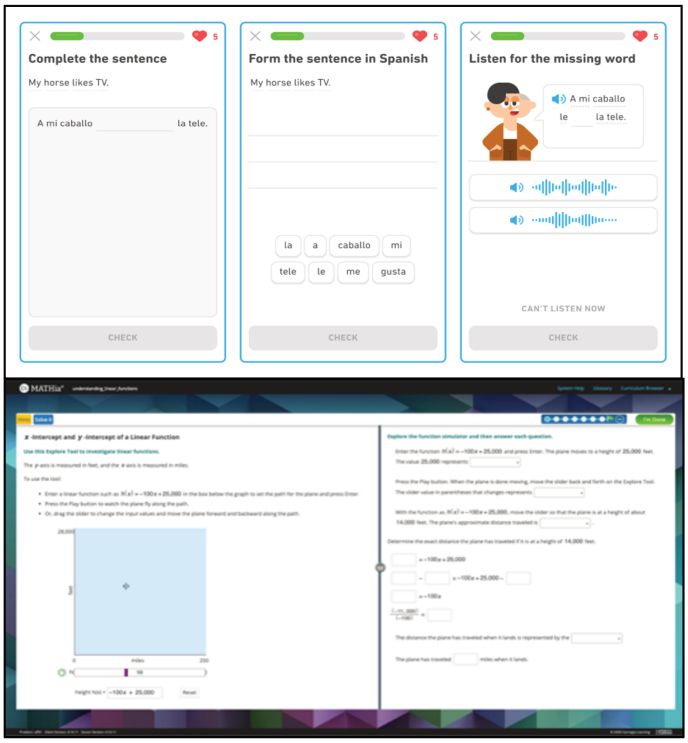


Figure 25. Exemples de diversification des formats d'exercices : en haut, trois variations d'exercices sur le verbe gustar *au sein de* Duolingo *; en bas, un espace «* Concept builder *» dans* MATHia*. Sources : Pajak et Bicknell, 2022 ; Fancsali et Ritter, 2020, p.5*

La partie inférieure de l'image nous montre la récente diversification qu'a entreprise Carnegie Learning pour *MATHia*, en incluant, aux côtés des « *Mastery workspaces* » usuels depuis *Cognitive Tutor*, des « *Concept builders* ». Comme l'explique Steven Fancsali et Steve Ritter (2020), « ces espaces de travail ne sont pas associés à des ensembles de *knowledge components* qui déterminent la maîtrise. Au contraire, les étudiants disposent d'outils qui permettent l'exploration,

de vidéos et d'autres médias, ainsi que d'un ensemble fixe de questions axées sur les concepts »[131] (Fancsali et Ritter, 2020, p.2). Ainsi, des étapes de compréhension sont intercalées pour varier l'approche de l'étudiant, qui se voit proposer dans cet exemple une fenêtre permettant de visualiser la trajectoire d'un avion épousant la fonction linéaire saisie par l'apprenant.

Ces dispositifs très variés, qu'ils soient directement pédagogiques ou ludiques, doivent conduire la quête d'engagement de l'élève.

La gamification

Le deuxième ingrédient typique est la gamification, « l'usage d'éléments de *game design* dans des contextes autres que de jeu » (Deterding et al., 2011). Comme l'ont souligné Sarah Labelle et Aude Seurrat, les promesses entourant le *serious game,* le jeu à vocation pédagogique, sont nombreuses dans l'espace marchand, découlant à la fois de son aspect innovant et de ses supposées vertus pédagogiques (Labelle et Seurrat, 2015). Elles tiennent d'« un raisonnement en deux temps : le *serious game*, par ses spécificités « ludiques », permettrait l'immersion, immersion qui serait facteur de motivation pour l'apprenant » (Seurrat, 2018, p.35).

Le lien entre la gamification et la pédagogie behavioriste derrière l'industrialisation *par l'engagement* est symptomatique d'une *conception médiagénique*. Le jeu doit en effet instrumenter cette doctrine en réduisant le sentiment d'effort des apprenants et en accroissant leur motivation par des récompenses, ce qui les conduirait à prolonger leur présence et maximiser leurs interactions avec le système.

Sur le marché de l'alphabétisation, le recours au jeu est une norme éditoriale pour capter l'attention de l'enfant, comme le révèlent les fonctionnalités de Lalilo. L'élève progresse ainsi dans des « mondes » et y récolte des trésors (Figure 26, en haut à droite), ainsi que des badges et des histoires à lire. À la fin d'un monde, il reçoit comme récompense un « diplôme d'exploration » que l'enseignant peut imprimer et signer (Figure 26, en bas à droite).

Duolingo est devenue, comme on l'a observé dans la partie précédente, une entreprise emblématique de la gamification en éducation. La firme s'inspire du système BLAP (« *Badges, Levels/Leaderboards, Achievements, Points* ») théorisé par Scott Nicholson (2012). Celui-ci vise originellement à maximiser la motivation des joueurs. Duolingo l'applique notamment pour favoriser le retour fréquent de l'apprenant. Comme le révèle la partie gauche de la Figure 26, des récompenses graduelles comme des points d'expérience (« *xp* »), des cœurs (« *hearts* »), des couronnes (« *crowns* »), et des gemmes (« *gems* ») sanctionnent les bonnes réponses de l'apprenant. La mise en compétition d'utilisateurs du monde entier, classés en ligues selon leurs points d'expérience, doit favoriser chez eux un esprit de compétition et une quête de prestige, visible dans les « *leaderboards* ». En parallèle, la comptabilisation des jours d'utilisation consécutifs, les « *streaks* », doit créer chez l'individu une discipline d'autant plus forte que sa série s'allonge.

[131] « *Such workspaces are not associated with sets of KCs that determine mastery. Rather, students are provided with tools that allow for exploration as well as videos and other media, along with a fixed set of conceptually focused questions* ».

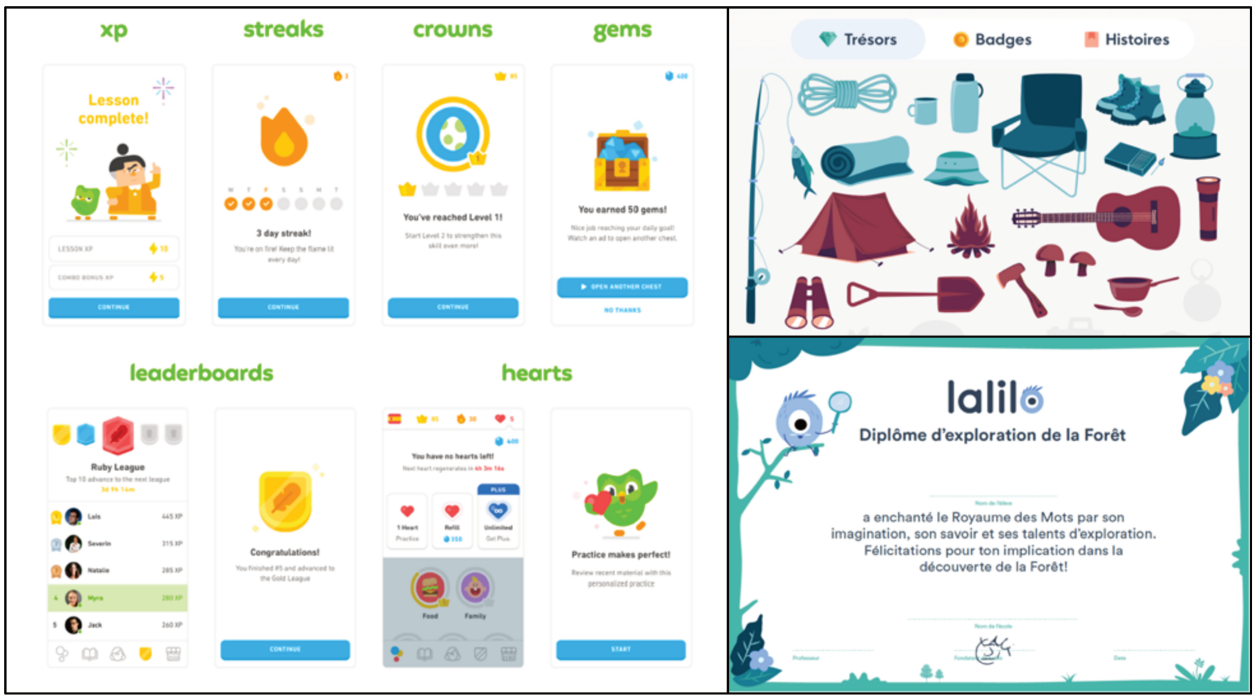


Figure 26. Exemples de formats gamifiés : à gauche, la présentation par Duolingo de ses fonctionnalités ludiques ; en haut à droite, inventaire des trésors d'un élève sur Lalilo *; en bas à droite, un diplôme d'exploration sur* Lalilo. *Sources : Duolingo, 2021 ; Lalilo, 2024.*

Ces dispositifs agissent comme des facteurs supplémentaires de convergence entre industries éducatives et industries culturelles, en l'occurrence vidéoludiques. Chose intéressante, le *quantified self* et la gamification ont partie liée depuis de nombreuses années, la seconde étant mobilisée par les industriels pour pousser les utilisateurs à générer volontairement toujours plus de données. La montre de fitness *Nike+FuelBand*, commercialisée en 2012, transformait ainsi les mouvements suivis en points *NikeFuel*, pouvant débloquer des succès, être partagés avec des amis ou être utilisés pour engager d'autres personnes dans une compétition, à l'image des fonctionnalités sociales de Duolingo.

La présentation linéaire des « learning experiences »

Enfin le troisième ingrédient stéréotypique est l'agencement linéaire des « *learning experiences* », concrétisant la valence behavioriste des algorithmes. Bien que l'adaptation « au jour le jour » des activités soit un principe fondamental des algorithmes d'apprentissage adaptatif, les médias tendent à afficher une linéarité du parcours d'activités auprès des apprenants, quitte à griser les suivantes ou à y apposer des cadenas. On observe en haut de la Figure 27 le parcours linéaire proposé par Lalilo, avec la métaphore d'une route sur laquelle l'élève avancerait d'étapes en étapes.

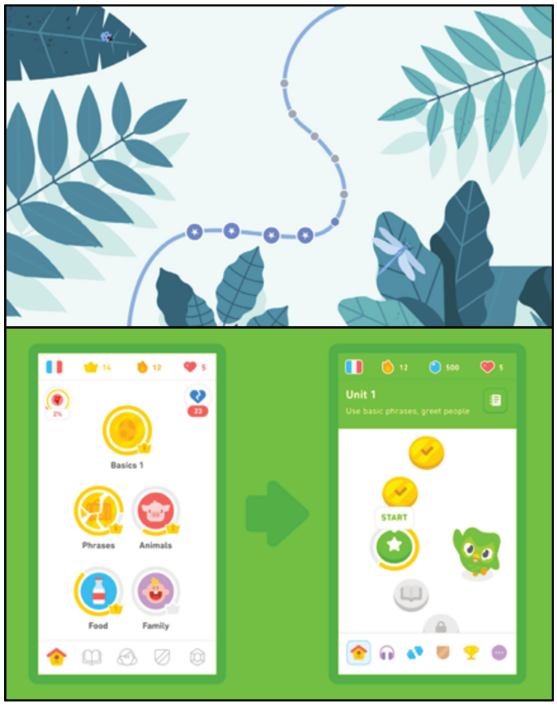


Figure 27. Visualisations des parcours d'apprentissage chez Lalilo en haut et Duolingo en bas. Sources : Lalilo, 2024, p.10 ; Munson et al., 2022

Chez Duolingo, la partie basse de la Figure 27 annonce la refonte de l'interface d'apprentissage. Généralisée en novembre 2022, la nouvelle version révèle, derrière la linéarité, l'enjeu de l'engagement poursuivi par la firme. Jusque-là, l'apprenant évoluait dans des unités séparées par des « *checkpoint tests* ». Il pouvait, à l'intérieur de ces unités, sélectionner les leçons de son choix – celles-ci sont représentées sur le *smartphone* de gauche par des cercles imagés –, et y effectuer des exercices. Dans l'annonce publiée sur le blog de la firme (Munson et al. 2022), il est expliqué que, devant une prétendue perdition des utilisateurs dans cette ancienne version, serait dorénavant proposé un « chemin clair à suivre ». Si l'ancienne version de Duolingo était déjà porteuse de dispositifs de conditionnement opérant, la nouvelle organisation a marqué un renforcement du behaviorisme et de la référence à l'enseignement programmé. Elle contraste d'une part avec l'« expérience » éducative au sens de Dewey (Dewey, 1968) et d'autre part avec les références socio-constructivistes (comme la ZPD) mobilisés par la firme dans son livre blanc. Eu égard aux critères de réussite des expérimentations de la firme, cette refonte a vraisemblablement été justifiée par des gains en termes d'engagement, valorisables économiquement pour l'entreprise.

Conclusion provisoire

En définitive, un modèle industriel, émanant des EIAH disponibles sur les divers marchés éducatifs, est dressé vers la poursuite de l'engagement, tant sur ses versants pédagogique qu'économique. Il consacre d'une part des recherches en laboratoires orientées vers cet indicateur, reposant sur une alliance entre psychologie, sciences cognitives, neurosciences et intelligence artificielle. L'effort des entreprises est de diversifier les dispositifs de captation de l'attention et de la motivation de l'apprenant, pour maximiser ses interactions et donc son apprentissage. Il s'agit aussi de maximiser les traces qu'il laisse en ligne, pour les valoriser ensuite auprès des institutions en tant qu'indices

de l'efficience des produits, ou tenter d'accroître, sur le circuit « *BtoC* », la propension des usagers à s'abonner.

Cette rationalisation métrique s'observe en *back-end* par des méthodes de *data mining* empruntées à l'« épistémologie de la donnée » pour modéliser l'apprentissage, ainsi que des algorithmes de *data visualization*. Les premières sont tendues vers la sélection de contenus pédagogiques suffisamment motivants pour les apprenants, quand les seconds doivent donner aux enseignants des capacités de diagnostic plus approfondies. Ils véhiculent ainsi des habitus d'élèves friands de divertissement et rétifs à l'effort, et d'enseignants relais de normes statistiques et productivistes.

Cette rationalisation est relayée en *front-end* par des dispositifs d'engagement stéréotypés, consistant dans la diversification des contenus, la gamification et l'affichage linéaire des parcours. L'industrialisation *par l'engagement* institue donc une certaine organisation technologique de l'éducation, oscillant entre technologisation additive et substitutive selon les contextes d'usage. Elle ré-instrumente le behaviorisme sur le plan pédagogique, et tend à converger avec la quête d'effets de réseau propre aux plateformes culturelles sur le plan économique. Capitalisant en outre sur les habitudes des usagers en termes de personnalisation et de formats de divertissement, ce modèle est un indice supplémentaire de la convergence industrielle entre l'éducation et la culture à l'heure du numérique.

Cependant il faut rappeler le caractère caricatural de cet idéal-type. Ses caractéristiques, poussées à l'extrême, sont reprises seulement graduellement par les éditeurs d'EIAH, selon plusieurs facteurs. Un premier est les dissensions d'ordre scientifique qui animent, comme on l'a vu, l'espace académique dont sont issus ces algorithmes. Un autre est l'ensemble des contraintes et stratégies socio-économiques des acteurs industriels à même de porter ces dispositifs (start-ups, acteurs anciens du *e-learning*, éditeurs scolaires, *Big Tech*), selon les spécificités de chaque pays et de chaque secteur éducatif. En synthétisant ces divers points de vue, il s'agit maintenant de considérer les principales tensions scientifiques, politiques et économiques que pose le déploiement de ce projet techno-éducatif en France et aux États-Unis.

Chapitre IX. Des tensions posées par le déploiement de l'industrialisation *par l'engagement*

Les tentatives de concrétisation de l'industrialisation *par l'engagement* menées par les acteurs de notre corpus font apparaître diverses tensions. Celles-ci prennent la forme d'incohérences entre les discours d'escorte et les normes instituées par le média d'une même firme, de positionnements adverses entre des entreprises, ou d'apories à même de menacer la légitimité pédagogique de ce type de média. Ils reprennent en partie les dissensions scientifiques observées dans la première partie, entre techno et anthropocentrisme, behaviorisme et constructivisme. Ces tensions ont également partie liée avec les spécificités organisationnelles de chaque niveau d'enseignement, aux cultures éducatives nationales, voire aux positionnements stratégiques des firmes dans la socio-économie des médias adaptatifs étudiée en deuxième partie. Apparaissent en creux des contre-modèles, industriels ou artisanaux, et avec eux de potentiels projets éducatifs concurrents de celui de l'industrialisation *par l'engagement.*

Ces tensions posent au final la question de la *trivialité* de l'adaptativité, de son adaptation et de sa transformation par les acteurs éducatifs qui composent le *champ* social. Nous reprenons la métaphore bourdieusienne du *champ* pour souligner les conflictualités et les rapports de force entre les divers acteurs socio-économiques.

Nous détaillons trois tensions principales à nos yeux : la concurrence de méthodes algorithmiques déductives pour endosser l'ethos de l'*adaptive learning*, les manquements et rigidités méthodologiques de ces médias, et les dissensions d'ordre théorique sur la pertinence d'ingrédients d'engagement et sur l'agentivité à réserver aux enseignants face aux algorithmes.

1. Des alternatives algorithmiques à rebours de l'« épistémologie de la donnée »

Comme nous l'avons dit plus haut, l'*adaptive learning* s'apparente à ce que Gaston Bachelard nommait une « expression obstacle » (Bachelard, 1938) : sa popularité absorbe le récepteur dans un imaginaire empêchant de discerner les méthodes algorithmiques qui le définissent. Cette confusion incite des firmes éditrices développant peu ou pas d'algorithmes de *machine learning* à se réclamer de ce courant industriel. Elles font ainsi concurrence aux entreprises menant bel et bien ce type de recherches dans l'obtention de marchés publics et, pour les outils les plus simplistes, affectent la crédibilité du secteur.

Rappelons d'abord que des médias conformes aux trois critères définitionnels de l'adaptativité (médiatisation, captation en continu des besoins individuels et flexibilité de la pédagogie) peuvent s'appuyer sur des technologies autres que le *machine learning.* Celles-ci peuvent être basées sur des protocoles statistiques et une démarche déductive propre à un domaine, souvent développée depuis des décennies. Celle-ci est étrangère à l'« épistémologie de la donnée », et peut pourtant présenter des fonctionnalités proches de celles que nous avons attribuées à l'industrialisation *par l'engagement*.

L'adaptativité de médias symboliques

C'est le cas de Carnegie Learning et de son ITS *MATHia,* qui repose sur le *Bayesian Knowledge Tracing* (BKT). Cette technique de traçage des connaissances conçue par John Anderson et Albert Corbett en 1995 ne repose guère sur des algorithmes de *machine learning* comme ceux liés à IRT ou Elo, mais sur une IA de type symbolique.

Pour rappel, la spin-off a été créée en 1998 par des chercheurs en *computer* et *cognitive science* de l'université Carnegie Mellon, à partir de l'ITS de mathématiques *Cognitive Tutor,* que John Anderson et ses collègues ont développé dès les années 1980. Il concrétise la théorie ACT–R, qui stipule que l'apprentissage d'un individu est analogue à la transformation de connaissances déclaratives en connaissances procédurales. Le BKT simule alors la progression de l'apprenant en calculant la probabilité qu'il réponde correctement à une question à partir de règles de production. Selon la réponse effective, il corrige ou non ses « croyances ».

Le BKT n'a donc pas recours directement au *machine learning*. Sa formalisation mathématique et notamment la définition de ses paramètres comme la probabilité d'apprentissage ou la probabilité de deviner, sont fixés à l'avance par les chercheurs, selon des théories pédagogiques et des expérimentations effectuées sur des cohortes d'élèves[132]. Cette différence notable avec les autres EIAH de notre corpus peut être expliquée par l'histoire de Carnegie Learning, développé dans un cadre académique, à une époque où la tradition de l'IA symbolique était dominante dans le champ scientifique. Cette méthode a depuis été développée et affinée par les ingénieurs de l'entreprise.

MATHia poursuit cependant une adaptativité variée, selon des modalités conformes à l'industrialisation *par l'engagement*. Son fonctionnement général demeure similaire à celui du *Cognitive Tutor* étudié en première partie. Chaque classe a des cours divisés en 60 à 90 « *workspaces* ». Parmi ceux-ci, l'élève est invité à réaliser des exercices, les « *Mastery workspaces* ». Le BKT estime l'évolution de la maîtrise de l'apprenant sur les *knowledge components* en fonction de ses réponses aux exercices associés et de ses interactions avec le produit. Parmi les interactions de l'apprenant donnant lieu à des adaptations, il y a les demandes d'indices : quand l'élève demande trois indices pour une même question, ses mesures de maîtrise sur les différents KC associés diminuent. Comme pour *Cognitive Tutor,* ceux-ci sont visibles dans le « *skillometer »* (en haut de la Figure 28). Sur son site, Carnegie Learning en fait un vecteur de confiance pour l'apprenant : « Les élèves voient leurs progrès dans plusieurs domaines sur le *skillometer* de *MATHia*. Et à mesure qu'ils deviennent plus compétents dans chaque domaine et voient clairement leurs progrès, leur confiance en eux grandit ! »[133] (Carnegie Learning, 2024). Ainsi Carnegie Learning diversifie les fonctionnalités de *data visualization* caractérisant l'industrialisation *par l'engagement* et pousse plus loin la référence au *quantified self,* en proposant des analytiques aux élèves eux-mêmes.

La firme pousse également plus loin les analytiques à destination des enseignants en leur proposant des « *Adaptive Personalized Learning Score* (APLSE) *Reports* ». Comme le relate un article tiré du blog de l'entreprise ayant trait à son usage de l'IA (Sloan, 2024a), les scores APLSE de la classe et de chaque élève sont calculés à partir des analyses du temps passé par les étudiants dans *MATHia*,

[132] Dès lors, il présente l'avantage de fonctionner convenablement à partir de jeux de données restreints. Comme l'a formulé Marie Sacksick dans une comparaison entre l'IRT et le BKT, le premier apparaît « plus prédictif qu'explicatif, et le BKT […] plus explicatif que prédictif » (Sacksick, 2019, p.76).

[133] « *Students see their progress toward multiple skills on MATHia's Progress Meter. And as they become more proficient in each skill and clearly see their progress, their confidence grows!* »

du nombre de *workspaces* qu'ils maîtrisent, du nombre d'erreurs qu'ils ont commises et des indices qu'ils ont demandés. Le tableau de bord donne à voir la distribution des scores dans la classe, ainsi qu'un graphe indiquant la prédiction de la maîtrise des élèves en fin d'année, comme on peut le voir en bas de la Figure 28 (Sloan, 2024).



Figure 28. Affichage, dans l'espace « skillometer *», des* knowledge components *plus ou moins remplis d'un élève, en haut ; aperçu d'un* APLSE Report, *en bas. Sources : Fancsali et al., 2022, p.3, Sloan, 2024.*

L'entreprise pratique en parallèle une forme d'*adaptive content* sur la sélection de l'indice adéquat. Celui-ci est adapté à la conceptualisation de l'apprenant, dans un répertoire de deux à trois indices par question, comme l'expliquent Fancsali et Ritter (2020). Un exemple est proposé dans l'article sur l'IA (Sloan, 2024). Deux élèves réalisent le même exercice de géométrie lié à la transformation de figures sur le plan. L'un a opté pour une réflexion du triangle par rapport à un axe, tandis que l'autre a choisi une translation. Face à ces deux approches différentes, un algorithme de *MATHia* détecte la stratégie utilisée par l'élève et les erreurs qu'il a commises pour ajuster ses indices, comme le montre la Figure 29.

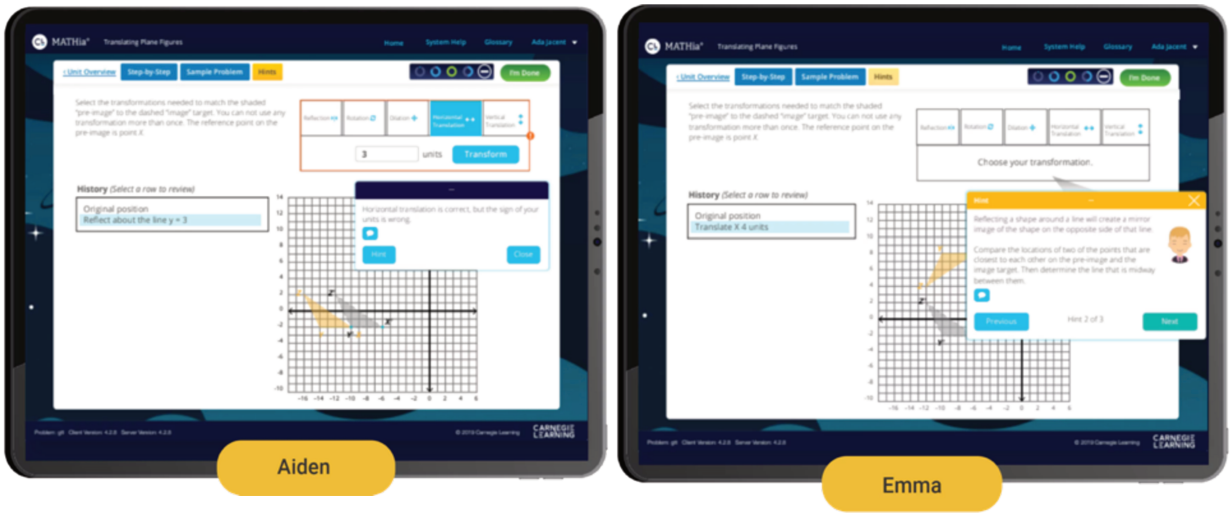


Figure 29. Comparatif d'indices différentiels apportés à des apprenants selon leurs méthodes de résolution. Source : Sloan, 2024.

Ainsi, Carnegie Learning propose différentes formes d'adaptivité sans emprunter pour autant à la démarche inductive centrée sur les données qui caractérise l'industrialisation *par l'engagement*. La firme emprunte néanmoins de nombreux aspects de l'idéal-type. Il poursuit le principe d'engagement dans ses discours, comme on l'a vu dans le chapitre précédent, pratique une sélection algorithmique d'exercices selon la ZPD, et fournit des analytiques étoffées aux élèves comme aux enseignants. Ce choix trouve un écho dans le positionnement ancien de l'entreprise en faveur d'une technologisation additive, complémentaire des pratiques des enseignants en classe voire des tuteurs humains – la firme propose ainsi des séances de tutorat à distance en complément de *MATHia.*

D'autres firmes parent en revanche leurs technologies des atours de l'*adaptive learning* alors qu'elles ne couvrent que des aspects répétitifs de l'apprentissage et/ou ne ne collectent que ponctuellement les données des apprenants.

*Des médias non-adaptatifs se réclamant de l'*adaptive learning

Dans le premier cas de figure, des médias auto-proclamés « adaptatifs » fournissent des services de mémorisation qui ne leur permettent pas d'afficher l'ambition pédagogique générale des EIAH. Leur faible focalisation sur la remédiation simplifie ainsi certains domaines comme l'orthographe, qui serait une affaire de mémorisation plutôt que de compréhension. Ils supposent en outre une configuration « en appoint » du cours d'un enseignant, auprès d'élèves qui auraient déjà compris les notions traitées et n'auraient plus qu'à les mémoriser. Ainsi ces dispositifs projettent-ils une maîtrise chez tous les élèves, à rebours d'une promesse fondamentale de l'adaptativité qui est la lutte contre le décrochage.

Témoin le positionnement de Woonoz, l'éditeur derrière *Le Projet Voltaire* évoqué dans la partie précédente. L'entreprise lyonnaise fondée en 2005 a breveté sa technologie d'« ancrage mémoriel®», inspirée du principe de répétition espacée, permettant pour chaque nouvel étudiant « d'établir [un] profil de mémorisation et [un] niveau initial de connaissance dès les premiers exercices » selon le site du *Projet Voltaire* (Projet Voltaire, 2024). Il y est fait cette promesse au lecteur : « Vous êtes ainsi automatiquement dirigé(e) vers un parcours entièrement personnalisé, à même de corriger vos lacunes

tout en optimisant votre temps de formation » (ibid.). À travers la mémorisation, la promesse de gain de temps et d'efficience prend le dessus sur celle d'une compréhension fondamentale du domaine.

Par ailleurs, sur un plan technique, l'outil repose sur un arbre de décision dans lequel l'élève est « envoyé » selon son « profil de mémorisation » estimé en début de formation. Il n'y a pas de collecte ni de traitement de données en continu qui permettrait d'adapter des contenus, notamment de remédiation, comme dans un EIAH.

Dans un second cas de figure, d'autres firmes proposent cette fois des LMS à l'ambition pédagogique plus généraliste. Parés des atours de l'*adaptive learning,* ces médias pêchent cependant eux-aussi sur le plan technique, leurs algorithmes ne collectant que ponctuellement les données des apprenants. Témoin les groupes anciens du *e-learning* évoqués en première partie, comme Educlever ou Maskott. On a vu qu'ils se présentaient en pionniers de l'*adaptive learning* en France. Pourtant leurs technologies ne répondent pas au critère de l'adaptativité qu'est la collecte de données en continu. Ce sont des arbres de décision dirigeant l'étudiant dans un parcours d'activités, sur la base de tests de positionnement effectués en début de parcours.

Ces entreprises ont cependant remporté des appels d'offres ministériels ayant trait à la différenciation technologique, et jouissent d'une certaine popularité médiatique. Leur positionnement éducatif compatible avec la forme scolaire traditionnelle, sans menace réelle de substitution pour l'enseignant, porterait ainsi ses fruits.

Témoin l'article cité plus haut, tiré du blog *Experiences* de Microsoft, citant Maskott comme un représentant de l'*adaptive learning*. Pascal Bringer, le dirigeant de Maskott, expliquait alors sa posture résolument axée sur l'outillage de l'enseignant : « Nous ne sommes pas là pour remplacer le professeur, mais pour l'armer avec des outils d'aide à la décision. Nous allons lui remonter des informations plus fines que celles qu'il aurait pu avoir, pour l'aider à être plus précis dans sa pédagogie » (ibid).

Se dessine ainsi un modèle alternatif à l'industrialisation *par l'engagement*, où la perspective de l'« humain dans la boucle » serait plus marquée et l'ambition technique des médias moins poussée. Il serait moins industrialisant dans la mesure où les enseignants auraient davantage de marge de manœuvre. Lui aussi ferait cependant évoluer le métier d'enseignant vers celui d'un analyste, devant prendre le temps de s'approprier les analytiques en mobilisant des compétences interprétatives nouvelles. Ce contre-modèle équivaudrait donc à une redéfinition à la marge des pratiques éducatives vers le numérique et les analytiques, au sein d'une forme scolaire pour l'essentiel préservée.

Ainsi le projet industriel *par l'engagement*, basé sur des méthodes inductives de fouille de données en continu, est-il questionné par des alternatives technologiques. Certaines poursuivent en partie son agenda mais selon des méthodes probabilistes « expertes » et « non-apprenantes », à l'image du BKT. D'autres restreignent l'ambition pédagogique en se positionnant sur la mémorisation et/ou en ne collectant les données d'apprenant que de façon ponctuelle. Plus simples à développer pour des structures anciennes du numérique éducatif ou des groupes de l'édition scolaire, ces médias tendent à promouvoir un contre-modèle au service de l'enseignant et compatible avec la forme scolaire. Ils exercent ainsi une concurrence économique auprès des médias proprement adaptatifs, via des candidatures aux mêmes marchés publics.

L'éventail très large de dispositifs qui se réclament de l'*adaptive learning* participe en définitive de son statut d'« expression obstacle » dans l'espace public, au risque de favoriser des malentendus chez les publics éducatifs. Cependant, des incompréhensions peuvent aussi naître des médias adaptatifs ancrés dans la quête d'engagement, à cause de leurs apories et de leurs angles morts techniques.

2. Les limitations techniques du paradigme

L'insuffisance sémantique des données

Une limite fondamentale du projet d'industrialisation *par l'engagement* est la pauvreté sémantique des données collectées en amont. Celles les plus communément traitées, comme l'exactitude des réponses, les temps de réflexion, ou les espacements entre les sessions, reflètent uniquement des comportements en ligne de l'apprenant, sans que des facteurs psychologiques, de stress ou d'inattention puissent être pris en compte pour opérer des diagnostics transverses.

Duolingo a tenté de répondre à cette limite avec l'algorithme *Blame*, visant à identifier la cause derrière l'erreur commise par un apprenant au sein d'une typologie de facteurs (un concept non-maîtrisé en amont, une faute de frappe notamment). Sa finalité pratique est de rendre plus pertinente la sélection des exercices. Pour cela, *Blame* étiquette automatiquement chaque exercice avec les différentes compétences visées. À partir de ses données d'apprentissage ainsi que celles de millions d'autres utilisateurs, il donne la cause la plus probable ayant occasionné une erreur. Cette ambition explicative peut être atteinte quand une ou deux erreurs sont commises, mais devient une gageure quand la réponse accumule trop de fautes. Dans ces situations où le processus cognitif de l'individu est complexe, le système finit par abandonner, comme l'a reconnu Luis Von Ahn, le dirigeant de Duolingo (Protalinski, 2020).

On en revient à la limite cruciale du secteur, exposée au troisième chapitre, selon laquelle les traces numériques demeurent des « traductions partielles et limitées des pratiques » (Ouakrat, Mésangeau, 2016). Pour cette raison, tout projet de technologisation substitutive est une gageure. Un plafond de verre est l'insuffisance de variables quantitatives pour appréhender l'apprentissage. Ceci quand bien même les données seraient enrichies de traces comportementales ou contextuelles, qui ne pourraient de toute façon être exhaustives. Par ailleurs, la recherche automatique de régularités statistiques dans les jeux de données réduit le raisonnement pédagogique à une simple prédiction, quand l'enseignement fait appel à la sensibilité et aux intuitions face à un éventail de facteurs hors ligne.

L'unicité méthodologique

Transparaît enfin dans ce paradigme une rigidité en termes de méthodologie pédagogique. Les firmes conçoivent leurs contenus, aussi divers soient-ils, suivant un cahier des charges globalement unique : ceux-ci doivent être courts et agréables pour maximiser la fréquence d'observations et diminuer l'effort perçu par l'utilisateur. Les médias ne peuvent dès lors prétendre à un enseignement allant au-delà de la transmission et de la vérification de comportements attendus, exigeant par exemple de la part des élèves des productions libres et/ou laborieuses. La quête d'engagement cantonne au contraire la pédagogie à des services d'initiation, d'entraînement voire de remédiation sur des notions ciblées.

Les algorithmes ne peuvent en effet sélectionner parmi une pluralité de méthodes d'enseignement celle qui conviendrait le mieux à un apprenant pour progresser dans une discipline. Dans son livre blanc, Duolingo clame ainsi concevoir ses leçons selon une méthode unique et censément efficace, l'*implicit statistical learning* (Freeman et al., 2023, p. 3). Celle-ci établit que le cerveau humain apprend par observation d'événements dans son environnement, le conduisant à en déduire lui-même des règles statistiques et opérer des prédictions de façon inconsciente (Conway, 2020). Ainsi toutes les leçons sont conçues pour se concentrer sur un nombre limité d'objectifs d'apprentissage. Les algorithmes sélectionnent les mots et formats d'exercices à même d'optimiser l'apprentissage par l'observation. Cette méthode révèle paradoxalement une manière unique d'envisager l'utilisateur et sa cognition. Exit des méthodologies plus laborieuses mais éprouvées comme par exemple l'enseignement grammatical ou la fourniture de tables de conjugaison.

Une opposition entre savoirs scientifiques et narratifs en question

Le paradigme de l'engagement s'accorde en outre avec l'acquisition de savoirs « scientifiques » décomposables (vocabulaire, bases grammaticales, lois algébriques, gestes professionnels normatifs) au détriment de savoirs « narratifs ». Ces deux formes de savoir proposées par Jean-François Lyotard dans *La Condition post-moderne* (1979), puis revisitées par Daniel Peraya à l'aune du web (2012), se distinguent selon cinq critères :

- Le statut discursif des énoncés : « la valeur de vérité de l'énoncé constitue [le] critère d'acceptabilité » du savoir scientifique (Lyotard, 1979, p.36), alors que le savoir narratif relève des « rapports de la communauté avec elle-même et avec son environnement » (ibid., p. 40) et intègre des valeurs politiques et morales, comme la justice, la beauté ou le bonheur.
- Les sujets reconnus comme légitimes pour les produire : des professionnels regroupés dans des institutions de savoir comme les universités formulent les savoirs scientifiques, alors que tout membre d'une communauté sociale peut exprimer un savoir narratif.
- Le rapport au référent : le savoir scientifique est voulu objectif en vertu de « preuves », excluant la subjectivité du chercheur, contrairement au savoir narratif qui se base sur le ressenti de son auteur.
- Le mode de validation et le lien aux savoirs existants : à la validation par les pairs et une accumulation des savoirs scientifiques au fil du temps s'oppose la validation du savoir narratif « dans le fait d'être rapporté, répété et réaffirmé au sein d'une communauté sociale qui en constitue l'espace de légitimation » (Peraya, 2012).

Alors que l'avènement du « web social » a permis une circulation décuplée de savoirs narratifs à travers les forums, messageries et réseaux sociaux, les médias adaptatifs conduisant une industrialisation *par l'engagement* reposent quasi-exclusivement sur des savoirs scientifiques. Ils sollicitent très peu des productions discursives libres des élèves, qui constituent pourtant une étape importante dans le processus d'apprentissage.

La *trivialité* des savoirs narratifs, c'est-à-dire le « caractère transformateur et créatif de [leur] transmission et de [leur] réécriture [...] à travers différents espaces sociaux » (Jeanneret, 2014), est de fait minimisée dans les médias adaptatifs actuels. La nécessité de codifier en amont le domaine avec des contenus de « grains » et de niveaux de difficulté variés semble être la principale raison du recours prioritaire aux savoirs scientifiques.

Par exemple, la méthode de Duolingo peut contribuer aux compétences réceptives (de compréhensions écrite et orale), mais la mise à l'épreuve de compétences productives (d'expressions écrite et orale) est pour le moins rudimentaire. Dans le cas de l'expression écrite, l'application demande à l'utilisateur de rédiger tout au plus une phrase, en tolérant, pour l'espagnol par exemple, des réponses non-accentuées. Elle ne présente guère de tâches qui seraient plus exigeantes en termes d'effort et de temps, comme la rédaction de textes longs requérant la bonne graphie.

En pâtit l'idéal initial de personnalisation pédagogique. Pourtant les éditeurs tendent à présenter leurs outils comme des « solutions » d'apprentissage de disciplines entières, limitant rarement leurs offres à l'initiation ou à la révision de matières. Le recours à l'expression d'*adaptive learning*, faisant miroiter un service d'*apprentissage* complet et autonome, est à cet égard un abus de langage, même pour des médias pratiquant des formes d'adaptativité. En découle un risque de malentendu entre les firmes et certains de leurs apprenants qui auraient des ambitions poussées en termes d'apprentissage.

Les médias numériques intégrant des savoirs narratifs sont pourtant légion depuis l'avènement du « web social ». Ils représentent des contre-modèles en puissance pour les ITS. Par exemple l'aménagement d'espaces de production libre, sans quête d'adaptation, peut inspirer un mode alternatif de conception. Cela suppose, dans une vision ambitieuse de la trivialité, d'ouvrir les ITS en créant, aux côtés des parcours d'exercices adaptatifs, des espaces collaboratifs et/ou de communication entre pairs. Ceux-ci permettraient de faire circuler des savoirs narratifs dans une communauté d'apprenants, à l'image des forums de MOOC. Dans un cours de langue, il s'agirait de permettre à des élèves d'exprimer et de partager leurs pensées à l'écrit ou à l'oral. Dans une matière scientifique, des apprenants pourraient être conduits à interpréter et expliquer à leurs pairs l'origine et le sens d'un théorème. Dans une formation, des élèves pourraient être invités à témoigner de la légitimité de tel geste professionnel dans l'exercice de leur métier.

Dans une voie plus faible de trivialité, l'intégration de ce type de savoirs peut également passer par des formes de différenciation technologique. Témoin l'agent conversationnel de Duolingo, *Roleplay*, basé sur l'IA générative de *ChatGPT,* devant permettre aux apprenants de s'entraîner à converser à l'écrit ou l'oral dans des scénarios (comme aider un touriste à trouver son chemin, en Figure 30), en dehors des exercices adaptatifs. Ce type d'IA permet à l'apprenant de produire et communiquer des savoirs narratifs, certes contraints par un dialogue scénarisé, auprès de la machine. Une visée pédagogique est de lui faire gagner en aisance grâce aux *feedbacks* de cette forme de tuteur automatisé. Celui-ci ne cherche plus à modéliser ses capacités comme un ITS, mais évalue simplement ses productions discursives.

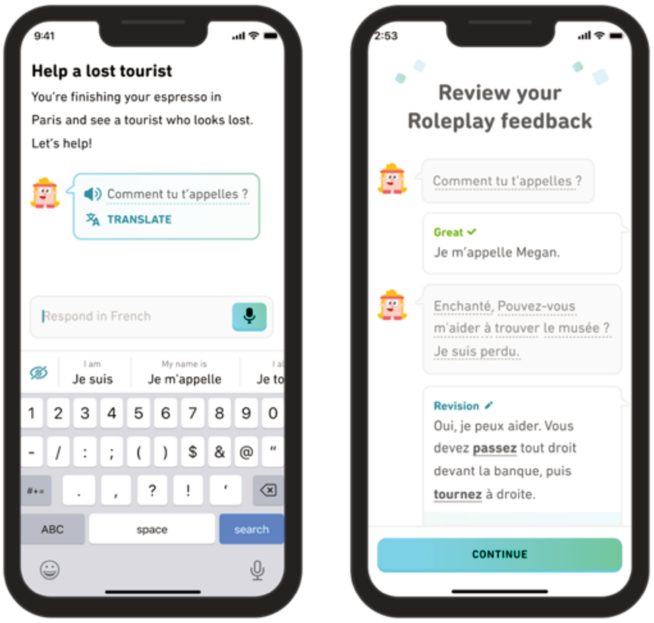


Figure 30. Aperçu d'une conversation dans la fonctionnalité Roleplay *de* Duolingo. *Source : Duolingo, 2023*

3. Des positionnements contradictoires entre constructivisme et behaviorisme

On a observé dans la partie précédente que les médias adeptes de méthodes inductives centrées sur les données prescrivaient le plus souvent des usages complémentaires à des cours en présentiel, en réservant un espace de supervision à l'enseignant. Cette répartition des usages pédagogiques, entre des temps où l'enseignant reste seul en charge des activités, et d'autres où l'outil doit fournir à l'élève un espace de travail en autonomie, crée une ambivalence dans la conception des médias éducatifs sur la question du guidage automatique. L'adaptativité doit-elle, dans une veine constructiviste, porter sur de simples étayages, en laissant les élèves expérimenter et travailler en groupes, ou doit-elle encadrer étroitement les parcours des apprenants dans un schéma behavioriste ?

Un conflit entre autonomie et contrôle des apprenants est réactualisé par le recours croissant aux dispositifs d'engagement observés plus haut. Cependant, la référence au behaviorisme, qui a perdu ses lettres de noblesse dans l'espace public à partir des années 1960, est diversement reconnue par les firmes. Celles-ci se positionnent de façons diverses dans ce débat, la plupart connaissant des conflits intérieurs entre leurs discours publics et la technologisation de leurs médias.

La gamification, sujet d'ambivalences

L'un des ingrédients d'engagement suscitant le plus de dissensions est la gamification, assimilée par certains à un renouvellement du conditionnement opérant skinnerien. Reprenant la théorie de l'auto-détermination de Ryan et Deci, certaines tendent alors dans leurs discours publics à rationaliser leur recours aux récompenses par l'objectif de stimuler la motivation intrinsèque des étudiants.

Dans son livre blanc, Duolingo s'appuie sur les travaux de deux chercheurs en linguistique appliquée ayant avancé les bénéfices pédagogiques des deux types de motivations (Ng et Ng, 2015).

Ils ont insisté sur le rôle joué, dans la motivation extrinsèque, par la satisfaction d'autorités extérieures comme les parents, et sur celui de l'enseignant et des pairs dans la constitution de la motivation intrinsèque. La firme de Pittsburgh en tire de manière cavalière une justification pour la création de contenus non seulement gamifiés mais aussi « agréables », capables selon elle de générer les deux types de motivation :

> « Il a été démontré que la motivation intrinsèque […] et la motivation extrinsèque […] sont liées à une amélioration des performances des apprenants (Ng et Ng, 2015). Duolingo aide à développer les deux types de motivation de nos apprenants. La motivation extrinsèque passe par les fonctionnalités de gamification : l'application utilise des séquences et des notifications pour inciter les apprenants à revenir chaque jour, ainsi que des fonctionnalités telles que les classements pour augmenter le temps que les apprenants passent à apprendre. Parallèlement, nous créons des espaces permettant aux apprenants de développer leur motivation intrinsèque en les invitant à fixer leurs propres objectifs et en construisant un contenu agréable qui suscite leur intérêt et les aide à trouver l'envie d'apprendre » (Freeman et al., 2023, p.9).

D'autres firmes louent les retombées pédagogiques de la gamification, mais les développent dans des produits séparés de leurs ITS, à l'image de Carnegie Learning. L'ancienne spin-off de Carnegie Mellon vante sur son blog les retombées de certains jeux en termes d'apprentissage actif, de collaboration, de pensée critique ou de retour sur erreur. Ainsi son ingénieur en chef Steve Ritter confie-t-il dans un article du blog de l'entreprise : « Des jeux bien conçus donnent souvent aux étudiants des espaces sécurisés pour explorer et se tromper » (Sloan, 2023). Toutefois l'entreprise a préféré appliquer ces principes à des environnements consacrés à l'apprentissage des plus jeunes enfants, comme *MATHia Adventure*, distincts de son ITS historique, qui garde une structure relativement austère et exempte de gamification.

Dans son livre blanc en anglais, Lalilo tente pour sa part un grand écart théorique en se réclamant du conditionnement opérant de Skinner, tout en prétendant faire primer la motivation intrinsèque sur celle extrinsèque, qui dévoierait l'apprentissage :

> « Les récompenses de Lalilo sont cohérentes avec le conditionnement opérant […], dans la mesure où le feedback positif favorise l'acquisition de comportements, ce qui à son tour augmente l'engagement et la motivation. Lalilo vise à se concentrer sur la motivation intrinsèque plutôt que sur la récompense extrinsèque. Le programme évite trop de gamification, qui pourrait entraîner un excès de motivation extrinsèque, entraînant une réduction des gains d'apprentissage pour les étudiants. La recherche montre que des récompenses peuvent amener les élèves à les privilégier par rapport au contenu et à l'expérience d'apprentissage du jeu »[134] (Renaissance Learning, 2021, p.15).

On a pourtant observé dans le chapitre précédent la diversité des récompenses proposées par la filiale de Renaissance Learning (trésors, badges, histoires à lire, diplômes). Pour prouver la prévalence des dispositifs de motivation intrinsèque, la firme prend appui sur la typologie des « joueurs » de Richard Bartle (1996), adaptée par Andrzej Marczewski à la gamification éducative (2013). Dans cette théorie, le profil du « joueur » est motivé par une récompense extrinsèque, et quatre autres sont motivés de façon intrinsèque : l'« extraverti » par les relations sociales, l'« esprit libre » par l'autonomie et la découverte, le « performant » par l'idée de maîtrise et le « philanthrope » par un idéal propre.

[134] « *Lalilo's rewards are consistent with operant conditioning (efforts to increase or decrease a behavior by providing or taking away reinforcement or punishment; Skinner, 1971), in that positive feedback increases behaviors, which in turn increases engagement and motivation. Lalilo aims to focus on intrinsic motivation rather than extrinsic reward. The program avoids too much gamification, which could consequently result in excess extrinsic motivation—resulting in a reduction of learning benefits for the students […]. Research shows tangible rewards can lead students to value the reward over the content and learning experience of the game […]* ».

Lalilo prétend laisser de côté le « joueur » étant donné que « presque tous les badges du jeu sont remportés sur un critère de maîtrise ». La firme confond ainsi moyens et fins pour considérer qu'un effort d'apprentissage suffit à exclure une motivation extrinsèque. Pourtant la définition première du concept par Ryan et Deci (2000) insiste plutôt sur la volonté d'atteindre un résultat séparé. L'application d'une récompense extérieure à l'apprentissage comme un badge, quelle que soit la manière de l'atteindre, engendre par définition une motivation extrinsèque.

La firme prétend au contraire s'adresser aux quatre autres profils d'élèves qui seraient motivés de façon intrinsèque. Par exemple, l'« esprit libre » serait une cible de Lalilo, étant donné que, dans son service, « un étudiant ne peut acquérir une compétence qu'en la maîtrisant par lui-même ». Une acception très restreinte de l'autonomie est ainsi retenue, loin de celle prônée par le constructivisme et explorée dans les micromondes ou les environnements qui en ont découlé, permettant aux apprenants de découvrir par eux-mêmes le domaine.

En définitive, la gamification fait l'objet d'une certaine ambiguïté dans les stratégies discursives et technologiques des éditeurs. Si ses retombées en termes de motivation extrinsèque sont admises, celles en matière d'apprentissage sont encore discutées. Des firmes comme Duolingo promeuvent le recours au jeu tant dans leurs discours que dans leurs services. D'autres en reconnaissent l'utilité sans intégrer de gamification à leurs ITS, à l'image de Carnegie Learning. À l'inverse, des firmes comme Lalilo proclament éviter ce type d'énonciation pour se concentrer sur l'enseignement, tout en ayant recours à des dispositifs variés de récompenses. Une telle rhétorique semble être liée à une volonté de la firme de se parer d'un ethos scolaire « légitime », éloigné des codes des industries culturelles, afin de mieux pénétrer les institutions éducatives.

Ces dissensions parmi les acteurs de notre corpus renvoient à celles existant aujourd'hui chez les chercheurs des *learning sciences* entre courants constructivistes et behaviorisme. Apparaissent également des différences de positionnement sur la marge de manœuvre à réserver à l'enseignant face aux décisions algorithmiques des ITS.

Les formes d'agentivité de l'enseignant face aux algorithmes

Si le rôle du formateur est reconnu par l'ensemble des éditeurs qui tendent à « s'institutionnaliser », des divergences restent palpables sur la technologisation adéquate des activités pédagogiques, entre addition et substitution à l'enseignant pour certaines d'entre elles. L'influence constructiviste commande à première vue de laisser à ce dernier le soin de planifier les activités, alors que la veine behavioriste suppose une programmation étroite des parcours, l'enseignant devenant un relai de sa mise en application.

Trois types d'agentivité

Trois types d'agentivité prévue pour l'enseignant peuvent être dénombrés parmi les médias de notre corpus. Ils donnent à voir des choix différents opérés par les entreprises. Nous les présentons ci-après dans l'ordre croissant des capacités d'action données au formateur. Le premier est l'*analyse*. On l'a observé, dans le modèle de l'industrialisation *par l'engagement*, et en particulier chez *Scholastic FIRST*, où le formateur est projeté dans un rôle d'analyste par des visualisations graphiques des apprentissages à l'échelle collective et individuelle. La possibilité lui est offerte d'intervenir, selon ces analyses, en pratiquant une différenciation en classe. Nulle capacité d'intervention au sein

de l'ITS n'est donc possible, les algorithmes d'apprentissage adaptatif assurant intégralement la pédagogie.

Vient ensuite une forme d'agentivité plus marquée : l'*assignation* de contenus pédagogiques, en particulier d'exercices, en complément de l'adaptativité algorithmique. On la retrouve chez *Lalilo* par exemple, qui propose aux enseignants de créer des « activités ciblées » sur des leçons autres que celles déjà identifiées par les algorithmes comme « non-maîtrisées », et qui sont automatiquement réaffichées auprès des élèves (Figure 32). Sur son blog, la filiale de Renaissance conseille d'utiliser ce dispositif autant « en réinvestissement d'une leçon vue en classe », qu'en « remédiation d'une notion acquise par la majorité de la classe », ou « en découverte en groupe-classe ou de niveau » (Lalilo Blog, 2024).

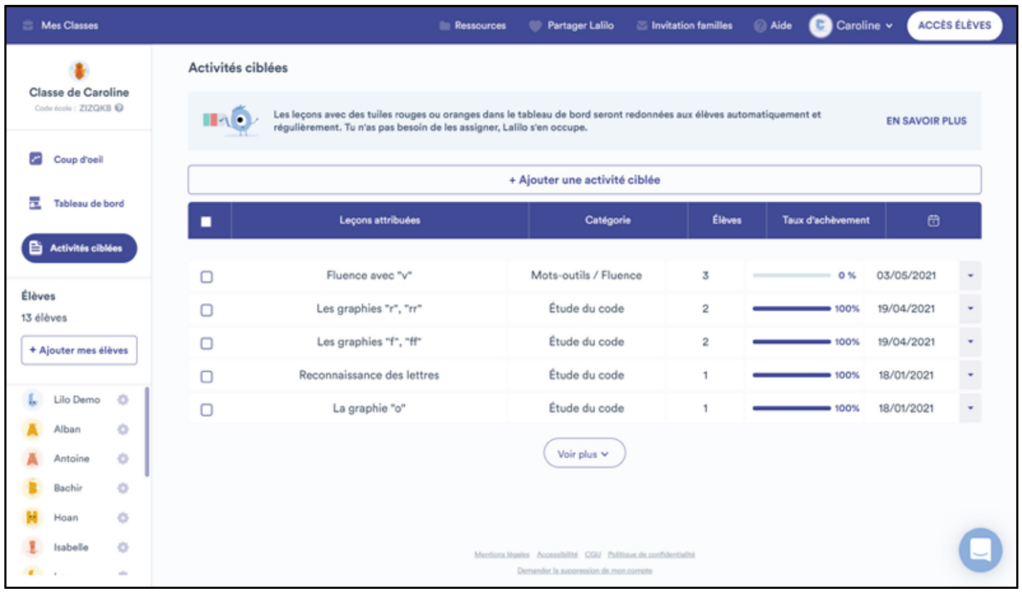


Figure 32. Aperçu de l'onglet « Activités ciblées » dans l'interface enseignant de Lalilo. *Source : Lalilo, 2024.*

Duolingo for Schools propose également d'assigner des leçons sur des notions particulières, avec des degrés d'importance figurés par le nombre de points d'expérience (XP). L'enseignant relaie ainsi la gamification de l'application originelle en décidant lui-même de la quantité de motivation extrinsèque. Cette assignation se fait cependant de façon collective et non individuelle.

La troisième forme d'agentivité, la plus poussée, est l'*encadrement* de l'adaptativité au sein de l'ITS. Incluant l'*assignation* d'exercices, elle offre d'abord à l'enseignant des recommandations diverses pour ses propres pratiques de différenciation. Celui-ci a en outre la possibilité de cadrer les interventions des algorithmes d'apprentissage adaptatif, en décidant d'une part des leçons qui composent le programme d'élèves ou de groupes d'élèves, et en contrôlant d'autre part les modélisations de l'apprenant opérées par les algorithmes.

C'est le cas d'Evidence B qui a mis en place ce type d'agentivité dans *Adaptiv'Math*. Du côté des recommandations, l'enseignant se voit épaulé par un algorithme de « *clustering* » qui regroupe les élèves par niveaux : l'algorithme SACCOM (pour *Student Activity Clustering for Classroom Orchestration & Monitoring*). Il a été développé par l'équipe MOCAH du laboratoire d'intelligence artificielle LIP-6, dont est membre François Bouchet, conseiller scientifique d'Evidence B.

Dans l'article de 2021 cité plus haut qu'il a co-écrit avec Didier Roy, il explique que cet algorithme procède en deux temps. D'abord il formule des *clusters*, ou groupes d'élèves homogènes, à partir des données de l'ensemble des utilisateurs, pour chaque objectif et chaque niveau de difficulté, suivant quatre principales variables :

- Le score moyen obtenu, devant permettre de déterminer le degré de maîtrise sur un niveau ;
- La durée moyenne passée sur un exercice, devant donner à la fois une idée du degré de maîtrise d'un objectif et du degré d'effort impliqué.
- Le niveau de difficulté maximum atteint au sein d'un objectif, devant refléter la progression de l'élève sur cet objectif ;
- Le nombre d'exercices, qui donnerait une idée de la proportion de temps passé par l'élève à travailler sur un certain niveau de difficulté (Bouchet et Roy, 2021, p.10).

Chaque *cluster* est réduit à un élève « virtuel », représentatif du groupe. C'est par rapport aux élèves virtuels des différents *clusters*, figurés dans un graphe comme celui en haut de la Figure 33, qu'est comparé dans un second temps tout nouvel apprenant, pour être ensuite placé dans le groupe le plus proche. Une visualisation de ces regroupements est proposée au bas de la Figure 33.



Figure 33. Aperçu de la méthode de clustering utilisée par l'algorithme SACCOM en haut ; visualisation d'un regroupement proposé à l'enseignant dans Adaptiv'Math en bas. Sources : Evidence B, 2021, p.29 ; Adaptiv'Math, 2024.

Les usages de ces regroupements sont laissés à l'avenant. L'idée principale est de donner à l'enseignant les « outils » pour opérer lui-même une différenciation en classe. Des scénarios d'usage de type constructiviste lui sont cependant conseillés dans le guide pédagogique publié par Evidence B à l'issue du P2IA en 2021. Dans un premier scénario, la firme propose d'organiser des groupes

de remédiation en classe, avec un temps de quinze à vingt minutes durant lequel l'enseignant rappellerait les concepts mathématiques importants et proposerait des problèmes oraux ou écrits à résoudre collectivement, avant d'organiser un temps de travail en autonomie sur l'outil (Evidence B, 2021, p.21). Un autre scénario consiste à organiser des groupes de besoin pour permettre à des élèves qui seraient en avance de découvrir de nouvelles notions, en étant seuls sur l'outil puis en groupe (ibid.).

Le professeur a ensuite la possibilité de cadrer les interventions des algorithmes d'apprentissage adaptatif sur les trois étapes principales de la pédagogie distinguées par François Bouchet et Didier Roy : le test de positionnement, le parcours adaptatif au sein des modules et l'évaluation finale. Il peut ainsi décider quand doit se tenir le test de positionnement, et à quels moments les données des apprenants remontent ou non vers la base de traces d'activités où s'opère la modélisation de leurs niveaux. Il peut aussi insérer des exercices, et, depuis 2024, les constituer en « playlist » à réaliser, dans un ordre prédéfini, par un élève ou un groupe. Il peut également modifier la position d'un élève d'un *cluster* à un autre s'il juge l'affectation non-pertinente (Figure 34). Il peut enfin verrouiller une partie du programme pour que l'ordre des notions soit stabilisé et que ZPDES, l'algorithme de sélection d'exercices, n'intervienne pas (Bouchet et Roy, 2021).

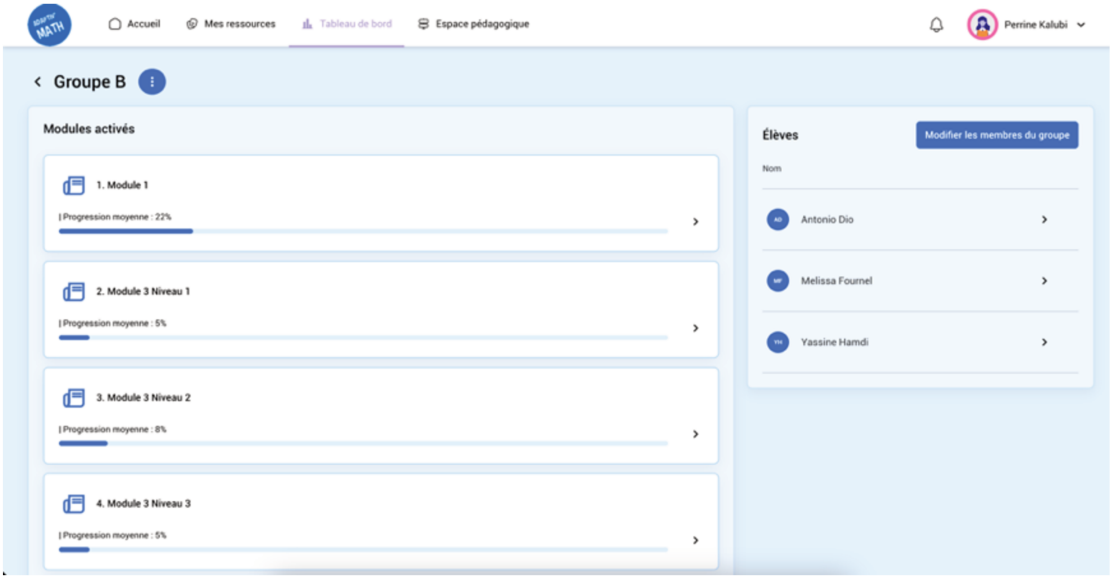


Figure 34. Interface permettant à l'enseignant de veiller aux modules assignés et aux membres d'un cluster. *Source : Adaptiv'Math, 2024.*

Autant de possibilités d'intervention restent exceptionnelles dans le cadre de notre corpus, et semblent fortement liées à un facteur culturel. L'enseignant français jouit en effet d'une liberté pédagogique forte dans sa classe par rapport à son homologue états-unien, dont les pratiques sont traditionnellement plus encadrées. Le mode d'agentivité plus simpliste qu'est l'*analyse* est ainsi davantage présent dans les ITS états-uniens de notre corpus. Dans le mode au contraire le plus poussé, *Adaptiv'Math* semble miser sur un habitus d'enseignant « chef d'orchestre », sans doute en vue de mieux pénétrer le marché scolaire et de remporter des marchés publics. Néanmoins l'ITS projette des compétences en informatique et *digital literacy* poussées chez les enseignants, qui manquent pourtant dans leurs formations actuelles.

Des formes d'agentivité revisitant la problématique de l'explicabilité

Ces formes d'agentivité jettent un éclairage nouveau sur le sujet de l'explicabilité des décisions algorithmiques. On a vu au troisième chapitre qu'il s'agissait d'un sujet brûlant pour les recherches en *Learning sciences*. Devant la complexité croissante des EIAH reposant sur le *machine learning*, de nombreux chercheurs en EDM et LA tentent d'une part de neutraliser les biais contenus dans les données, et d'autre part de penser les conditions d'usages éclairés de ces technologies. Quels sont les positionnements des firmes éditrices pour permettre en particulier à l'enseignant d'exercer son agentivité de manière raisonnée ?

Plusieurs médias proposent d'abord des visualisations des décisions algorithmiques auprès des enseignants. *Scholastic FIRST*, qui projette une agentivité restreinte à l'*analyse*, donne par exemple à voir pourquoi un élève a « sauté » plusieurs livres dans les exercices de compréhension écrite (en haut de la Figure 35) : ses scores sur les quinze premiers livres ont excédé les 70% de probabilité de réussite – seuil retenu par la firme pour inférer sa maîtrise. Dès lors, il a été directement envoyé par l'algorithme d'*adaptive sequencing* au 35^{e} livre.

Dans le cas d'*Adaptiv'Math*, l'enseignant peut visualiser les écarts entre les prévisions de succès de l'élève, calculées par ZPDES pour chaque leçon, et la réussite *a posteriori* de l'élève (Figure 35 en bas). Cela lui permet de saisir la manière dont l'algorithme ajuste ses prises de décisions : quand l'élève rencontre des difficultés alors que l'algorithme a formulé une probabilité de réussite élevée, ce dernier baisse drastiquement ses attentes à l'activité suivante, comme entre les activités A1.E4 et A1.E5. L'élève ayant réussi cette dernière activité, l'algorithme tend inversement à remonter ses attentes à l'activité A1.E6.

Quels que soient les degrés d'agentivité de l'enseignant, ces explications offrent un niveau de détail très superficiel – la cause générale du « surclassement » d'un élève sur tel livre, ou la manière dont s'ajustent les prévisions de succès d'un élève par rapport à ses réponses précédentes. Nulle explication en revanche sur les paramètres qui ont entraîné le choix du 35^{e} livre ou de telle question plutôt qu'une autre au sein de l'activité A1.E5. Dans le cas d'*Adaptiv'Math* qui projette un enseignant relativement autonome, il pourrait être utile à ce dernier de se faire un avis éclairé concernant la qualité de prédiction de ZPDES – qui paraît faible au vu des écarts contenus dans le graphe de la Figure 35 –, afin de le garder actif ou non selon certains points sensibles du programme, ainsi qu'il en a le pouvoir.

On a vu dans le troisième chapitre que des méthodes de *computer science*, comme les analyses de sensibilité, pouvaient permettre de saisir en partie les causes d'une décision de *machine learning*. Cependant les obstacles évoqués restent entiers : la multiplicité des variables présentes dans les calculs opérés par les algorithmes de *machine learning* peuvent se compter en milliers voire millions d'opérations par seconde lorsque que les bases de données des firmes sont très vastes. Ainsi Duolingo, qui dispose de la base d'usagers la plus vaste de notre corpus, n'offre aucune explication des arbitrages pédagogiques de *Birdbrain* ou de *Half-Life Regression*. En outre, l'explication d'une décision algorithmique isolée, fastidieuse pour la firme et potentiellement inintéressante pour le formateur, pose la question des aspects pédagogiques à privilégier pour l'interprétation.

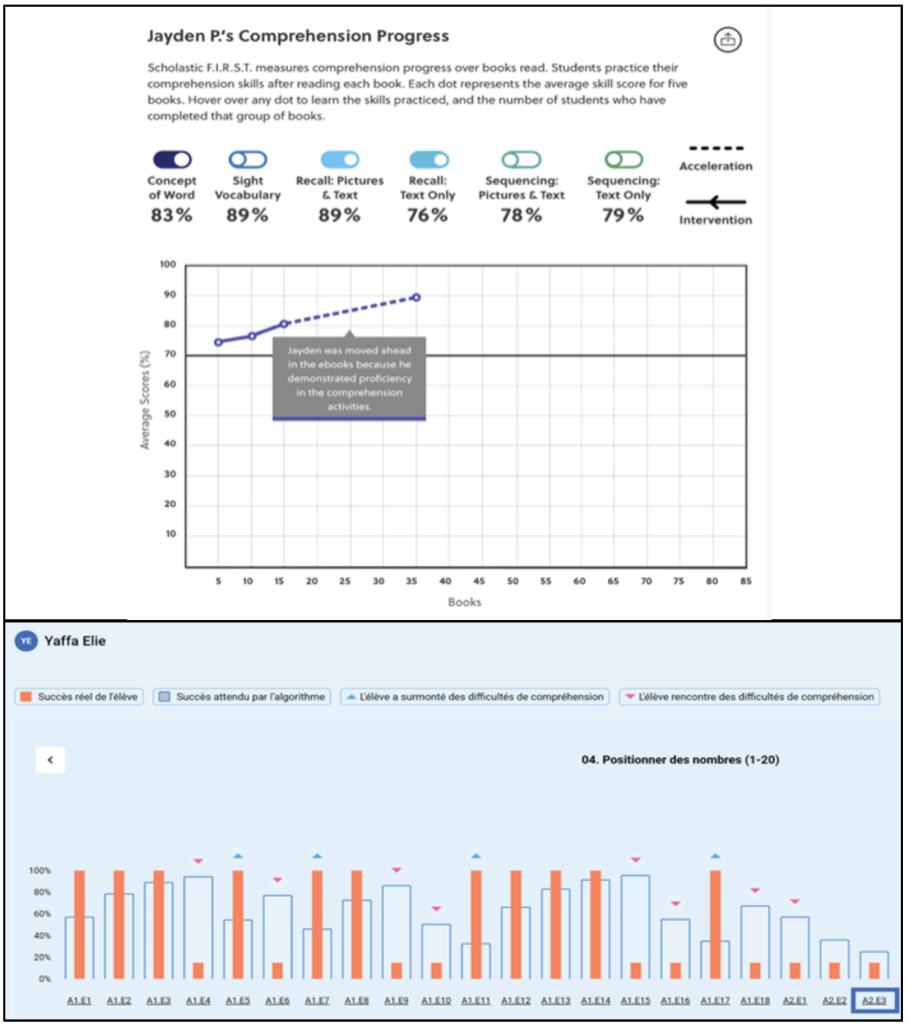


Figure 35. Visualisation des progrès d'un élève de Scholastic FIRST *en compréhension écrite (en haut) ; visualisation des écarts entre les estimations du succès d'un élève d'*Adaptiv'Math *par ZPDES et ses succès réels. Sources : Scholastic Research & Validation, 2019, p.17 ; Adaptiv'Math, 2024.*

En définitive, lorsque l'enseignant confie à un ITS la tâche d'exercer ses élèves sur le long terme, il ne peut comprendre fondamentalement pourquoi un enchaînement spécifique d'exercices a été assigné à un élève. Il doit ainsi déléguer une part importante de sa pédagogie à ces dispositifs, et leur vouer une confiance « par défaut ». Si les pouvoirs publics affirment vouloir « donner un sens à l'IA », pour reprendre la formule du député Cédric Villani, et faire rentrer ces technologies dans un cadre éthique, l'absence de directives claires pour y parvenir se vérifie dans l'offre disponible.

Faire de l'explicabilité une exigence première pour les médias adaptatifs semble induire deux conditions : d'abord que les firmes limitent les capacités prédictives de leurs algorithmes de *machine learning*, afin que leurs décisions restent « symbolisables » ; ensuite que les pouvoirs publics financent un effort important de formation des enseignants à la compréhension et au jugement critique des *analytics,* afin que l'explicabilité se traduise en retombées pédagogiques.

Néanmoins, dans la pratique actuelle, la délégation d'activités pédagogiques repose non pas sur des preuves de pertinence mais sur des croyances. Or, on a observé dans ce chapitre que les dispositifs reposaient eux-mêmes sur des présupposés pédagogiques troubles, oscillant entre constructivisme et behaviorisme, où transparaissaient des projets de société distincts. En s'imposant par des décisions

algorithmiques, aussi fines soient-elles, dans divers niveaux d'enseignement, ces projets risquent d'être naturalisés sans formes de débat possible.

Conclusion provisoire

Il est temps de revenir sur notre questionnement exposé en début de partie. En nous tenant au croisement des théories de l'industrialisation de la formation et de la trivialité, nous avons étudié la manière dont les codes communicationnels des médias adaptatifs conduisaient des stratégies industrielles, dans l'alignement ou non des industries culturelles. À partir d'un corpus de sept dispositifs, il s'agissait de dégager des permanences dans la rationalisation, l'idéologisation et la technologisation pratiquées.

Nous avons mis au jour dans un premier chapitre un idéal-type industriel tendu vers la poursuite de l'engagement. Sur un versant pédagogique, la rationalisation procède de recherches orientées vers la motivation dans les EIAH, reposant sur une alliance entre psychologie, sciences cognitives, neurosciences et intelligence artificielle. Sur le plan commercial, cet objectif se traduit dans la diversification par les entreprises des dispositifs de captation de l'attention pour maximiser les traces que l'apprenant laisse en ligne, afin de les valoriser ensuite auprès des institutions, ou tenter d'accroître, sur le circuit « BtoC », la propension des usagers à s'abonner. Cette rationalisation métrique s'observe en *back-end* par des méthodes inductives de *data mining*, inspirées de l'« épistémologie de la donnée » pour modéliser l'apprentissage, ainsi que des algorithmes de *data visualization*. En *front-end,* des dispositifs d'engagement stéréotypés, consistant dans la diversification des contenus, la gamification et la linéarité des parcours, prennent le relai. L'industrialisation *par l'engagement* est un projet techno-éducatif renouvelant le behaviorisme sur le plan pédagogique. C'est également un modèle de développement économique tendant, par la quête d'effets de réseau et la capitalisation sur les habitudes des usagers en termes de personnalisation et de formats de divertissement, à converger vers les stratégies des plateformes culturelles.

L'observation pratique de notre corpus a également mis au jour, dans un second chapitre, des lignes de crête sur lesquelles les médias adaptatifs se tiennent. Des alternatives technologiques se développent à travers des méthodes d'ingénierie écrites et déductives, plus simples à développer pour des structures anciennes et des acteurs exogènes, confortant la forme scolaire tout en prétendant pratiquer de l'*adaptive learning*. Elles exercent une concurrence à la fois économique et idéologique vis-à-vis des firmes liées à la quête d'engagement. Parmi les acteurs de notre corpus sont également apparues des dissensions sur la pertinence pédagogique d'ingrédients d'engagement comme la gamification. S'y ajoutent des positionnements divers des éditeurs sur la posture de l'enseignant à privilégier face aux algorithmes qu'elles développent, entre analyse, assignation ou encadrement. Subsistent dans tous les cas des défauts d'explicabilité des décisions algorithmiques, qui croissent avec la complexité des technologies et empêchent un débat éclairé sur les projets techno-éducatifs contenus dans ces dispositifs.

Nos avancées ont été nombreuses depuis que nous avons posé notre problématique de départ. Les constats tirés de nos réflexions partagées entre les projets scientifiques, les stratégies d'acteurs industriels et l'étude des médias adaptatifs méritent maintenant d'être synthétisés.

CONCLUSION. Vers une adaptativité au service de l'individuation

Il est temps de revenir à notre questionnement de départ, qui trouvait son origine dans les citations d'Henri Claparède et d'Edward Thorndike. Derrière la critique commune de la standardisation de l'éducation, les deux penseurs affichaient au début du XX^{e} siècle des visions différentes sur les moyens pour parvenir à un « sur-mesure pédagogique » : l'un évoquait une différenciation émanant d'un enseignant « tailleur » qui disposerait implicitement d'outils adéquats ; l'autre une différenciation opérée par un média éducatif (à l'époque le livre), sans intervention de l'enseignant. Ils annonçaient ainsi un débat intestin au sein de la recherche en technologies éducatives entre technologisations additive et substitutive. L'arrivée du *Big Data* et de techniques de fouille de données depuis quelques années a bouleversé de nombreuses positions épistémologiques et relancé l'espoir d'un « sur-mesure de masse » appliqué à des domaines variés de la pédagogie. Notre questionnement était non seulement de savoir comment le déploiement d'algorithmes d'apprentissage adaptatif revisitait cet idéal et ce débat, mais aussi de déterminer dans quelle mesure il était susceptible de véhiculer un paradigme éducatif singulier, cohérent et transversal aux cultures françaises et états-uniennes et aux secteurs éducatifs.

Nous formulions l'hypothèse selon laquelle ces technologies algorithmiques traçaient un paradigme industriel bel et bien transversal et antinomique vis-à-vis de la forme scolaire. Restait à déterminer la manière dont ces dispositifs projetaient l'apprentissage de l'élève ou la place de l'enseignant face à la machine, entre la valence potentiellement behavioriste et substitutive des techniques de *data mining*, et les directions socio-constructivistes et additives héritées des technologies numériques.

Restait également en suspens la morphologie industrielle de ce nouveau paradigme, les nouveaux rapports de force qu'il serait à même de favoriser, et ainsi le projet de société qu'il représenterait face à la forme scolaire : conservatisme social par le cloisonnement de chaque élève dans un parcours, annihilation de la méritocratie républicaine par la responsabilisation non plus de l'élève mais de la machine, ou projet connexionniste d'émancipation pour des individus dont les besoins seraient priorisés ?

Pour cerner les dimensions techniques, pédagogiques et socio-économiques de ce projet adaptatif, nous avons observé trois foyers différents où il se constitue : les recherches académiques, les rapports socio-économiques entre acteurs industriels, et les médias adaptatifs où est codifiée la pédagogie.

Principaux enseignements

Sur le premier niveau, nous avons procédé à une étude sélective d'anthologies retraçant l'histoire des technologies éducatives, et à des entretiens avec des chercheurs français et états-uniens, pour saisir les évolutions techniques et paradigmatiques dans le domaine de l'adaptativité. Est apparue l'ampleur du chemin parcouru sur le plan technique, de l'adaptation sommaire permise par l'enseignement programmé aux tuteurs informatiques reposant sur l'IA symbolique, puis aux EIAH usant d'ontologies, de techniques de fouille de données et de grands modèles de langage. Avec chacun de ces artefacts, les champs d'adaptativité se sont étoffés : adaptation des tâches à réaliser dans les machines à enseigner, premières adaptations des feedbacks et de la difficulté des questions dans les tuteurs modélisant la maîtrise de l'apprenant, puis adaptation des remédiations (*adaptive content*),

du contenu des questions (*adaptive assessment*) et de l'ordre des notions (*adaptive sequencing*) dans les EIAH. Ceci grâce à la collecte et au traitement en continu des données des élèves. Les modes pédagogiques ont eux aussi varié, d'un « *drill and practice* » behavioriste à un tutorat d'abord très encadré puis assoupli par des formes de pédagogie actives et collaboratives, sous l'influence du constructivisme et de ses courants affiliés. La singularité des projets adaptatifs, à rebours de la forme scolaire collective et standardisée, est ainsi apparue clairement.

Toutefois, les divergences théoriques voire philosophiques évoquées plus haut pénalisent de longue date l'impulsion d'un paradigme éducatif cohérent. Depuis le tournant des sciences cognitives se font face deux visions différentes sur la place à accorder au guidage automatique dans la situation pédagogique : d'un côté une approche *technocentrée*, héritière du behaviorisme, cherchant à coder des capacités pédagogiques autonomes dans la machine, et de l'autre une approche *anthropocentrée*, d'influence constructiviste, recherchant une assistance adéquate des élèves et du formateur dans leurs usages. Les communautés actuelles des *Learning Analytics* et de l'*Educational Data Mining* réactualisent ces débats. L'une prône un usage de la technologie en appoint des pratiques pédagogiques, vers davantage de personnalisation et de collaboration entre pairs, quand l'autre recherche la substitution de l'algorithme à l'enseignant sur de nombreuses tâches. Si une technologisation totalement substitutive est une gageure, le degré d'agentivité des publics éducatifs vis-à-vis des technologies amène des positionnements divers sur plusieurs problématiques : le dilemme entre le volume et l'« épaisseur » des données récoltées, la nécessité de contenus de remédiation variés ou l'explicabilité des décisions algorithmiques auprès des publics. Au mieux ces différences de vues offrent-t-elles des modèles « en puissance » pour les acteurs de l'éducation et de la formation.

Dans ce deuxième foyer de constitution paradigmatique que sont les systèmes éducatifs et de formation, nous avons questionné la manière dont les médias intégrant des algorithmes d'apprentissage adaptatif favorisaient l'émergence d'une socio-économie cohérente, conduisant certains acteurs à influer sur la définition des offres médiatiques et d'un potentiel paradigme. Nous avons observé les stratégies d'acteurs industriels sur trois niveaux. Le premier, celui des pouvoirs publics, a mis au jour des politiques volontaristes en matière d'IA éducative – un courant technologique vague auquel tend à être assimilé l'*adaptive learning.* Par une veille documentaire et une analyse de discours publics en faveur du déploiement des médias algorithmiques, nous avons révélé des permanences entre les deux pays. Les projets techno-éducatifs tracés par les textes gouvernementaux formulent des attentes d'obédience socio-constructiviste. Elles visent une différenciation de l'éducation et de la formation au service des besoins d'apprenants de tout âge et de tout parcours, des pratiques d'apprentissage collaboratives et autonomes, ainsi que l'affirmation des évaluations formatives. Les pouvoirs publics défendent également la figure de l'enseignant, qui devrait voir ses pratiques pédagogiques « augmentées » par l'analyse et la visualisation automatisées des apprentissages numériques. Sur le plan organisationnel est espérée une industrialisation empreinte des préceptes du *Social Efficiency Movement*. L'administration des établissements doit ainsi échoir aux acteurs éducatifs « de terrain », dont des professions spécialisées liées à l'ingénierie pédagogique, et s'ouvrir aux acteurs innovants issus de l'industrie. La vocation démocratique de ces technologies est également un point d'attention des gouvernements, veillant au respect des libertés individuelles et notamment à la confidentialité des données numériques traitées. Toutefois cette posture proactive est apparue affaiblie par la marge de manœuvre budgétaire contrainte d'un grand nombre d'établissements publics, de même que les doutes subsistant sur l'utilité pédagogique des technologies numériques dans les collectivités et les organisations éducatives.

Ces frictions tendent à rendre l'intervention des acteurs privés d'autant plus cruciale dans le déploiement de ces technologies.

C'est sur ce deuxième niveau, celui des firmes impliquées, que nous avons observé la formation d'une socio-économie cohérente autour de trois domaines générateurs de valeur : la conception d'algorithmes d'apprentissage adaptatif, la production de contenus pédagogiques et la distribution des médias adaptatifs. Notre analyse socio-économique a d'abord fait apparaître des différences nombreuses entre l'industrie française, faisant l'objet d'interventions substantielles de l'État, et son homologue états-unienne, où financiarisation et marchandisation sont plus marquées.

Néanmoins des permanences ont émergé dans les stratégies adoptées par chaque type d'acteur. En amont de la filière, des start-ups prennent en charge la majeure partie des recherches algorithmiques, avec le concours régulier de laboratoires. Elles bénéficient, *a minima,* des infrastructures *cloud* des *Big Tech,* et, au plus, de leurs modèles de langage et autres briques algorithmiques. Les acteurs anciens du numérique éducatif adoptent des positions volontaristes en se repositionnant pour éditer ce type de média, de façon plus ou moins ambitieuse selon leurs ressources. Des éditeurs scolaires et organismes de formation se positionnent également en éditant leurs LMS, capitalisant sur leurs catalogues de contenus avec le concours plus ou moins marqué de start-ups. En aval, du côté de la distribution, des plateformes d'intermédiation publiques comme privées sont devenues des acteurs incontournables de l'accès aux publics éducatifs – étape cruciale dans ces filières, tant du point de vue technique que commercial. Les *Big Tech* sont quant à eux présents sur les principaux domaines générateurs de valeur, vendant leurs technologies en amont et éditant leurs propres plateformes en aval, tant auprès des particuliers que des organisations éducatives, sous le contrôle accru des pouvoirs publics inquiets de leurs pratiques hégémoniques. Ainsi l'évolution des rapports de force industriels font apparaître trois tendances transversales : l'importance des start-ups *EdTech* en tant que moteurs d'innovation pour les offres adaptatives, celle des plateformes de distribution, et les stratégies hégémoniques mais de plus en plus régulées des *Big Tech*.

Selon ces dispositions communes, des singularités se manifestent selon les deux pays. Ainsi, les circuits états-uniens de distribution aux institutions et aux consommateurs tendent à converger vers le modèle du courtage, consacrant la centralité des plateformes d'intermédiation privées. En France, les deux circuits gardent des spécificités marquées, la présence des pouvoirs publics dans l'intermédiation auprès des établissements maintenant le modèle du club et la centralité des éditeurs concentrés dans la filière. Dans les deux pays se précisent en définitive les modalités d'une lutte entre éditeurs concentrés et acteurs de l'intermédiation pour s'installer au centre d'une nouvelle socio-économie.

Au cœur de cette tectonique des plaques industrielles se jouent des enjeux politiques où émergent des projets de société concurrents. Deux tendances transparaissent. D'abord, l'ensemble des firmes tendent à s'accommoder aux attentes étatiques en termes de technologisation additive – *a fortiori* lorsque les firmes sont dépendantes des subventions étatiques comme les start-ups françaises, ou soumises aux régulations des législateurs comme les *Big Tech.* En outre, la concurrence entre les modèles du club et du courtage induit celle entre leurs projets de sociétés afférents. Le premier véhicule le cloisonnement de l'expérience éducative à l'échelle de l'organisation ou d'une communauté d'abonnés. Visant à maximiser la présence des adhérents, il promeut non pas une vision démocratique ou massifiée mais plutôt une solidarité communautaire (Mœglin, 2005, p.232).

Le second modèle porte l'idéal d'une assistance personnalisée au cœur d'une conception réticulaire et connexionniste de la société (ibid., pp. 233-234). Il vise à maximiser la participation des individus dans les dispositifs techniques selon des effets de réseau. Ce projet est traversé par le « nouvel esprit du capitalisme » théorisé par Luc Boltanski et Ève Chiapello (1999), qui tend à s'appliquer au travail, aux relations privées et aux politiques publiques. Tous deux empreints d'un individualisme fort, ces modèles socio-économiques se distinguent en définitive sur une tension inhérente à l'individu-apprenant, entre quête d'appartenance à une communauté et recherche de prestations « sur-mesure ».

Des volontés réformatrices ont été observées au troisième niveau, celui des réseaux de professionnels constitués autour du mouvement de l'*EdTech.* Les associations à but non-lucratif et les incubateurs constituent des agents très importants pour la rencontre de professionnels technophiles, travaillant tant dans les systèmes éducatifs et de formation que dans les firmes. Par l'analyse de leurs discours publics, nous avons mis au jour le soutien voire l'influence qu'ils exerçaient sur les projets technopédagogiques des pouvoirs publics. Ils tendent à se focaliser sur l'assistance à l'enseignant, ainsi que la diffusion de compétences liées à l'usage des technologies. Ces réseaux visent ainsi une légitimation de l'utilité pédagogique de ces médias dans l'espace public, et la diffusion d'habitus d'entrepreneurs et d'acheteurs parmi les publics éducatifs. S'y jouent également des tentatives de codification des professionnalités par certains acteurs, à l'image des *Big Tech*, à la fois engagés dans ces communautés et lançant leurs propres réseaux.

Pour déterminer si les offres médiatiques sont en mesure de trancher les tensions paradigmatiques présentes à l'échelle socio-économique, notre troisième partie a cherché à mettre au jour un modèle industriel cohérent, par l'étude d'un corpus diversifié de médias adaptatifs. Nous avons recherché les trois marqueurs de l'industrialisation, à commencer par l'idéologisation, via une analyse sémio-discursive des discours d'escorte des firmes éditrices. Nous avons recherché des indices de rationalisation et de technologisation dans les algorithmes, identifiés à partir des productions scientifiques des firmes, et sur les interfaces des usagers, via une étude techno-sémiotique des « formes » stéréotypées à l'écran. Est apparu un modèle que nous avons nommé « industrialisation *par l'engagement »*. Sa cohérence, sa transversalité et sa singularité devaient être jaugées avant de pouvoir le qualifier de paradigme.

La cohérence et ses hypothèques

Sa cohérence a été attestée à plusieurs niveaux. D'abord sur le plan de l'idéologisation, il prétend répondre aux deux promesses d'un tutorat personnalisé et d'une assistance auprès de l'enseignant. Le moyen privilégié pour ce faire est la maximisation d'un indicateur cardinal et polysémique, l'« engagement ». Son acception pédagogique découle d'une double inférence. Celle-ci fait, premièrement, de la mise en place d'une pédagogie « positive » le moyen pour maximiser les interactions de l'usager dans le produit, équivalant à son engagement. Elle fait ensuite de l'engagement le corollaire de gains d'apprentissage pour l'élève et donc de l'efficacité pédagogique du produit. Cette dernière équivaudrait donc, par relation transitive, à des mesures du temps passé et de la quantité d'interactions de l'individu avec le produit – un motif fort de rationalisation. Concernant la technologisation, les firmes visent ainsi à personnaliser les dispositifs de motivation en *back-end,* par des méthodes de fouille de données servant à modéliser l'apprentissage, ainsi que des algorithmes de *data visualization*. Les premières sont tendues vers la sélection de contenus pédagogiques suffisamment motivants pour les apprenants, quand les seconds

doivent donner aux enseignants des capacités de diagnostic plus approfondies. Ils véhiculent ainsi des habitus d'élèves friands de divertissement et rétifs à l'effort, et d'enseignants relayant des normes statistiques et productivistes. Cette rationalisation est concrétisée en *front-end* par des dispositifs d'engagement stéréotypés, consistant dans la diversification des contenus, la gamification et la présentation linéaire du parcours. Ce modèle cohérent engage donc une organisation technologique de l'éducation, oscillant entre addition et substitution à l'enseignant selon les contextes d'usage. Il ré-instrumente cependant puissamment le behaviorisme en se concentrant sur le conditionnement de comportements attendus.

Le projet apparaît aussi cohérent et transversal sur un plan économique. Que ce soit dans le modèle du club ou du courtage, l'engagement est visé à travers la maximisation des traces collectées. Le but est de valoriser celles-ci auprès des institutions en tant qu'indices de l'efficience des produits, ou de tenter d'accroître, sur le circuit « *BtoC* », la propension des usagers à s'abonner, à consommer des publicités ou à acheter des bonus « à l'acte ».

La cohérence de ce potentiel paradigme est apparue plus friable sur certains points, qui réveillent des spécificités entre les cultures nationales, les filières et les niveaux éducatifs. D'abord sur un plan technique, des méthodes alternatives à la fouille de données comme des systèmes-experts codés depuis plusieurs décennies peuvent conduire certaines caractéristiques de ce projet. D'autres EIAH n'opérant pas d'adaptations dynamiques et cantonnés à des aspects comme la mémorisation prétendent au syntagme de l'*adaptive learning,* concurrençant les dispositifs adaptatifs auprès des publics éducatifs. Ensuite les limites techniques de ces médias, comme les angles morts contextuels et l'unicité des méthodes d'enseignement, sont de nature à en détourner des publics en attente d'une formation poussée. Enfin des dissensions d'ordre théorique entre behaviorisme et constructivisme surplombent les offres des firmes sur deux principaux aspects. Premièrement, le recours à la gamification cristallise des divergences entre les discours publics des firmes et la technologisation de leurs médias. Alors que des acteurs de l'autoformation ou de la formation continue assument le lien entre ce mode énonciatif et le conditionnement opérant skinnerien, une forme d'euphémisation est palpable chez des acteurs de l'enseignement primaire et secondaire qui cultivent une proximité avec la vision constructiviste des pouvoirs publics. Deuxièmement, les postures codifiées de l'enseignant face aux algorithmes révèlent aussi des sensibilités théoriques différentes chez les firmes. Nous avons identifié trois régimes d'agentivité pour le formateur, de l'*analyse,* ne lui laissant qu'un rôle consultatif, à l'*assignation* et *l'encadrement,* le plaçant en « chef d'orchestre » au sein de l'EIAH. Le choix de ces régimes varie selon l'autonomie traditionnelle laissée au professeur dans les cultures nationales. La question de l'explicabilité des décisions algorithmiques met enfin les médias aux prises avec la superficialité des explications et la capacité réelle des enseignants à s'en saisir.

Une singularité relative vis-à-vis de la forme scolaire

La singularité du projet est enfin attestée vis-à-vis de la forme scolaire collective. Si l'on reprend les huit caractéristiques avancées par Maulini et Perrenoud pour décrire cette dernière, les différences sont patentes. Les médias adaptatifs mettent d'abord au défi la clôture de l'institution éducative par rapport au reste de la société, en permettant des usages en classe et en dehors des institutions. D'autre part, le contrat didactique entre un formateur et un apprenant est revisité en profondeur par la présence tierce de l'algorithme, dans un rôle de tuteur, sélectionnant des activités d'apprentissage

et de remédiation codifiées par les firmes éditrices. Du côté de l'apprenant, une simplification behavioriste de l'apprentissage en termes statistiques se manifeste dans les traitements algorithmiques. Le principe de « discipline » de la forme scolaire, supposant que l'apprentissage soit « laborieux », est fortement amendé par la quête du « juste » effort qui permettrait de maintenir la motivation et le « plaisir d'apprendre » de l'élève. Enfin la référence à des normes d'excellence et des critères d'évaluation est ambivalente : elle est très palpable du côté de l'enseignant qui, en surveillant les apprentissages, veille à leur correspondance avec des normes statistiques et productivistes. Elle est en revanche moins visible du côté de l'élève, qui en évoluant dans un parcours cloisonné et « *feel good* », ne doit pas éprouver un potentiel sentiment d'échec.

La filiation entre ce paradigme et celui qui accompagnerait, selon de nombreux chercheurs, le déploiement de technologies numériques dans l'éducation est également palpable. Il s'inscrit pleinement dans certains antagonismes qu'avait déjà synthétisés en 2005 Pierre Mœglin entre « nouveau » et « ancien » paradigmes éducatifs. En effet, il se focalise également sur l'apprentissage plutôt que l'enseignement, sur le « sur-mesure » au lieu du standard, et sur la validation de compétences plutôt que le diplôme (Mœglin, 2005, p.188). Les autres tendances du numérique éducatif ajoutées par Laurent Petit, telles que la pression nouvelle des parents ou la formation tout au long de la vie (Petit, 2018) se retrouvent aussi dans les médias adaptatifs. Les parents reçoivent des identifiants propres sinon des mails réguliers des firmes pour suivre les progrès de leurs enfants, et les applications d'autoformation comme Duolingo s'inscrivent dans la perspective d'une « montée en compétences » tout au long de la vie.

Toutefois, l'industrialisation *par l'engagement* trace une voie propre sur plusieurs aspects. Nombre de technologies numériques se distinguent de la forme scolaire en se concentrant sur l'appropriation par les apprenants de savoirs scientifiques et narratifs plutôt que sur la transmission exclusive des premiers, sur l'autonomie des élèves plutôt que sur leur hétéronomie, sur l'individualisation et non sur la massification, et sur une pédagogie constructiviste plutôt que behavioriste. Les médias adaptatifs suivent à ces différents égards un agenda plutôt réactionnaire : ils tendent vers un retour à la transmission verticale de savoirs scientifiques, à un guidage étroit et au renouvellement de préceptes behavioristes, tout en opérant, grâce à leurs algorithmes, une tentative de conciliation entre la massification et l'individualisation (entendue dans un sens pédagogiquement faible). Leur projet paradigmatique tend à rompre avec la tendance à l'ouverture des médias numériques antérieurs vers des pratiques horizontales et collaboratives, pour en revenir à une fermeture paradoxale des ITS sur la relation tutorale entre l'élève et le système informatique, intégrant l'enseignant de façon essentiellement périphérique.

Pour toutes ces raisons, il apparaît encore précipité de désigner l'industrialisation *par l'engagement* comme un paradigme. Ce projet éducatif et socio-économique présente des traits saillants, de même que des caractéristiques encore instables à nos yeux. Ces apories l'empêchent de représenter à ce jour une alternative solide à la forme scolaire. Ceci d'autant moins qu'il semble, selon les niveaux éducatifs, pouvoir être adapté à l'organisation collective et verticale existante.

L'hypothèse de l'établissement d'un éventuel paradigme sur les bases de ce projet n'est cependant pas écartée, loin de là. Il serait intéressant de continuer à observer les évolutions de ses diverses caractéristiques dans les offres de firmes des ITS en France, aux États-Unis et dans d'autres pays aux socio-économies différentes. Cela permettrait de faire évoluer et de préciser nos résultats, en particulier sur des aspects comme l'universalité du projet.

Il serait notamment intéressant d'interroger plus avant, en partant de la philosophie politique, l'idéologisation des médias adaptatifs, tant dans les nouvelles directions technologiques qu'ils pourraient prendre que dans l'évolution des positionnements éducatifs au sein de la société civile. Les élections récentes de régimes conservateurs questionnent les idéologies qui pareront les technologies éducatives à l'avenir. À rebours de pédagogies progressistes et artisanales, le behaviorisme s'en trouvera-t-il renforcé voire légitimé par les responsables politiques ? Au contraire, les attaques croissantes contre la technocratie pourraient-elles affecter le déploiement de dispositifs algorithmiques de plus en plus puissants dans l'espace social ? Si, aux États-Unis, le recours aux technologies algorithmiques a jusqu'ici semblé susciter un consensus bipartisan, l'administration de l'éducation attire au contraire des opinions radicalement différentes – notamment des critiques portées par le camp républicain sur la légitimité de l'intervention fédérale. Dès lors, il n'est pas à exclure que, loin d'un consensus officiel autour d'un paradigme socioconstructiviste des deux côtés de l'Atlantique, l'éducation et la formation fassent de plus en plus l'objet de choix éclatés en matière de médias éducatifs.

Par ailleurs, notre questionnement paradigmatique mériterait d'être mené, au sein des sociétés, à des niveaux méso et micro, par l'étude circonstanciée des usages dans des organisations éducatives adoptant des ITS. Cela permettrait d'enrichir nos résultats d'observations localisées qui donneraient d'une part à voir des concordances ou des incohérences avec les stratégies « macro » des acteurs industriels. Ces analyses permettraient d'autre part de révéler des discordances entre des pratiques liées au médias adaptatifs et des dimensions essentielles de la forme scolaire. À l'origine de ces blocages apparaitraient ce que Laurent Petit a nommé des « nœuds » – tels que l'évaluation sommative à l'université (Petit, 2018) –, qui feraient tenir l'ancien paradigme dans les organisations, jusqu'à leur dénouement et l'adoption d'un nouveau modèle.

La question de la convergence avec les industries culturelles

Se pose également, sur le plan socio-économique, la question de la singularité de l'industrialisation *par l'engagement* vis-à-vis des stratégies entourant les médias algorithmiques dans les industries culturelles. Cette question recoupe *in fine* celle de la convergence entre industries éducatives et de la culture. Parmi les diverses tendances paradigmatiques qui soutiennent cette convergence et que nous avions soulevées au quatrième chapitre, les médias adaptatifs entérinent la marchandisation croissante de l'éducation. L'engagement tend en effet à être employé autant par des médias éducatifs que par des plateformes culturelles comme un indicateur de leur capacité à développer leurs bases d'utilisateurs et à les monétiser. Les modèles socio-économiques issus de la théorie des industries culturelles ont par ailleurs été utiles pour caractériser ce paradigme, et notamment la prééminence des modèles du club et du courtage, avec des déclinaisons différentes selon les circuits de distribution.

Une seconde tendance paradigmatique à la convergence est la multiplication des innovations technologiques, comme les médias basés sur des algorithmes. Les acteurs des technologies, qu'il s'agisse d'équipements ou de logiciels, diversifient et renouvellent en permanence leurs offres culturelles dans des marchés très concurrentiels et soumis à l'obsolescence. Les médias adaptatifs s'inscrivent dans une tendance similaire dans l'éducation, que Louis Porcher a décrit comme une « économisation éducative », c'est-à-dire le recours croissant à des outils et médias coûteux dans les salles de classe (Porcher, 1998, p.2), à l'image des outils et médias numériques (calculatrices

connectées, tablettes, tableaux blancs interactifs). En tant qu'infrastructure, les algorithmes d'apprentissage adaptatif profitent de la diversification des supports et terminaux, s'affinant avec la diversité des terminaux éducatifs et des données d'usage collectées sur les apprenants. Le revers de la médaille est la possibilité que de nouvelles innovations, qui dépasseraient notamment certaines limitations des algorithmes actuels, rendent d'autres formes médiatiques plus désirables à l'avenir.

Une autre tendance paradigmatique qui est apparue avec acuité dans les médias adaptatifs et ceux de la culture est le recours à l'individualisation. L'analogie entre individualisation des consommations culturelles et adaptativité a semblé pertinente à plusieurs égards. D'abord, les médias adaptatifs capitalisent sur des modélisations des préférences individuelles selon la logique du *ranking,* et notamment des techniques de filtrage collaboratif usuelles dans les plateformes culturelles. Ils s'appuient également sur les habitudes des usagers en termes de formats de personnalisation, comme le *quantified self* et la gamification. Cependant, des points de friction sont apparus chez les éditeurs rejetant ces dispositifs exogènes, pour des raisons politiques et/ou de positionnement commercial.

Une quatrième tendance rapprochant l'éducation et la culture, qui s'est manifestée dans les trajectoires de nombreux éditeurs de médias adaptatifs, est la financiarisation. Ainsi les investissements venus des fonds de capital-risque se sont grandement accrus en France et surtout aux États-Unis avec le développement de l'*EdTech*, tandis que des start-ups ont été acquises par des acteurs concentrés de l'édition scolaire ou du numérique éducatif. Par ailleurs, un acteur historique du secteur, Duolingo, est parvenu à introduire son capital en bourse. Comme dans les industries culturelles, le développement de l'actionnariat est un moyen de financement central face à la pression concurrentielle, l'augmentation des coûts de R&D et de marketing.

La hausse générale des dépenses en marketing est une dernière tendance symptomatique de la convergence entérinée par les médias adaptatifs. Les enjeux de *branding* et d'image sont cruciaux, comme dans le cas des plateformes culturelles, quand la quête d'effets de réseaux est clé, tant pour des motifs techniques (l'entraînement des algorithmes) que commerciaux (l'accroissement des revenus tirés des abonnements et licences).

Les industries éducatives conservent toutefois des spécificités fortes. La marchandisation poussée des industries culturelles, ayant débouché depuis plusieurs décennies sur cinq modèles socio-économiques plus ou moins stabilisés, demeure moins marquée et variée dans l'éducation. Dans le circuit institutionnel, le système ancien des licences par élève et enseignant, pratique ancienne du numérique éducatif, est ainsi conservé et pris en charge par les pouvoirs publics et non par les publics éducatifs eux-mêmes, contrairement au domaine culturel, davantage privatisé. Continue ainsi de prévaloir, derrière le financement des médias adaptatifs, une logique de service public qui accompagne l'industrialisation de l'éducation depuis les manuels scolaires. Cette logique milite pour une lecture non seulement économique mais aussi politique de l'industrialisation *par l'engagement* et de ses alternatives.

L'adaptativité algorithmique, renouvellement de la méritocratie ?

Nous interrogions en introduction la rupture du projet porté par les médias adaptatifs vis-à-vis de la méritocratie, idéal concomitant de la forme scolaire. Nous nous demandions si la démocratisation

de systèmes de tutorat adaptatifs pouvait permettre d'égaliser les chances, dans une veine méritocratique, ou au contraire conduire à rompre avec la méritocratie sur ses aspects les plus critiqués (*hubris* des « vainqueurs » et humiliation des « perdants », reproduction forte des inégalités sociales). Nous faisions l'hypothèse selon laquelle les médias adaptatifs portaient les germes d'une annihilation de la méritocratie républicaine : un éventuel échec scolaire pèserait non plus sur l'élève mais sur la machine, qui aurait pêché dans l'attribution des contenus d'apprentissage et de remédiation. En outre, la logique de compétition entre les individus semblait en mesure d'être abolie par des parcours d'apprentissage sur-mesure.

Cependant, notre thèse a mis au jour une autre perspective : celle d'une réaffirmation de la méritocratie, accompagnée de l'invisibilisation technologique de ses aspects les plus critiqués. L'*adaptive learning,* pris au sens de courant industriel paré d'imaginaires stéréotypés dans l'espace médiatique, est plutôt apparu comme un nouvel avatar dans la quête méritocratique de dispositifs d'égalité des chances.

Les médias adaptatifs pourraient en effet pousser plus loin la constitution d'inégalités que les médias standardisés de la forme scolaire. Ils supposent d'une part une tendance accrue des administrations éducatives à diagnostiquer les performances des élèves par rapport à des normes statistiques. D'autre part, la focale individualiste de l'apprentissage tutoré est de nature à amplifier la logique méritocratique en faisant avancer plus loin les élèves les plus doués. Ceux-ci sont en mesure de se hisser dans une discipline sans devoir attendre le reste de la classe, comme dans la configuration traditionnelle d'un cours collectif et simultané.

La méritocratie se verrait par là-même euphémisée sur plusieurs de ses aspects délétères, que de nombreuses voix ont dénoncé, de Michael Young à Michael Sandel. Le premier est le sentiment d'échec pour les élèves qui échouent aux tests et épreuves d'admissions, et l'impression latente de ne devoir cet échec qu'à eux-mêmes. Comme nous l'avons dit, les EIAH adaptatifs, par le cloisonnement des parcours individuels, tendent à réduire voire dissimuler aux yeux de l'élève son écart à la norme, par une adaptation de la difficulté et des formats « *feel good* ». Plus encore, la promesse des firmes de permettre à tous de réussir, même aux élèves les plus en difficulté, grâce à des parcours optimaux au regard de leurs besoins, vise à traiter le problème de la marginalisation d'une partie des citoyens (comme les élèves « décrocheurs » ou les adultes non-diplômés).

Un autre reproche fait par Michael Sandel à la méritocratie est, derrière l'*hubris* des « vainqueurs », le profond mal-être d'individus soumis à une pression institutionnelle et/ou parentale dès les premières années de leur scolarité, pour performer davantage que leurs pairs. Leur réussite se fait régulièrement au prix de la construction de leurs aspirations personnelles et de leur santé mentale (Levine, 2006). En invisibilisant la logique de compétition entre individus par des espaces d'apprentissage cloisonnés et voulus bienveillants, les ITS tendraient à diminuer la pression pesant sur les élèves.

Toujours est-il que cette invisibilisation n'est que de façade, et les ITS poussent plus loin la surveillance métrique des apprentissages par les enseignants, administrateurs d'établissements, voire les parents. La logique de compétition continue en outre de régner dans les processus d'admission de l'enseignement supérieur et d'embauche sur le marché du travail. Il nous semble donc que le projet tende, en déléguant la différenciation de parcours d'apprentissage à la technique, à euphémiser la méritocratie.

Il réactualiserait en fait le projet de James Bryant Conant, président de Harvard de 1933 à 1953 (Sandel, 2021, pp.242-245). Ce dernier, espérant en finir avec la reproduction d'élites aristocratiques dans l'admission aux grandes écoles, prônait la refonte du système scolaire américain en vaste « machine à trier ». Celle-ci devait être capable de promouvoir les enfants de toutes les classes sociales selon leurs seules capacités intellectuelles. Ce discours revêtait une certaine critique de la forme scolaire existante dans la mesure où il déniait aux notes scolaires la capacité de révéler les capacités réelles des enfants – la maîtrise de disciplines scolaires donnait selon lui un avantage aux élèves formés dans les meilleurs établissements. Il appelait de ses vœux le déploiement d'une variante d'un test de QI utilisé par l'armée américaine pendant le Première Guerre mondiale : le *Scholastic Aptitude Test* (SAT). Sous l'influence de ses prescriptions, l'administration de tests standardisés s'est généralisée dans les procédures d'admission des grandes écoles et universités du pays. Le *No Child Left Behind Act* a instauré les tests standardisés plus en amont dans le système scolaire pour réhausser le niveau global du système éducatif et identifier les écoles en situation d'échec – une autre forme de tri bureaucratique. Devant l'incapacité de ces évaluations statiques à permettre davantage de mobilité sociale et à réhausser le niveau général du système éducatif, la « machine à tri » appelée de ses vœux par James Bryant Conant tendrait aujourd'hui à reposer sur des évaluations adaptatives, à partir de dispositifs algorithmiques de collecte et de traitement de données massives.

Ces quelques observations d'ordre politique mériteraient d'être explorées plus avant, en étudiant d'une part les évolutions des fonctionnalités et de la place des médias adaptatifs dans les systèmes éducatifs français, états-uniens, et d'autres pays aux cultures éducatives différentes. Il serait d'autre part intéressant, avec le recul des années, de mesurer la mobilité sociale réellement permise par le recours à des médias adaptatifs tout au long de la vie scolaire, universitaire et/ou professionnelle des individus. Nous faisons l'hypothèse que ces technologies apportent finalement peu de chances de progrès et contribuerait au contraire à embarquer les élèves, au-delà de leurs parcours d'apprentissage, dans des orientations professionnelles cloisonnées. Des risques existent, notamment avec l'existence de biais de genre ou de race dans de nombreux jeux de données. Des dérives sont également possibles, comme le recours, par des organisations éducatives ou des employeurs, à des données d'apprentissage en amont d'examens, de processus d'admission ou d'embauche, afin de prédire la réussite d'un individu dans tel ou tel cursus ou tel ou tel métier. Si la généralisation de telles pratiques paraît encore abstraite grâce aux législations en vigueur, des usages détournés seraient potentiellement critiques pour la vie des citoyens et la société.

Pour une éthique de l'engagement

Pour clore notre propos et ouvrir d'autres perspectives de recherche, il nous importe de questionner la possibilité de mettre les algorithmes d'apprentissage adaptatif au service d'autres projets de société que ceux soulevés à travers l'industrialisation *par l'engagement*. Il nous semble en particulier intéressant de reposer les termes du débat en redéfinissant l'objectif d'engagement qui guide actuellement les firmes du numérique éducatif. Son sens industriel, équivalent à la motivation de l'apprenant à revenir sur le média, est instrumental, visant l'efficacité pédagogique de l'outil. Promouvant une rationalité en finalité plutôt qu'en valeur, il exclut ainsi une réflexion sur les fins politiques de l'éducation. Cet escamotage procède à nos yeux d'une dépréciation ancienne du sens et de la valeur du travail dans les sociétés occidentales, qui se répercute sur l'école. Tout au long de sa vie, l'individu est tenu d'apprendre et de se former dans des sociétés de plus en plus tournées vers

la quête d'efficience et d'adaptation au changement, sans réelle réflexion sur les fins humaines, politiques et écologiques à atteindre au niveau individuel et collectif.

Dans *L'Éloge du carburateur*, le philosophe états-unien et réparateur de motos Matthew B. Crawford réhabilite l'engagement comme le lieu intime de construction de sens personnel, permettant de dépasser l'opposition courante entre travail et loisir. Qu'il y ait présence ou non d'une gratification extérieure, il s'agirait pour l'individu de « prendre prise » sur le monde extérieur en trouvant dans une activité une motivation intrinsèque – concept que nous avons déjà vue chez de nombreuses firmes qui prétendent la viser. Ainsi entendu, l'engagement ne peut être conditionné par une autorité extérieure, car il émane de la subjectivité de chaque individu : la perception, dans une activité donnée, d'une poétique liée à ce qu'il considère comme une « vie bonne » émane de facteurs multiples dans l'histoire personnelle d'un individu, impossibles à prévoir ou à mettre en données. Crawford met alors en avant dans l'engagement la poursuite par l'individu de sa propre excellence dans un domaine, et ce faisant de sa dignité au sein d'une communauté.

Cette acception de l'engagement, tournée non pas vers le mérite mais vers la dignité, nécessite de sortir d'un paradigme économique de l'attention, qui considère cette dernière comme un capital à allouer de façon optimale entre des objets multiples. Cette focale est certes heuristique face à la profusion de contenus culturels et éducatifs en ligne, mais demeure instrumentale. Il faut lui ajouter une autre perspective axée sur la construction personnelle et collective de significations politiques dans l'éducation. Cela nécessite de rechercher non seulement une *individualisation de* l'éducation, mais surtout une *individuation par* l'éducation, qui offrirait la possibilité à chaque individu de trouver sa motivation intrinsèque.

Nous reprenons l'idée d'*individuation* à Gilbert Simondon, qui à rebours de l'individualisme propre au libéralisme, considère qu'il n'y a pas d'individus au sens étymologique, c'est-à-dire indivisibles, à partir desquels les sociétés seraient construites. Il pose au contraire un principe d'inséparabilité, selon lequel l'individu est embarqué dans des processus permanents d'*individuation*, de construction personnelle par l'interaction avec un milieu et un collectif, impliquant notamment des dynamiques transindividuelles (Simondon, 1989).

Selon Crawford, « si nous suivons à rebours les traces de nos actions jusqu'à leur source, celles-ci peuvent nous instiller une certaine compréhension de la vie bonne » (Crawford, p.228). L'analyse des traces, que celles-ci soient des résultats empiriques de l'action individuelle sur l'environnement – la réparation d'un objet, la rédaction d'un texte – ou des bits accumulés dans des *data centers*, peut guider l'éducation dans cette mission. La diversité des traces ouvre la voie à une collaboration entre humains et algorithmes pour permettre à chaque individu de construire son identité en lien avec les autres et le monde qui l'entoure.

La quête des sources de l'engagement nécessite de distinguer les échos multiples qui conditionnent nos attentions au quotidien. Dans *Pour une écologie de l'attention,* Yves Citton a repris le principe d'individuation pour inciter à « ré-envisager nos comportements attentionnels à partir de leur conditionnement par leur milieu » (Citton, 2014, pp.344-345), pour une prise de conscience des éléments polyphoniques qui conditionnent ce à quoi nous prêtons attention : attention collective du grand public envers des « envoûtements médiatiques », attention conjointe au sein de groupes par le jeu de situations relationnelles, et attention individuante chez des individus construisant leur identité en priorisant certains objets de leur environnement. Dans une logique éthique

et écologique, l'éducation peut permettre à chacun de non pas s'émanciper vis-à-vis de conditionnements inévitables, mais de se constituer une attention réflexive, une « méta-attention » réellement individuante. Le but serait de lui permettre d'aménager en conséquence son environnement pour réaliser sa façon propre d'habiter le monde et d'y trouver sa dignité.

Pour permettre à chacun de chercher du sens dans ce jeu d'influences, on peut repartir de la définition de l'« expérience » donnée par William James, philosophe pragmatiste états-unien qui a autant inspiré Edward Thorndike, inspirateur du behaviorisme, que John Dewey, figure tutélaire des théorie constructivistes. Il définissait l'expérience comme « ce à quoi j'accepte de me rendre attentif »[135] (James, 1890, p.402), en opérant une sélection parmi la multitude de choses qui composent le monde extérieur. On a vu que le behaviorisme s'était chargé d'opérer cette sélection, d'abord au sein de programmes skinneriens, puis, sous sa forme la plus récente, par des classements de contenus pédagogiques opérés par des algorithmes.

Il s'agit, plutôt, pour proposer une éthique de l'engagement, de reprendre possession collectivement et individuellement de ce à quoi nous choisissons de nous « rendre attentifs ». Il convient pour cela de déconstruire la logique de *ranking* installée par les algorithmes de fouille de données, faisant miroiter une pseudo-objectivité en vertu des comportements passés de milliers ou de millions d'autres élèves. Comme l'a écrit Yves Citton en référence à l'algorithme *PageRank* de Google, « cette automatisation, basée sur des dynamiques d'auto-renforcement circulaire, devient [...] préoccupante lorsqu'on ne se contente plus de mobiliser les dispositifs machiniques pour nous aider à trouver ce que nous valorisons, mais lorsqu'on leur abandonne le travail même de valorisation » (Citton, 2014, p.142).

L'auteur réaffirme certains principes éthiques qui prennent ici tout leur sens. Partant du constat d'échec de la forme scolaire à intéresser nombre d'élèves dans l'apprentissage, il entend dépasser le débat entre gamification des savoirs et conservatisme pour penser l'éducation du point de vue d'une écologie de l'attention. La salle de classe, lieu canonique de la forme scolaire, recèle ainsi des possibilités d'attention non pas passive mais « conjointe » entre les élèves et le professeur, qui permettraient un enrichissement de chacun selon quelques conditions. Ce cadre éthique postule un travail synchrone où l'enseignant, loin d'une déclamation magistrale et standardisée de savoirs, ferait attention aux « retours attentionnels » de ses élèves et notamment aux évolutions de leurs émotions. Il modulerait ses stratégies pédagogiques lors d'un processus d'« invention collective » où il guiderait les élèves dans leur recherche de sens par l'investigation de sujets.

C'est dans ce projet éducatif, consacrant la centralité de l'apprentissage simultané comme lieu de construction d'attention conjointe, que l'on peut adjoindre des techniques numériques, comme autant d'occasions pour chacun de chercher du sens dans les savoirs. Citton vante ainsi, à travers les MOOC, les promesses de démocratisation de pratiques éducatives collaboratives, tout en mettant en garde contre la tentation de substituer ces cours à distance au fonctionnement présentiel des universités, pour des raisons d'économies d'échelle :

> « S'ils sont bien réalisés, ils peuvent contribuer à donner à certains savoirs et à certains gestes de recherche une diffusion absolument inédite et proprement réjouissante. Ils ne sauraient toutefois pas davantage se substituer à l'enseignement présentiel qu'une photographie ou une vidéo de pizza […] ne peuvent se substituer à la pizza elle-même aux yeux d'un affamé » (Citton, 2014, p.175).

[135] *« My experience is what I agree to attend to »*.

Les algorithmes d'apprentissage adaptatif apparaissent quant à eux incompatibles, on l'a vu, avec l'enseignement mutuel et la rareté des contenus qui caractérisent les MOOC. Ils sont davantage liés au mode tutoral individualisé et aux contenus préconçus des ITS. En cela leur contribution aux modalités d'attention conjointe est réduite à une portion congrue. Quels que soient les résultats impressionnants que ces techniques peuvent montrer en termes de prédictions de succès ou de décrochage, on a en outre observé les effets simplificateurs de la mise en données de la cognition humaine et les standardisations qui affectent la capacité de ces médias à permettre un apprentissage véritable.

Cependant, leurs retombées dans les domaines liés à la révision ou à l'assistance peuvent être intégrées en appoint, et dégager du temps pour des activités simultanées d'« invention collective ». Les médias adaptatifs ouvrent ainsi des perspectives intéressantes pour assister les élèves selon leur niveau de maîtrise, sur des bases plus inclusives. Dans l'idée d'une initiation, les ITS peuvent concrétiser des dispositifs de découverte personnalisés selon le degré de maîtrise ou les centres d'intérêt de l'élève. Par leur logique axée sur le *ranking*, les ITS peuvent aussi permettre à l'élève de travailler sa mémorisation pour arriver plus confiant en cours. Celui-ci peut également approfondir sa compréhension d'un domaine en y trouvant des significations qui ne lui étaient pas apparues auparavant.

Ces médias peuvent aussi assister les enseignants dans leur compréhension d'une part non-émotionnelle mais quantifiable des apprentissages. Dans un schéma empreint par la quête d'attention conjointe, les *analytics* délivrées auprès des enseignants gagneraient à se décentrer de l'assiduité de l'élève, qui indique peu sur les manières d'apprendre et véhicule plutôt un esprit de surveillance. Elles gagneraient au contraire à être pleinement orientées vers l'identification de la manière d'apprendre de l'élève, pour instrumenter des différenciations lors d'enseignements en présentiel. L'enseignant pourrait y trouver un moyen supplémentaire de « faire attention » à l'attention des élèves, dans une logique que les sciences humaines nomment le *care* (Molinier et al., 2009). En cela, les algorithmes de *clustering* semblent être des voies intéressantes, ouvrant la voie à des travaux de groupes, tant qu'ils servent la découverte et la socialisation de l'élève.

La condition principale pour une intégration des algorithmes d'apprentissage adaptatif dans une éducation « écologique » est à nos yeux que leurs usages soient explicitement limités, notamment par les pouvoirs publics, au tutorat et non à l'enseignement. Il importe ainsi qu'au-delà de l'*individualisation* algorithmique et de la modélisation de l'individu à partir de ses seules traces individuelles, soient enrichis les espaces d'*individuation*, de construction par l'individu de son identité personnelle à partir de ses relations à son environnement et aux autres, en ligne et hors ligne.

Tient ici notre contribution, encore liminaire, à la réflexion sur un cadre éthique d'intégration des algorithmes d'apprentissage adaptatif à l'éducation. Sa pertinence est, nous en sommes conscients, conditionnée par la présence d'opportunités socio-économiques pour les acteurs industriels. Nous faisons l'hypothèse que ce cadre, dégonflant la bulle d'un *adaptive learning* prenant en charge des prestations entières d'apprentissage, ne peut que bénéficier à une industrialisation réussie, car choisie, de médias qui ne seront pertinents qu'en complément de pratiques humaines autonomes et réfléchies. Nous ouvrons par là des pistes de recherche qui ne demandent qu'à être explorées à l'avenir.

ENTRETIENS

Entretien 1. Carolyn Rosé, le 28 septembre 2023

– Do you consider there is a gap between collaborative technologies you develop and those which are present on the market?

Carolyn Rosé: I think in the US there's a huge gap and I think that it's because all the marketing just focuses on whatever people will buy and not what actually teaches people. And so I think the focus of the research is on how people learn and how we can better support the learning but everything that we do it doesn't look flashy and so it doesn't help sell textbooks, it doesn't necessarily make things look more attractive to people who don't really understand the science under it anyway. And I think that there isn't an appreciation for good science in the market you know, maybe that's not surprising but that's where the disconnect comes I think. And so and now I think there's a huge number of new startups because of, you know, people think that there's like this huge advance in generative AI but the truth is that it's just more visible now. You know like all the things that people think just like suddenly appeared are things that people have been working on for decades and they're like thousands of people who've been working on those things, each one working on a tiny little thing. And now it's like… it's all like being packaged up as this generative AI that just appeared, you know. But it's not really what's happening but enough people believe that now is the time when if lots of companies start and they say that that's what they have, that people will buy it because they think it's something when it's really not. And so I think there's a total disconnect between the science and what sells yeah so yeah that's that and so in my career I think that I've mostly ignored that, because I'd rather do good science and you know um but I haven't I think maybe pushed enough on the side of making sure like fighting against this tendency that things that get widely deployed are not necessarily the best things. And I don't know if that was the right thing to do or not but it made me feel better about my career (laugh). Hopefully someday someone will push whatever you know... I in terms of adaptability I guess um there's a lot in what I do that had over time focused on the idea of kind of micro adaptations during the learning process but I actually think it's the macro level adaptations that are the most important. You have to make sure that for whatever is your learner population that what you're teaching them is well calibrated and I think that the pedagogy is way more important than the technology… but that's you know what I've seen over and over. It's what I really believe.

– You call that macro perspective?

CR: Yeah that's how I would, that's how I would describe it. I think if you are working in a context where the learner population is very diverse like we have community colleges in the United States that um… they're not universities and people come to them for lots of different reasons. Some people come to them to do their general requirements for college before they go to a university, other people come to get further credentials after they've already started working, some people come there just for taking a course or two, or like they couldn't afford to go to a university at all they're just going to get a TR degree there… just like people are there for so many different reasons that within the same classroom there's a lot more variation than there would be in a typical classroom anywhere else so like K-12 is organized in such a way that people are in groups that are more or less homogeneous. Here it's like that. I think in most colleges there are enough courses and paths through that they can organize it that way so… And it's much easier to teach when you have a more like homogeneous

student population I think some of our courses probably have more variability in the student population than others and in those courses I do think it's much more challenging and I think those are the environments where micro-level adaptation actually matters, or like figuring out a strategy where you can divide the time and target differently at different times and so you're sort of doing an integrated approach. I think that that can be really important and I've seen that as being really important in one of the big classes that I regularly teach here, because that course is not a core course in one curriculum it's like a course that people from all over the campus take as a… like a side course so that's why there's more variability so I think in those environments the micro level adaptation actually matters. I think that in other places there's enough homogeneity that probably just making sure that you've picked the right curriculum at all is the important thing (laugh). And um I think that there are ways in which if you design the pedagogy well, there are lots of different adaptive technologies that will all be good at serving that up to people and it's not like "magic technology is going to make everything better". So um yeah and I feel like it's like the cold sad truth in a way, because when I was a computer science major when I was an undergrad and then I made a decision that I wanted to go into educational technology because I wanted to have an impact on something that matters but then (laugh), I think that what I've learned is that the technology part of it is the least important part of the whole like big picture and then if you bring in the socioeconomic aspect of all of it, you realize that the people who are really in trouble it's not… there's no easy fix for the problems that are out there and… so if you think in terms of… like "how much impact by the end of my career am I going to really have on the people who really need it, from the work that I'm doing here ?" (laugh) it's kind of depressing actually (laugh).

– Did you take part in technology transfer initiatives?

CR: A little bit. So my work is in language technologies and so a couple of my PhD students actually started a company that was about automated essay scoring and I think that they've done a lot of work, they started their own company um that was called first Lightside labs and then they became part of turnitin.com. They left the PhD program and started this company. It was like right around the time when massive open online courses were like the hot thing and and so similar to what's happening there were lots of little startups that were starting and they figured they could make a lot of money, so they went off and they started this company. That was building on top of work that they were doing in their research here. So I consider that part of tech transfer that I was part (laugh) even though they kind of like took their stuff and left and I wasn't really part of it but I like to say that you know that was the tech transfer thing that I was part of you know and then they did actually make a lot of money and I didn't make any money but I didn't care, because I got to do my research and that's what I wanted to do (laugh) and they wanted to make money, that's what they did so… But then they were bought by turnitin.com. I actually avoid, I avoid being part of tech transfer because I feel like it just limits what you can do if you have to like hide things and make sure you do all the proper disclosures and to spend time with lawyers and I just want to do my work. So I feel like in a way other people who like doing that you know they are going to take the best that's out there and so if I publish my work you know it's contributing towards the big landscape that they're drawing from. So even if there's no direct way in which I can say that product was something that I worked on, I feel like I'm contributing to a field. But I do actually think that things are not transferring as well as they should. I think that if you look at what happened during the pandemic, you can see that you know it's especially sad from the the side of Computer-Supported Collaborative Learning, I think there's a lot that we know about how to sort of build community online, how to keep people engaged, how to use interaction between students to support learning, and it wasn't visible enough. And so then people went off and did crazy

things that were damaging. And I think um in a way like I should feel embarrassed about that because "what was wrong with us ?" you know... I mean there is an extent to which I think during that time that my community of researchers felt that it was actually kind of like good for us, that we did give us more visibility… I just don't think that it had the impact that it should have had. And so you know and now like we see students coming in to the university who were in high school during that time and it was damaging to them and we're still dealing with the aftermath in a big way.

– I would like to question you about the main fields that you target as research programs concerning the human computer interaction. So when you try to modelize the conversational aspects of pedagogy what is it about? Mainly the tone of the personalized tone of the conversation?...

CR: … I think that what we want to see is different perspectives being expressed, comparing across them, finding ways of like seeing how they relate to each other, so that's the core of what we really want to see. But I think that there are other layers that support that or not and so if you think in terms of tone it says something about the culture of the group and about the level of openness in a group and those things support the processes that we care about but I don't think they do directly impact how people learn. They might directly impact though how engaged yeah people are so I think there are indirect effects in there. So we look for all different kinds of things because all of them relate to what we really care about, and sometimes you have to address the environment in order to promote the processes that you really care about you know, and you can monitor those processes directly but knowing that they are not happening the way they should doesn't always tell you when you should intervene or how. So we have to think more holistically about it**.** So I think that we've looked at it over the years like I have like over 100 papers on this so it's… and each one is a tiny little piece of the puzzle um so we've looked at it from all different you know angles.

– How do you manage to judge the efficacy?

CR: I mean the bottom line is : how much do people learn between pre and post… I think that there are all different ways that you can look at collaboration and what would be a successful collaboration. So this is only one perspective on that. Some people think a good collaboration is one where people learn better collaboration skills that they can apply going forward. Or a good collaboration is one in which people feel good about the interaction that they had you know. Or a good collaboration is one in which they got a lot of work done during that session. And those things are all different from "they learned a lot during that session". And they're all valid ways of thinking about what a good collaboration is, and I don't think the same collaboration can be good on all of these fronts at the same time. But I would say that most of my work focuses on creating an environment in which people will learn more conceptual knowledge between pre-test and post-test. And part of that is because learning sciences at Carnegie Mellon is very cognitive (laugh) and we all like, we all care about those tests and so that's been my big focus but I do recognize that there are these different ways of looking at it. And every you know sometimes I've worked on projects where it was a different one of those like I'd say the most frequent second one is that when they come together they produce something good. So in engineering classes where they're working on a project together you want to support the project and a test is less important but it's important that they came together and they did something and so… I think that we support the those things in similar ways but there are other things that we have to think about, like sometimes a project is a longer term thing so then group composition is way more important for the long-term and how they're going to manage their time. Um then it is if they're like working together for 30 minutes and then you just want to know "did they talk about these things?

did they manage to learn concepts a b and c?". So we know one strand of work that I have done is on how do you actually form teams? Especially in a really big class where the teacher doesn't know all of the students or let's say even more challenging: the students have some choice who they're going to work with, how do you sort of create these affordances that draw people into groups that they're going to be productive in? Because a lot of times people want to work with their friends and on a long-term project I think we know that that's a bad thing but that's what they want to do, they want to be with their friends it's understandable uh but then there are all these breakdowns. So if what you want is for the teamwork to work well and the project to get done you don't want that, you want to try to encourage something else. So we've worked on some ways of orchestrating classrooms over time so that people… I mean I can't say we've totally solved the problem but that was the goal.

– So I understand you're mostly engaged in encouraging collaboration, more than forcing students… So how do you leverage these incentives?

CR: Sometimes… to be honest the most powerful thing, the most powerful incentive unfortunately is what's going to give them a good grade. So actually you can have a huge impact just in how you express the instructions, like what is it that you like… defining what is the scope of what it is they're supposed to be doing and how they're going to be evaluated. When people know how they're going to be evaluated that has a huge effect. But we still work on helping people actually pay attention to how they're using their time. I think recently a lot of the work that we've done has been in… um well we've been doing a lot of work in software development and how do people learn through that. And it is really important how they spend their time during the session and less important how much code they write. And so we've looked at how they should spend their time and then making them aware of how they're managing that.

– Let's talk about large language models. I see that there are many possible applications concerning the conversational adaptivity (adapting the answers, the feedback, the scaffolding or the even the context of the exercise). Do you have an overall view of the possibilities or do you target special applications?

CR: Oh okay yeah I think… I strongly advocate for targeting special applications because I think that again there's this all this rhetoric about generalized intelligence that's not true. It doesn't exist (laugh) it's a myth. Um so I mean I do think that there's a lot in general that you can do with something like ChatGPT until you really depend upon it to do it well, and then you see where it breaks down. But it's like you can literally go to it with anything and it will generate something, but you could also literally go to Google search and type something into your search box and use a more or less traditional search technique and still get a lot of useful information and ChatGPT was trained on the web essentially right? But I think that people forget that there are lots of like simpler ways that you could have gotten the same information need met. Actually I'll just say one problem with ChatGPT is that it isn't even just the large language model but they built some layers of safeguards around it that affects what you actually see as the response. And um so I was looking at something… um I don't know, do you know what permaculture is ? [*Yes.*] The interesting thing about permaculture is that some people think it's like this hippy fringy… whatever, and other people think it's very scientific. And if you look on the web you can see documents from either side but if you ask ChatGPT is permaculture like a hippy science? It'll be like oh no no there's lots of evidence that scientific and it's not a good idea to have prejudices against particular groups and there's no… So it gives you like all this like… Because I think that they're hedging you know, like they don't want to turn off any

particular group so they make everything very… But you know so it's coming from all these web documents but there's something that's actually skewing how it's actually presenting the information. And so one of the new things I'm working on is with one of my colleagues who does conversational search which used to be just using very traditional search methods but now they're looking at integrating those traditional search methods with large language models and so there's this question about how do people verify the sources that we're talking about, how do you make people aware of this. I think that it's dangerous that there's this layer that sort of prevents you from seeing what are the sources and is that really all that's out there, because it really insulates the end user from the things that we used to teach as information literacy. So we're looking at what do you do about that and I think actually what's going to happen is that ChatGPT is not going to continue to be the standard forever. People are eventually going to realize that that's a bad idea (laugh) but right now it just looks like this thing that everybody loves you know because it it sort of like makes you feel good about yourself (laugh) and gives you something. Anyway I definitely don't believe in that I think that I think eventually people will become disappointed with that. But there is I think a lot that can be done with large language models, and I think that it's not all together different from things that we were already using in our field like BERT. I don't know, are you familiar with BERT ? [*Yes*]. Okay so I think it's really just a really much bigger BERT in a way, and I think that there's a lot of experimentation that's going on now about trying to leverage what it does better than what we had been doing but it still doesn't solve all of our problems and so there are a lot of different areas where people are looking at the specific weaknesses. And then there are ways in which in particular application areas where you really care about the behavior that you want to see, that you realize that in order to get there you have to sort of have a very focused effort towards appropriate fine tuning. I think one thing that is interesting from a technical standpoint is the idea that um rather than have one very big fully focused dataset that's all that one problem, now there's much more work on how to take heterogeneous data sources and cobble something together that solves your problem. And for a really long time I think in the you know mid-like around 2010, 2011, I one thing that worried me about where machine learning in the field were going is that it seemed like people just ignored all the problems where there wasn't already a huge data set. And then you forget the fact that there's all these problems in the world you can't address and everyone was really focused on just these benchmark tasks that we had these big things. I think now people are thinking more broadly, and I think that's very productive and I think this area of learning from fine tuning over heterogeneous data sets towards a solution that's focused is a really productive direction for the field. And I also think that since there is a lot of people believing – even if it's not true – that like all of a sudden we have this new thing you know that everyone wants to put a lot of money into this and I think the fact that there are lots of companies will mean there will be innovation, because some of them will do something and others will just lose a lot of money but at least (laugh) there will be a big investment in trying to get somewhere so… Anyway that's how I see it. So I would say better investment is like specific application areas and I would also say that lots of people are thinking about applications but not everyone is asking themselves what can we do now that we couldn't already do. And I think it would be a much better investment to focus on the things we couldn't already do a cheaper simpler way and so I would encourage people to think in that term I think you know almost all my colleagues I think we all agree on this and we're all like pushing our students to be like "but yeah that's seems like you could do that but we could have done that like 15 years ago too**"** (laugh) so people forget...

– You mainly imply yourself with your students on these problems but I understand you do not want to be too much involved into tech transfers…

CR: Yeah I think I just don't want to care about like starting a company, I don't want to spend my time with lawyers, I don't want to try to restrict access to what I'm doing, I don't want to think about just what will make more money because I think that… um I think it's a distraction. It's not that I don't care about the problems in the world, it's not that I don't want to develop something that you know meets a new need or something like that, I care about those things, but I just don't care about spending my time with lawyers.

– Then still talking about these kind of technologies how do you define the right border between what the machine might do and what the instructor uh might keep as own…

CR: Like the role of the teacher and all of this ? I don't think the role of a teacher is going to change that much. I think that there's been a fear of AI taking over the classroom for a really long time and it hasn't happened yet and I don't think it's any closer to happening. Um I think that there's a lot about the experience that students have that is just where people play a role, not because they can provide a kind of stimulus from a cognitive perspective that you couldn't actually make a program do, but just because it's different interacting with a person than with the information. So I think that for example if you think in terms of collaborative learning, the value is in appreciating the fact that there's something in somebody else's mind that's different from what's in my mind, and there is another mind that I'm trying to understand. If I'm just talking to a computer agent there's information there but it isn't a mind that I'm trying to understand, it doesn't have a human experience that's going to help me look at the world through someone else's eyes who's experienced the world the way I did. And so I think that's what you can never do with the computer and um I think that people get something out of being with another person. And so I don't… and I think that if anything the whole Covid experience really highlights is that because it's like… I don't believe that the information that people were getting over the Zoom thing… well for kids maybe, I mean I was with college people so I can only talk about my own experience. I actually thought that in many ways the classes that I taught online went better as online classes, because in a physical classroom that's a big class people can hide in the back of the room, but in a Zoom everybody is equally close to you. And and we could have these hands-on sessions where if somebody was stuck on something they could share their screen and I could see exactly what they were doing. And I felt like in many ways it was better but still we see the effects you know and I think a lot of it was the isolation, the fact that it's hard to keep people engaged when they're by themselves. It's easy to then shut it out all together. And so that's why I just have no inkling that the role of the teacher is fundamentally going to change because of generative AI (laugh), I think that's all part of this myth of what it is.

– So if I understood well your methodology you blend a lot of quantitative and qualitative methods?

CR: Yeah. Even in the modelling work that I do in language technologies it's very much that way, because I think that language is something that is more often than not actually bending rules you know. Models discretize the world and so there's something fundamentally wrong about the nature of what a model does when it comes to something like language and how we actually use it. We fool ourselves into thinking that we're capturing language. I think that there's a lot that we can do on a practical level with language and still build something that does something semi-useful but I think that we're missing a lot of what's really there and that's why I blend those together in my work.

– Maybe last question about more lexical question: are you okay with calling your research as artificial intelligence or adaptive learning researches or do you prefer other terms?

CR: Sure, no no I'm fine with calling it that but I guess it's not how I usually label myself, I would I think the term that I identify with most is computer-supported collaborative learning, that's my community.

Entretien 2. Peter Brusilovsky, le 27 septembre 2023

– Would you personally qualify your works as artificial intelligence? with which meaning?

Peter Brusilovsky: Yes. That is a very standard expression and the field which I'm working on at least a lot of my expertise and students called "Artificial Intelligence in Education". It's a very big field but I've been there from pretty much the early days. I was the steering committee member of the first AI in education society which is really active. It organizes AI in education conference every year and I think it's the main place where people working on that come together. But just to make sure: AI and personalization are not the same thing. First of all we can use very simple personalization which might not need AI which I believe some of the companies you mentioned are doing but on the other hand there is a lot of use of AI in education which is not personalization. Like nature language dialogue: they might not be personalized, the AI is there to understand nature language, to produce nature language. There are many examples, but there is a very good overlap when you use real AI techniques, user modelling, personalization, student modelling to achieve goals and education, so that's not the same thing but overlapping.

– I would also like to speak about the technology transfers especially in the US and in Pittsburgh. What do you consider as the main drivers or constraints for a technology transfer between laboratory and educational markets?

PB: Well there is a lot of steps for a lab technology to become a part of the market, and the United States is really probably the best country to do that. There is a lot of support in the universities to help with that because every time the faculty becomes the owner of technology, they can license technology; you don't even need to start the company; you can copyright technology, patent, and license that anytime and university gets part of the money. So at University of Pittsburgh and Carnegie Mellon, there are very large programs to help faculty with creating startups or doing the licensing, and actually I went through the program like that here, you may see how it works. It's funded by university in itself, and then every time it brings in faculty, gives them starting grants and offers them courses on how to do enterprises, how to do startups, I think it's very real, so it's really an important driver for us. However in all the time at the University of Pittsburgh, only one company, which I know, kind of get through that in the field of education, the company founded by professor Chris Schunn, which focuses on kind of AI technology, he is a professor in AI in education and known in the field, but his company was based on peer review technologies, so he has I think produced one of the most popular peer review based approaches, which is kind of... based on human intelligence, but a lot of ideas are supported by AI. It’s called Peerceptiv. So I think it's in the lot of startups which have been pushed out by this by this kind of a program. In medicine specifically, huge number of startups. Computer science too, machine learning, but very little education. That means either education market is harder, or just a fraction of people who work on education is lower. And we have quite a number of researchers, Pittsburgh is one of the central places for research and education, definitely one of the top three in the world, and yet we have like 10 faculty who work on the topic,

but yet we only see one of this attempt to market the technology. So it might be hard. At Carnegie Mellon again even more people, but again you don't see that many startups, there are more, if you talk with people at Carnegie Mellon, they are doing more startups. Carnegie Learning is just the best known, but it's still not just like, well...

– And do you see an evolution in the way these projects go to markets? Are there especially concerning adaptive learning softwares are there new trends, new drivers or constraints for these types of technologies?

PB: I don't see a lot of new trends to be honest. I don't know why but generally a technology might have existed for 10 years and try it in libraries and explore it before it gets to the market and most of the time it may be a press from a person who might be a student, might be faculty or might be some external person to get it. Otherwise it might not happen because we're really busy as faculty and researchers and not many of us are motivated especially in Pittsburgh because price of living is very good. If you live in Silicon Valley, even if you're professor in Stanford, it's a very strong push to create a company**,** because otherwise you just have a one or two-bedroom apartment, that's very expensive place to live. In Pittsburgh, if you're a professor, you can get a big house, you don't really need to do that. Also infrastructure is completely different in smaller city than a bigger city, I don't know, maybe this or not. But there's not that many companies in e-learning, so that means that may be a tougher thing to do… and what usually the company do discover is that they might come to the market with one idea, but they need to adapt and change it, because the idea itself is not working, you need to figure out how to expand, how to make it work with a huge number of people right, which might be kind of infrastructure issues, and how to keep the engagement. And engagement means both engagement of the providers, I mean those who push use of the software, instructors in the school, parents and the home learning and the learners. And you kind of need to engage both, because even if students are happy about the software, but those choose to buy the software not, like instructors, they are not going to use it, the same as is the students are excited, but parents are not, it's not going to happen either, right? So that's a kind of tricky issue, so you're kind of right, engagement is the most important thing, and usually you don't learn it until you get to the real market. So that's the discovery. But again, while we get a lot of educational research here, social personalization is not the place where the majority of education software have been created. So I think it's not a good research, but mostly market pressure, opportunities. Venture capital is important, there are quite a number of these companies being in Silicon Valley, they might call themselves personalized, but there might have almost no personalization. It’s just like a very nice word, and yet they exist as companies, so that's really interesting thing.

– I've worked on a methodology paper which was launched by Pearson and Edsurge, maybe you know it, it's about adaptive learning, but on an industrial scope, considering different publishers who have different kinds of adaptivity. There is a trilogy of adaptivity: adaptive content, after a student makes a mistake, for example scaffolding, the selection of certain hints, visual cues; then adaptive assessment, when adapting the difficulty of the next question after another; and then adaptive sequencing...

PB: These are all traditional technologies, especially I like adaptive sequencing, which is my main field of expertise. Again, the trick is that in most of the cases, it's not very much personalized and we need to distinguish these three cases because the first one uh if the student don't really well answer the question we need to do something differently…It's one of the oldest technologies more than 50

years old. In the old good days it was called computer assisted instructions, than called branching instruction that has been pioneered by a person called Crowder. The challenge with that is it has very low adaptivity because it's adapted to the case, not adapted to learner' knowledge, and it doesn't really model learner's knowledge. As a result, it's very frequently inefficient, and that's exactly what is driving the negative impression on personalized learning. Also the amount of things you can do based on the results of single problem or single quiz is very limited. You really need to get a lot of content, so if for example you have an explanation and a student fails the problem, what you can do is basically repeat this explanation. You have two versions of an explanation, you can try another version. But it usually doesn't help because it doesn't address the roots why a student isn't doing that. So it's really inefficient technology. It is still better than nothing in many cases, because the most cases are not adaptive. In fact just the fact that the system can assess what you're doing and provide feedback and do something differently is very important engagement because the person feels like it's a dialogue. And when, whatever you do, the system is doing the same thing, it is becoming very low engagement and it's very well-known. So the system needs to be perceived as being responsive, that's it, but responsiveness is not exactly personalization. On the other hand computer adaptive testing is a very strong discipline connected to psychometrics. It has a very strong mathematical background in Rasch models and computer adaptive testing. So it can really produce very impressive results. But the whole research and the whole foundation of what has been developed for adaptive testing is when study knowledge do not change. So you get a battery of tests to estimate student learning and now people try to adapt it to very different applications where for example you learn and then you try to do testing adaptively and the classic technology is not going to work because they assume that the knowledge is not changing and trying to apply classic Rasch models when the knowledge is changing. It is like trying to make a surgery, and the patient runs or walks. Things change, and you just can't follow that. So that's the other story, but at least this one has a very solid background. So sequencing is kind of in the middle, because there's generally very good basis for that, it's an ability to model student knowledge as they learn, and this is done by student modelling approaches like Bayesian Knowledge Tracing for example, and there are a number of them, and they're very well explored, they are formalized and they can be assessed, they can be assessed offline using predictive modelling. But there are still a lot of factors in adaptive sequencing uh and just I can name you two. First of all when you do adaptive sequencing … it required that you have a lot of diverse content because depending on the level of student knowledge and different knowledge gaps you need to show students different things and maybe different types of things sometimes you need to present readings or examples, sometimes you need to present easy problem or challenging problem. So, the system needs to get about 10 times more content than you need to use for the normal case. As a result, well you try to pretend that you can do adaptive sequencing in courses and MOOCs, it never happens because MOOCs barely have enough content for a single student to learn, and creating 10 times more, which is required for good sequencing, it's just out of question. So there's only pockets of good adaptive sequencing for specific needs like programming, if there's a large set of problems created, then you potentially can do good sequencing, but overall in a big course like MOOCs… adaptive MOOC is a myth, they don't exist, because it just get everyone walk the same line. That's the first issue, right? we can do good sequencing because we can do good student modelling, but there's just not enough content. So it's like you understand what kind of tool you need for the job, but you only have a hammer, it really doesn't matter, I mean that you understand that you need a chisel, you only have hammer right, or screwdriver right? So the other problem is that knowledge modeling is kind of typically not perfect, we observe something that the student is doing, but not everything, and as a result these technologies are not really perfect. Assuming that they know everything about the student, they would do a good job, but usually they don't know, since students can do something elsewhere.

And as a result easier you need to make sure that almost everything which student is doing is inside the system, which might happen if for example students only work in class and every time in class they use the software. they usually don't work outside of the class in midle and high school. Another option is to provide flexibility for the students, so if the system is not correct in infering student knowledge because they did something outside of the system, they can correct. And it could be done by so-called open learning models, which requires students to see what system thinks about the knowledge and be able to correct it or act accordingly. But the majority of adaptive systems are not doing good work in open student modeling. By the way, what I demonstrated with this kind of with several different classes, it's a kind of very crude way of open learning modelling, we see what kind of students can do, but it's very poor, because it's not going to help students in deciding whether something is good idea or not, so if you like to move on with imperfect student modelling, you need to provide a way for human-AI collaboration. In the classic personalized learning AI is driving everything, AI chooses what is the best question for you in adaptive testing, AI decides what is the best content for you in adaptive sequencing, right? But the modern AI is more and more human-AI collaboration, because people understand AI is not perfect for many reasons. So we need to find a way for humans to contribute right? In different branches of AI, that is becoming very fruitful, but it's still very little in AI in education which focus on human-AI collaboration. A reason for that is that the biggest funding in education comes for middle school and sometimes high school – at least in the United States, nobody cares very much about the colleges, they are kind of private businesses to some extent or state-level businesses – and in middle school the ability of students to be in control of their learning is much lower, so here indeed AI might be the driving force, even if it's not perfect, but even starting with high school where modern students are very different, they get a lot of stake in controlling their applications on mobile phones, they know how to do it and treating them as completely dumb figures where AI is doing everything for them is not appropriate. However very little research is done on how to engage uh humans in the loop. So that's just two problems right, in sequencing, but at least sequencing has a good background and I think it will get through that because even now there are some really good pockets of sequencing like for example this home prep software in China. Most of the prep is done within the system so the system can assess student learning and as a result modelling is reasonably good. And it's a huge amount of students so the system can improve the algorithms. But again these are pockets and they just use a very small portions of software in a very small domains, but overall let's say you like to do personalized sequences high school algebra course, it's a lot of content right, and uh, just not enough students within the same system to do that. Some companies can be better opportunity to do it like Carnegie Learning, which I think one of the reason to success is that they really use this cognitive modelling and student modelling on the back, so they can do very good adaptive sequence, although they have not always done that. So that's basically the answer. So there are three promising opportunities, you're absolutely right, there are three things, but one of them is mostly hoax, I mean it's it personalized, but is not really solving units, and two others have very low small pockets of existence, so we kind of need to do more, but I think this is coming and this is coming in the way of natural language learning. So it has been pushed through different directions, so for example using Alexa right? People get used to talk with AI using natural language and also there is a converging research direction in AI in education, it's dialogue-based tutors and when you're doing the dialogue you kind of really need to do good modelling of a student. And as a result what you get here is a new way of personalization, it's not just learning content which is personalized but the whole dialogue, the way you construct the dialogue, the way you adapt. You can adapt also not to the knowledge and adapt to engagement and I think that is probably going to be the fastest developing field where you definitely get a bunch of enterprises and ChatGPT is a great example of that.

– That was question I was about to ask you : to your point of view large language models are a possible growing perspective for the whole adaptive learning research?

PB: That will be very strong push to move the research on adaptive learning in this kind of dialogue based-tutors which again… in Pittsburgh we have two faculty working on them so I know about that and they're very successful. However we need to get through the first stage when we can adapt this language models for education and we already get some failures and some successes. And in our case what we try to do… so we for example explore worked examples in programming where people can see the code of a problem and they get explanation why a specific line is important as a part of the code right? It's all for teaching computer science, and in the past, these explanations have been generated by experts. And also in our new work with University of Memphis we're using that in a reverse way, building a tutor which can interact with the students in a dialogue way. And we try to use language models instead of experts. But what we're finding out is that explanations produced by large language models are more complicated, more verbose, and farther away from students than an explanation produced even by experts. And I think that is the outcome of LLM being trained on a professional body of knowledge on the web, while to succeed on the educational domain, you kind of need to train it more in education literature which has been specifically produced to explain things clearly to people who don't have the knowledge. And I don't know exactly how LLMs have been trained, but I think this educational content is a very much minority. As a result, if you're an expert, LLM can provide very good explanation, if you 10-grader, maybe not, if you are middle school, definitely not, it just speaks in different way, it speaks to you as an expert, and so that's an issue. But LLMs are becoming prominent and there are open LLMs where you can use yourself, so technically you can use the technology and train it whatever you like, you get it and you train it in the body of open textbooks and a lot of publishers get a lot of textbooks available right, so if for example publishers and LLM kind of merge together, it can get versions of LLM, which indeed can talk to learners in kind of the same language, so that's the barrier, but well it's a small barrier, it just like take maybe two or three years until people will adapt on both sides because the technology become very… this technology is expanding so rapidly, it's taking everything by the storm, so far most of the instructors perceive it as a threat as a way to kind of cheat, but in fact is amazing opportunity, it's kind of to some extent like a calculator which a lot of math teachers perceive as a threat, but you can't stop that, I mean calculating is a part of our life, we can still teach people long division or other calculation, but it's not going to be what you're doing in your life as a professional, and in the past there were people who been just calculators right, it's was their job, we don't need it anymore, and as a result these people can do some something else. The same thing is happening with other jobs which can be done by LLMs. So I think it's an opportunity, how you can use the opportunity. And a lot of people trying it, so again, LLMs, at almost every conference attended in the last half a year, yeah, all the discussion has been focusing on LLMs on one way or another. Just an example of one of the recent AI conference, which was supposed to have three keynotes on three different topics, like knowledge graphs, I don't know, you are probably not an expert in AI, but knowledge graphs, NLP and some topic completely different from that. At the end all three keynotes were about large language models and ChatGPT specifically and the experts tried to compare what they did before in their fields and what in the fields is happening now and how potentially we can leverage LLMs together with what they do. So it's influenced all of us. That's why I believe it will be the next expansion in education too... *[silence]* by the way, what is really interesting is who succeeded on the market? In many cases, the technology which was originally created in labs and which has been focusing on supporting the class of learners might not be that successful because there is no easy client, so as a focal you can use

it on your class, you can use it in other classes, but if you try to sell that is very hard because who's the buyer? the student in the class is not a good buyer, he's in class, instructors don't have money, it took a lot of efforts to figuring out for example for Carnegie Learning, who could buy that. That's like school districts. On the other hand, if you focus technology on the home market, there are plenty of buyers. Even if the technology is very strange, it might succeed, and Chegg is an example, which has started in a completely different way by basically peddling solutions, but it that was appealing to individual people who are very willing to pay, so that's a market opportunity, but then they try to think what else? and they engage quite a number of people who do personalized learning, so one of my former collaborators from Carnegie Mellon and my former students both now work at Chegg because this is the place which is hiring and they have very nice ideas, so sometimes it does matter how you start, you are in the market and then you have to adapt to the market and get creative ideas.

– So you talk about this big step, which is the way to let the human enter into the pedagogy and the best way to find the collaboration between teachers and machines. What axis do you find the most important in this direction for a teacher? What collaboration do you find relevant today in the technologies that you study?

PB: It's a good question. I think the need for human-AI collaboration or basically it's more generally called 'human-centered AI', usually pushed by two groups of people: first of all researchers who understand it's important and every community has these researchers. Like in our community for example, professor Judy Kay from Australia, University of Sydney, was always pushing the things like scrutability in AI in education, so students need to see why the system is working like that, what the system knows about them, right? so there are always people like that. However the majority did not care very much, so what caused the majority to change their mind is the failures in just pure AI approach in practical cases and successes when you get collaboration. So in the case of machine learning and other domains, success cases and human-AI collaboration like for example 'human in the loop' machine learning, interactive machine learning, as long as people started to work in that and produce good results, you start seeing that your troubles might be resolved in the same way, but that takes time, but these movements kind of converge, those who believe in that from the very beginning and those who realize is important because what they do is not kind of quite working. But it's kind of more delayed in the field of education because to analyze that it's not quite working, you get it to scale, and other part of AI like machine learning is very easy to get up to scale, but some parts which require users to be part of that is not very easy to get up on scale, so not everyone is learning the lessons, it might take more time in education. In fact AI in education was one of the first places where human-centered AI started, and that was these open learning models, it was the first case in the whole AI where the AI system tried to show the user some knowledge about the user through the process, and engaged the user in potentially changing that. That was a completely groundbreaking. The rest of AI didn't get there until 10 years later. But kind of after that, AI in education hasn't been really contributing to this human-AI collaboration, but this is actually pushing now. So I've seen a number of people who started to work on that, and in our team it was always a topic, but it hasn't been really resonating, so we rarely presented to them. We presented, nobody cared. But now when we present on the topic, people get enthusiastic so we start doing more work about the same thing. So I think that that's a long way in the field of education, but since it's taking everything else quite by the storm just like large language models, which is actually an example. A large language models is a great example of human-AI collaboration, AI is trained, but humans need to work on specifying the task as a prompt, and humans need to know how to do it, and they are willing to do because results are great, so it's a

great example where, not just sitting passively, by contributing, you can get good results. So that might be another push to human-AI collaboration.

Entretien 3. Bruce McLaren, le 24 octobre 2023

– Do you want to introduce yourself to begin with?

Bruce McLaren: Yeah sure. I've been doing education technology research for about twenty years and I've had three main general areas that I've done research in. One is – I'll go from what I've done the least in to what I've done the most in, I would say – I've done collaborative learning research. I actually lived in Germany for several years and I was on some EU projects where I particularly focused on collaborative learning and technology to support collaborative learning. I've also done a lot of research with intelligent tutoring systems, including a project… actually two projects I have going on right now, but my sort of main area of interest and what I've done the most research in the recent years is digital learning games. My lab has developed several games, most prominently one called *Decimal Point*. That's for kids learning decimals and we've done a series of studies. In fact as we speak I just interrupted my writing a book chapter about the series of experiments I've done with the learning game *Decimal Point* that we've developed. And we've just recently developed some new learning games. So those are my major areas of interest. Prior to that I have a PhD essentially in artificial intelligence, the actual name is intelligent systems, but it's really AI. And I also worked for a number of years in industry, before becoming an academic. Primarily in like developing expert systems, AI expert systems, um, there was a company that was a spinoff of Carnegie Mellon called *Carnegie Group*, um, they're now defunct, but they developed expert systems for industrial applications, and I was doing that prior to becoming academic. So that's I guess my summary.

– All right, well there are a few elements we will talk about later, but first question is about, well, it's maybe a bit prosaic to you, but how do you define personally, artificial intelligence and especially in education?

BM: So in general, I define it – I know there's been sort of different ways it's been defined – but the way I view it is you know any kind of software activity that does something that more or less simulates what humans can do and maybe can even do it better and a lot of cases it can do it better and faster. So that's sort of my general description. I know back when I was first doing AI stuff there were two paths: one was like sort of doing cognitive modelling, so trying to really make you know the artificial intelligence work like a human's intelligence works, then the other path is just doesn't matter if it's like humans or not, as long as it's something you know that more or less simulates or emulates the results of intelligence. And a lot of machine learning is that I think where it's doubtful that the algorithms that have been developed in machine learning really look like human intelligence, but you know the learning that they do, is like humans, and even much faster and better in many ways. So anyway that's my answer to that question.

– Do you still notice these differences of understandings, of meanings between researchers in your discipline or between social players, politicians, industrial players?

BM: I don't actually… you know, but that might be also because I just you know my nose is down to doing my research and I'm not getting involved in those sorts of more general discussion. So it might

still be happening. I used to be involved, I used to go to like AAAI, the big conferences in AI, but I haven't gone to those for many years because I've been focused more on AI applied to education. So no, I can't say that I know like the… what the current state of argumentation is around those two perspectives on AI.

– All right. Let's talk about a chapter you recently wrote. I saw that it was about artificial intelligence in STEM education and that you presented three paradigms: AI-directed, AI-supported and AI-empowered STEM education. Could you clarify a little bit these this distinction and how did you come to pointing out these three paradigms?

BM: Oh yeah, I think you're citing the book chapter that I was involved with, I think last year, we wrote a book chapter… I'm trying to remember what our distinctions were, but I think maybe more relevant is my distinctions that I came up with in a chap, another chapter that I wrote, I'm not sure if you saw that, it was about learning games, and there we made the distinguishing characteristics of AI used to sort of personalize educational technology, in particular in games in that context. But I think it could be said in general, that's maybe the most common use of AI used in education and educational technology. So in other words, doing something like emulating mastery learning using Bayesian Knowledge Tracing so that you determine what skills students are most lacking in and then you have the students practice those skills more than other skills. Another way that AI is used in learning games and maybe a little less in educational technology in general is in providing like… in games there's always non-player characters, they're called NPCs – I'm not sure if you know much about digital games – and using AI to make those characters more intelligent. So for instance there's a game that a company out in California headed up by a colleague of mine named Lewis Johnson where they develop like basically natural language games that allow people to learn how to communicate. Particularly in different languages. He does a lot of military applications so for instance you know soldiers could learn how better to communicate with Iraqis in Iraq for instance. It's set up as a game environment. So there the intelligent characters are, you know, who you're communicating with as the human and the more intelligent you can make them the more interesting the gameplay is. Another use of AI is to do more like "after the fact" analysis. So there's a field called Educational Data Mining, which you're probably familiar with, which uses lots of AI techniques to be able to investigate what, you know, students have done in their use of educational technology. So for instance, trying to identify the affect that they have while they're using technology – so are they confused? are they frustrated? –, those kinds of things. So that also may… it's moving towards being more of an online activity, but it's been I'd say more typically been offline, meaning you use the technology and then you apply these techniques after the fact to be able to identify various behaviors and states of affect that students have. So I guess that my answer to that question is that those are kind of the different ways that I've seen AI being used in the context of educational applications. Now of course all of that… that's all going to open up quite a bit and there's probably going to be a lot more applications with you know the rise of LLMs, in fact they're probably going to dominate. And they might be used for those three purposes but they may also be used for other purposes too… like other kinds of language applications. Like for instance in our game we prompt students to self-explain their steps that they take in solving problems, and you know if they do that in an open-ended way where they sort of just type in their explanation, you could imagine using *ChatGPT* to analyze what they say and then provide some feedback using all of the internet knowledge that it's able to scoop up, you know, using the large language model approach. So I think there's going to be… so it's actually a pretty exciting time. Everyone is sort of rushing around, trying to figure out like what's the best way to use large language models. So that's probably going to be emerging as the most interesting area, both

research and practice I think. Yeah, although at the same time there's a lot of concern about the ethics and you know privacy of students and all of those issues I think are all going to play a big role, because of course large language model bring in all kinds of information, some of it which may be biased. There are going to be those issues to tackle and maybe get in the way of it really being used in a practical sense for a while until we have a better understanding of, you know, how they operate and whether or not they actually do exhibit bias or not.

– Okay thank you for bringing this subject. I saw you also wrote a paper about testing ChatGPT *for mathematics tools. And then I would like to talk about the mathematics ITS market for K-12 students in which you have expertise. So what gap do you notice between research state, the research you do at Carnegie Mellon, and the products that exist on the market?*

BM: Yeah, I guess the gap is sort of how much we push the envelope and how we're able to… and this is always the case with research is that practical applications are always going to be you know lagging a little behind, I think. Just because the scaling up process is pretty difficult and you know you have to sort of prove things out in many experiments rather than just one or two experiments. So you know one example Carnegie Learning is a company that I'm very familiar with, and I've collaborated with. And they have intelligent tutoring systems, cognitive tutors for mathematics that are used pretty much all across the United States. They have hundreds of thousands of students who have used and are continuing to use their product. But I would say that the state of their technology… I mean they made lots of cool software developments, but I wouldn't say that they've pushed the envelope on a AI and cognitive models much since the earlier days of cognitive tutors when the company was first founded. Again not to be disparaging because I think they've really done a lot with their software, but maybe they haven't you know always kept up with the latest and greatest… like some of that stuff I was talking about, using EDM to be able to analyze students states of affect and metacognition and those kinds of things. I don't believe that they've been able to incorporate that yet into their product but I think that's clearly a direction they'll go. I mean I know they're fortunate, they have a guy who's the chief scientist, I don't know, have you met Steve Ritter?

– No, I haven't met him, but I've met a few engineers at Carnegie Learning.

BM: Okay, well Steve is the chief scientist, and he originally came from Carnegie Mellon, and he was involved in early development of cognitive tutors. So he has like a research perspective, he's really good at sort of wearing both hats, both the research side and the practical side. And he acts as a conduit between CMU and Carnegie Learning. Which I think has been a huge benefit to their company, that he keeps his connections with people like Ken Koedinger and Vincent Aleven, who are both founders of Carnegie Learning, myself, John Stamper, there's a bunch of people at CMU who have collaborated with Steve over the years, and you know managed to connect. And I think there was a project that we had that unfortunately got torpedoed by the pandemic, and that was… we tried to do some experiments scaling up the approach that Ken Holstein took – I was one of his co-advisors – and he basically came up with this concept of "smart glasses" in the classroom, where the teacher is able to use augmented reality to be able to assess what students are doing, so if a student is struggling in the class and they're working on a computer because you know the data is being logged and we're doing data analytics on what they're doing, like there would be a virtual hand that would go up that would allow the teacher to know that the student's struggling. And he did like initial relatively small experiments to show that that approach really improved student learning, because the teacher is able to more directly know which students are struggling. Otherwise you can imagine even

with a regular dashboard on like an iPad or something, it was never as effective as this idea, I mean that combining dashboards with the augmented reality in the smart glasses was really the beauty of his idea. But anyhow those are the kinds of things that aren't quite yet – back to your original question – aren't quite yet out in, you know, regular use, but... I know, actually, again, the pandemic stopped us from being able to do classroom studies with that, but I know Vincent Aleven has continued to work with Steve Ritter on a version of that project that I just described, and I know they just submitted a proposal for that. So those are the kinds of things, those really state of the art technology things are mostly not out in regular use, but maybe aren't too far away from being regularly used, you know, in the marketplace.

– All right. Regarding the uses, do you mean that the teachers or the students are not ready yet for this type of technologies?

BM: You know, I don't think it's the students and teachers as much as it is just cleaning up the technology and making it, you know, being able to scale it up and used by in the bigger scope. I was actually really surprised and I have to say my own beliefs were sort of overturned by Ken Holstein's idea that, you know ,teachers would willingly wear glasses in a classroom and the kids would accept it. I thought "oh the kids are just going to make fun of that, that's too unusual" but it turned out it was not that way. There was a little bit of, you know, they probably chuckled and had some laughs at the teacher at the beginning – especially because the time he did his studies, I haven't kept up with that technology, but the glasses were more bulky and, you know, you looked kind of funny if you were wearing them, they weren't just like regular glasses. But I think it's moved more and more towards like regular glasses that you wear on your… you know, like you normally wear. So then the kids wouldn't even have any sort of adjustment time. Yeah, I mean, there's always going to be acceptance issues, so there's early adopters just like there is in all technologies that get transferred from research to the real world, you know, there's early adopters, those are people are very technology, science-oriented or willing to try stuff out, and then there's people who have been teachers for 30 years and you know they don't really want to rock the boat. You're always going to have that, so that's always going to be something that's going to slow things down. So I actually don't think though that that will be too much of a hurdle, I think the biggest hurdle is just going to be fine-tuning the technology and you know… because the other thing is that it's not going to work in just a regular classroom, it's got to be a classroom that's already technology-enhanced, you know, like students are already using the technology. And more and more, at least in my experience, that's not unusual, you know the kids will… like schools I work with, kids all have issued iPads to them or Chromebooks and you know as long as they have technology and they're given assignments where they're all working in a classroom then you can imagine that technology working. You know, you basically just need to make sure you get the whole software cycle working so that you log the data then you analyze the data and then you provide the data to the dashboard, to the glasses.

– Now I would like to talk about game-based learning and I heard you talking about the tension between engagement and learning in many softwares. So how would you define the good balance between these two objectives? In which conditions can they be simultaneously reached? and talking about data, how do you check if they are?

BM: Yeah, so actually one of the studies that I did with our game was exactly that. And by the way, that was the reason I got interested in game-based learning to begin with. It is that I was working in classrooms at that time with tutoring systems, in particular in middle school classrooms, and they

were using tutoring systems that I developed and even more specifically there were tutors that had worked examples and erroneous examples in them. What I noticed is that after our studies were done and the kids were allowed to sort of do what they wanted, that we'd always tell the teachers – you know, because there'd always be a delayed test, we didn't want the kids to be studying what the topic of that particular study, and that would be, you know, in our case it was decimals –, so the teacher would often just say "okay, you go ahead and play this game". And I just noticed kids were playing all these different games and it got me wondering "okay, so they seem like they really enjoy these games, but are they really learning anything from?" (laugh) because it's one thing and I think it's easy and especially for teachers you know, when they want to set kids off on doing technology if kids seem to be enjoying themselves there's a temptation to believe "oh yeah this is a good thing for them". But I wanted to know is there any evidence that they're learning from those games. And then what I did is I started to investigate that issue, I looked into the literature, in particular I looked at a paper which eventually became a book by Richard Mayer. He's an educational psychologist – do you know Richard Mayer is? He's somebody you might want to check out because he's actually pretty prominent. He's done a lot of studies of what are called "e-learning principles", he's identified these different principles and he's done studies – mostly lab studies, not classroom studies – of these different principles. But anyhow he also got interested in learning gains and he wrote a book and he did kind of, like a meta-analysis and it turned out there was very little data that showed that especially learning games and mathematics which is what I was targeting really led to any benefits. And the few studies that had been done at that time – of course I'm talking quite a while ago because this is when I first got into learning games – there wasn't much evidence. There was evidence that they were engaged but not that they were learning. So then that's what got me started on the research agenda, to see, well, could I develop a game that would be a balance between learning and engagement? and in particular you're thinking of a causal effect here where the engagement leads to learning, or at least is one factor that leads to learning. When students are engaged they're more likely to learn. So that's why I developed the *Decimal Point* game and eventually we did a study where we compared… we had like a dashboard that would emphasize having fun where they could just pick games that they enjoyed doing, versus one that used like Bayesian Knowledge Tracing to support their learning, to see which of those two led to better learning and which of them led to more enjoyment. (Laugh) Unfortunately we didn't really see a big difference, which kind of surprised me because I thought that the version of the game that was more learning-focused would, you know, lead to more learning, and the engagement-focused one would lead to more enjoyment. We didn't really see that effect. But the point being, the general point being that we were trying to explore just like that tradeoff between learning and engagement and where's the sweet spot between those. I can say that, you know, when my research agenda started, I compared the learning game *Decimal Point* to a tutor. And we did find that there was better learning and better engagement with the learning game. And so my study was one of the very first in the literature in the research literature that actually showed evidence of the benefits of a learning game over you know more conventional learning technology. I'm not sure I'm completely answering your question other than that I acknowledge that that's like… exactly what we're trying to do is find out where's the balance between those. There is a lot of evidence, there was some studies done with a game called Zombie Division by some colleagues of mine in the UK. They actually tested whether or not students learned better if the content was integrated in the game so somehow you make the gameplay use, you know, the content and the way they did it is there were these zombies that had on their chests numbers and then you'd have weapons that you could use to kill the zombies (laugh), a lot of schools frowned on this because of the kind of violent nature of it. But nevertheless, the zombie… so like if a zombie had 10 on its chest you wouldn't want to shoot a weapon that has three because then it doesn't… the idea was that you wanted to get the number

divided evenly. So five and two would be good weapons to have, like whatever weapon, like there'd be this variety of weapons that would be the divisor weapons if you will. So that was a really clever way of integrating the actual numerical, mathematical content in the game. And then you can imagine if it sped up and you had more and more zombies coming at you, you'd have to be get pretty fluent with doing division in a way where, you know, you have to evenly divide numbers. And I took a cue from that, that was actually done before I did my work, and so we also integrate pretty tightly the numerical, mathematical content within the game. And by the way, they were able to show that by integrating the content it was better than if you separated it. So they had a version of the game where you'd play the zombie game, shooting the zombies, but it didn't matter if you evenly divided the number on their chest. And then you do some… you know, after you finish a game, you do some division offline, you know, so that's separating, so you play the game, but the numerical mathematical content is not integrated in the game. So that's… I think one of the key lessons that we've learned about game-based learning is that integrating tightly… and it's a challenge because it's not easy, you have to be creative to think about like, well "how could I get numbers (laugh) integrated in what students are doing, that still makes it sort of fun gameplay?".

– And you saw more efficacy proofs about this type of integration?

BM: Yes, I did, and that study that I told you where we compared the game to tutor, that was the first of that, you know, like we designed the game from the beginning to integrate the mathematical content with the game features in a way, again, building on their research that had already shown that… I think they called it "intrinsic integration" versus "extrinsic" where you play the game that wouldn't have the mathematical content and then afterwards you solve some division problems.

– All right, I saw you had four patents, so what was the objective of these patents and did they lead to licenses?

BM: Sadly none of them actually did, and largely, I guess because nobody on the team who did the patent really wanted to sort of push through and work in a commercial context, including myself. I told you a little bit about my career trajectory, you know, I spent a lot of time in industry, and I was really involved a lot in like, getting products out. And you know, to be honest, I kind of got tired of that, and I felt like it constrained my ability to sort of have more challenging intellectual problems. Because it's a completely different set of problems, I don't want to be little, the problems in industry, but it's more like very practical problems, making sure the software runs right and making sure that you get lots of sales of your software and stuff like that. So you know you always have to, I think whenever you do patents, if you don't have somebody on the team who's going to be like the entrepreneur, oftentimes these things just die in the vine. I will say that none of the patents were my idea to pursue. One of the patents was about the smart glasses that I told you about before, and I think we've just gotten the final approval of that patent. And I'm not sure again if I'm not going to pursue commercializing. I don't know if my colleagues are either. They've both gone on to different things. Except for Vincent Aleven who is the one I said has been working with Steve Ritter and continuing that work. Yeah so short answer is no, they have not moved on but it's really clear that the reason is because there was no entrepreneur on the team who wanted to then like take that patent and say "okay let's make a company, develop a product and make a company out of this".

– Okay all right. Then in general how do you judge from your experience well all the tech transfers and the collaborations in general between AI in education laboratories and educational companies (Edtech, startups, publishers) in the US? What drivers or constraints or difficulties do you perceive?

BM: I mean it's a difficult transition for the… partly for the reason that I said, you need to have somebody who's not just interested in the cool research, but really wants to make a company of it. In my experience too, because I did work for small companies and startups, it also takes a huge effort, of course, you know, it's got to be somebody who's willing to put in long hours and be totally dedicated to the task. And even then you're more likely to fail than you are to succeed. These are all reasons why I decided I didn't want to go down that path. I just, you know… if I'm going to put long hours in, I wanted to be doing research rather than you know that kind of work. And also the high probability of failure kind of turned me off too, because you really have to have the right circumstances things to work. We do have I think in the US a good model for this. I mean, I think probably the US is known better than any country in the world for transferring research to practical use. And I think that's at least in part because of an entrepreneurial spirit that's you know general in the US, maybe a better tax situation so you're not like overly taxed. But then also government agencies like the National Science Foundation and Department of Education they also have these programs, I don't know if you've heard of them, SBIRs – Small Business Innovative Research, I think that stands for. And there the whole idea of that is to prompt people to take their laboratory results and whatever they've developed in maybe a pure research grant and then transfer that into practical use. I'm actually consulting for somebody who's trying to do that right now, again I'm helping and giving advice, but I'm not, you know, the entrepreneur in that particular case. But I think those are the main factors that play a role. And yeah, I guess in my experience and from talking to people, I do a lot of travelling and I have a lot of colleagues around the world, I don't think any place other than the US… I think the US has the best model, but even in the US, it's no guarantee of success (laugh). It's just there's more of a context, you know to make it work. I think a issue in Europe, is largely… you know, and again I'm not criticizing this because personally I wish we had a better sort of model for healthcare and all of that, but I think you know you have high costs and high taxes that you know support a better environment for humans, but that doesn't foster the creativity, I think that you see more in the US, which is more like capitalistic. And we don't have that same model. And I'm sure you're well aware of that, not telling you anything you don't already know. So there's an upside and a downside to that, I think the model in the US creates more, you know, entrepreneurial spirit, it leaves more space for people to take things to product, but it leaves a lot of people behind too, okay. Whereas you, in Europe you take better care of the people, but that also doesn't support business quite as well.

– Did you also notice that when you traveled to Germany or France?

BM: Yeah I particularly… because I have lots of European colleagues and I lived in Germany for a while and so I got used to the different kinds of context and the different social environment and… yeah from a personal point of view I actually prefer the European model to be honest. But I do think as I said before that in terms of supporting startups and companies and so on, it's better in the US, with respect to making a success of those companies. Of course a lot of them fail too and then they go by the way side too (laugh)…

– Okay and then my very last question: how do you perceive all the federal and the states standpoints regarding AI in education? Have these standpoints evolved through times when you look at grant research programs?

BM: Yeah so I think we have a pretty good support system. I mean, I personally have been pretty successful at getting grants for, you know, pushing the AI and education agenda… Maybe it's even gotten better. (Laugh) The pandemic may have actually helped a little bit, because of course everything had to switch to online education, and a lot of people of course didn't like that. But bottom line is that there are lots of situations like people who live in rural areas or, you know, people who don't have the ability to travel where providing online education is a huge benefit and I think that became more obvious during the pandemic when everybody had to switch to that. And I think it's made the US, you know, in terms of… I must admit I haven't tracked this carefully, but I bet if I looked at what Congress has been providing, that it's gone up in terms of the money that's been provided to Department of Education and National Science Foundation for instance. We get very little… I've never gotten any state funding, so state funding is pretty sparse for AI in education. The other place where we get funding in the US are foundations. Maybe those are to be honest, I don't even really know what the situation is in Europe with respect to that, but there are lots of foundations like the Gates foundation, you know, like Bill Gates and Melinda Gates started a… and part of the… I think they have three different areas that they focus attention on, and one of them's education, not AI in education, but education. So I know lots of people, I haven't personally gotten grants from the Gates foundation, but I know people who have. And then there are other foundations too – that's probably the biggest one –, but there are a bunch of other ones… I don't know, have you heard about this Learning Tools competition that's going on in the US right now? That's funded by foundations, I think Gates is one of the foundations, it's multiple organizations, and basically they're giving relatively small grants to lots of people who have new ideas with hopes that you know and then they'll pick a few that will move on to the next stage. I don't know how successful that is, I don't play in that arena much because money is, you know, for me, is a little bit too small, but I think it's really helpful for somebody who has a new idea and is really just trying to get something off the ground. For me the work exceeds the pay (laugh). Yeah I've been pretty successful and have a good reputation, I think with NSF and Department of Ed, so that's kind of my bread and butter as far as getting support. But you know that is another – back to your original question about you know the way tech gets supported and moving into industries –,I think that's what the foundations who are supporting that program are really all about. They want to support, you know, people with new ideas, small ideas. But there might be a few diamonds in the rough that can then get better supported and ultimately become companies.

Entretien 4. Laurent Jolie, le 7 mars 2023

– *Nous nous étions vus fin 2021* [à l'occasion du salon Educatech-Educatice]*, alors que Lalilo venait d'être racheté par Renaissance. D'un point de vue personnel, comment ton travail de CEO a-t-il évolué depuis ? Comment répartis-tu ton temps entre le marché français et les États-Unis ?*

Laurent Jolie : D'un point de vue personnel, moi je développe surtout Lalilo sur le marché français. Du point de vue de l'entreprise elle-même, l'entreprise qui nous a racheté a une grosse flotte de vendeurs donc elle s'occupe beaucoup de la commercialisation aux États-Unis, c'est globalement des gens qui sont directement rattachés à cette entreprise qui s'en occupent. Comment est-ce que les choses ont changé ? Bon d'un point de vue général on a une stabilité financière qui est beaucoup plus grande. On a une grosse boîte derrière nous donc c'est agréable. Ensuite d'un point de vue développement de l'entreprise, forcément tu as l'avantage et l'inconvénient d'un grand groupe. On a une inertie qui est un peu plus grande, c'est un peu plus dur de prendre des décisions, d'avancer, il y a des processus de décision qui sont plus lourds.

– Tu veux dire en termes de décisions produit ?

LJ : Oui, d'une manière générale, ça peut être sur des décisions produit mais ça peut être sur un peu tout et n'importe quoi, même sur des décisions business effectivement. D'un point de vue produit, les directions produit n'engagent plus uniquement nous, donc forcément c'est un peu plus long de prendre des décisions.

– Ensuite j'ai une question sur la répartition actuelle de la base d'utilisateurs de Lalilo entre la France, les États-Unis. Est-ce que tu as des chiffres ?

LJ : Il faudrait que je regarde… je pitche en permanence qu'on a la moitié de nos utilisateurs aux États-Unis, 25% en France 15% au Canada et 10% dans le reste du monde, mais pour être assez honnête je pitche ça depuis deux ans donc (rires) il se peut que ça ait évolué, mais l'ordre de grandeur doit être celui-là. Je connais pas bien la dernière évolution mais si ça t'importe beaucoup je peux aller regarder les derniers chiffres. En gros la plus grosse partie de nos utilisateurs sont aux États-Unis.

– D'accord. J'aimerais maintenant en venir plus particulièrement au marché américain des environnements de early literacy. *Comment délimites-tu le marché de Lalilo aux États-Unis ? Qui considères-tu comme étant vos concurrents ?*

LJ : C'est intéressant. À nouveau j'ai un peu deux ans de retard sur toutes ces questions. Donc nous on va cibler du K-2, donc des élèves qui sont surtout en *kindergarten, first grade* et *second grade*. Notre principal concurrent ce serait Lexia, qui a tendance à faire un peu plus que du K-2, donc on considère qu'on est un peu meilleurs sur un segment plus spécifique. Après sur des gens qui sont plus spécifiques K-2, tu as un truc qui s'appelle Reading Eggs, un truc qui s'appelle Reading A-Z, tu as plein d'applications comme ça. Mais en réalité on s'est pas mal développés parce qu'en K-2 on était assez *pure player*, c'est-à-dire que nos concurrents avaient tendance à faire quelque chose de beaucoup plus large sur l'apprentissage de la lecture, et nous en étant très spécifique on a pu faire un produit mieux adapté.

– Quand tu parles de qualité, c'est en termes de quoi ?

LJ : Pédagogique. Non alors pédagogique et design. Vu qu'on a eu cette approche où on s'est connecté très vite à l'utilisateur, on a un produit qui a une plus forte capacité à être utilisé, et donc un peu mieux *designé* je pense d'une certaine manière.

– Le PDG de Toca Boca a dit que son plus grand concurrent était YouTube, mais que ça pouvait être aussi des studios de jeux vidéo, et que la comparaison pouvait être un peu difficile notamment du point de vue design, sur des applications que les enfants utilisent déjà. Est-ce que tu considères ces acteurs non-éducatifs au premier abord comme des concurrents ?

LJ : Non ce ne sont pas des concurrents, car on est vraiment sur des populations plus jeunes, et en réalité l'usage de l'ordinateur dans l'école est assez encadré. Ce ne sont pas des âges où on va dire aux enfants « Allez sur l'ordinateur et faites ce que vous voulez » et l'enfant va sur YouTube. Globalement... tu as quand même ça, aux États-Unis, ce truc d'« Allez sur la tablette et faites ce que vous voulez », mais c'est sur un choix restreint d'applications. Donc nos concurrents, ça va être justement les autres applications que je te citais, qui sont éventuellement en *freemium* et qui sont

disponibles sur ces tablettes de base. Et donc les élèves, ou bien ils aiment aller sur Lalilo ou bien ils n'aiment pas, et dans ce cas-là notre concurrent est lié au choix de l'enfant. Mais dans la plupart des cas c'est vraiment le prof qui va choisir ce sur quoi va travailler l'élève dans la salle de classe aux États-Unis. En tout cas c'est vraiment le vécu qu'on a. Je ne te dis pas qu'il n'y a pas des petites exceptions, mais le cas majoritaire qu'on observe c'est ça.

*– D'accord. Et mis à part la pédagogie et le design, quelle importance stratégique, en termes d'avantage concurrentiel, accordes-tu à l'IA (*adaptive learning*, reconnaissance vocale) sur ce marché ?*

LJ : Alors c'est deux sujets différents. La reconnaissance vocale, pour le traiter d'abord, on a mis beaucoup d'efforts dedans. L'*adaptive learning* c'est un peu le cœur de notre proposition de valeur de dire « on va s'adapter à chaque élève ».

– Et quelles perspectives en termes économiques, de concurrence, cet argument-là peut ouvrir pour un acteur comme Lalilo ?

LJ : Ce qui est un peu difficile c'est que c'est un truc qui se markette mal. C'est à dire que l'*adaptive learning* c'est assez dur d'expliquer aux gens quels sont les différents niveaux d'*adaptive learning* que tu peux mettre dans un produit. Tu vas avoir des gens qui vont avoir une progression et avoir les élèves en différents endroits dans cette progression et en fait tu peux dire que tu fais de l'*adaptive learning*, tu t'adaptes à l'élève, chaque élève est dans sa progression à un endroit différent. Du coup c'est assez désagréable parce que cette personne va dire si elle trouve ça bien d'un point de vue marketing « *adaptive learning* » et en fait elle va te dire « ouaip mais en fait on fait un peu mieux »… et en fait c'est assez intangible. Du coup, du point de vue proposition de valeur : assez dur à positionner. On est beaucoup sur la perception de nos utilisateurs sur ce sujet-là, et franchement cette perception elle est hyper biaisée, c'est assez frustrant. Mais biaisée positivement. Une grande bataille que j'ai essayé de mener c'est qu'on vende non pas de l'*adaptive learning* mais de l'efficience pédagogique, de dire « Si tu utilises Lalilo plutôt qu'un autre outil pédagogique, voilà ce qui va se passer pour les résultats scolaires de tes élèves ». Et on est encore dedans, il faut continuer de faire des études, montrer… c'est un très gros sujet de réussir à témoigner de ce qu'un outil pédagogique apporte, et personne n'arrive à bien témoigner de ça je trouve.

– Je veux bien en venir maintenant à la distribution. Tu as commencé à parler de la commercialisation du produit aux États-Unis. Vous avez opté pour un business model freemium (gratuit pour les profs, premium pour les écoles et districts). Quel est le contenu de l'offre premium, le prix et le mode de facturation ?

LJ : Alors le contenu de l'offre *premium*, en gros : on a quelques fonctionnalités en plus surtout pour les enseignants. Le modèle, on est récemment en train de le changer pour un *free trial*. Avant c'était sans limitation de durée, les élèves pouvaient essayer la plateforme. Maintenant il y a une limitation de durée. Globalement on te montrait un peu les premières lignes de ton tableau de bord mais pas les dernières... enfin c'était plus sur le *reporting* prof qu'on mettait des blocages. On vendait par prof, puis depuis qu'on a été rachetés on s'est alignés sur un modèle par élève, qui est un modèle majoritaire de la boîte qui nous a rachetés.

– Ca, c'est à l'abonnement…

LJ : Annuel.

– Et pourquoi avoir fait ces changements ?

LJ : Alors sur le changement *free trial*, c'est parce que on a fait pas mal d'études sur ce sujet

– J'ai une question sur un choix que vous avez fait il y a longtemps, celui d'un environnement accessible sur le web, et non d'une application distribuée sur les app stores. Pourquoi ce choix ?

LJ : (Hésitation puis rire) alors je pense que tous les mois on remet ça sur la table… alors pourquoi ce choix ? Parce que ça crée beaucoup plus de complexité… parce qu'on a une vraie diversité de matériels présents dans les écoles (en France), on a beaucoup d'ordinateurs, je pense que ça doit être 70% d'ordinateurs, 30% de tablettes. Et donc en fait le coût supplémentaire de maintenir quelque chose pour une tablette est assez élevé. Un peu compliqué à gérer aussi avec pas mal de problèmes de connexion. En fait quand tu as un prof qui va te faire remonter un problème, le premier truc que tu vas faire en support utilisateur c'est de lui dire « Est-ce que tu es sûr que tu as la dernière version ? ». Donc déjà avant d'identifier si on a un bug, c'est d'identifier si la personne a bien sa dernière version. Bref c'est un rapport coût de développement/opportunité qui mène à ça. Et si notre truc est utilisable sur tablette. Alors pourquoi pas sur les stores ? De toute façon on ne vend pas via les stores. On est sur un modèle *freemium* donc les profs pour qu'ils puissent y accéder, les tablettes elles arrivent quand même à accéder à notre produit via n'importe quel navigateur internet. Donc en fait la proposition de valeur c'est d'avoir une app native, donc derrière on a un peu hacké le truc, on dit aux gens « Vous pouvez faire un raccourci sur notre page sur le fond d'écran ». Il y a plein de petits trucs qui pourraient être mieux si on faisait du natif, et même la somme de ces petits trucs ne justifient pas de faire du natif. À la fois en coût initial pour développer ça et en coût de maintenance.

– Du point de vue produit, en écho à une difficulté souvent évoquée pour la early literacy *: comment arbitrer entre les deux publics à satisfaire : profs et enfants ?*

LJ : Je ne sais pas trop, de quel arbitrage tu veux parler ? Si on a une fonctionnalité à ajouter sur l'interface élève et une fonctionnalité à ajouter sur l'interface prof, laquelle on choisit, c'est ça ?

– Oui et puis en termes d'expérience et d'engagement, comment arbitrer entre l'engagement d'un public et celui d'un autre ?

LJ : De toute façon l'engagement des deux va être important. Si les élèves n'aiment pas notre produit il n'y a pas beaucoup de profs qui vont l'utiliser et s'il n'y a pas beaucoup de profs qui aiment notre produit ça ne va jamais aller dans les salles de classe. Donc de toute façon il faut être bon sur les deux. Nous on a deux interfaces à part. Ce qui est assez marrant c'est qu'en ce moment notre organisation c'est d'avoir deux équipes produit qui travaillent sur chacun des deux sujets, donc pour ça en plus les *roadmaps* ne sont pas en concurrence l'une avec l'autre. Et après c'est de la priorisation classique que n'importe quelle entreprise qui a créé un département de R&D a : « Est-ce que c'est mieux la *feature* A ou B ? », (rire) c'est des choix au cas par cas, selon la direction dans laquelle on va à un instant t, la priorité du moment, la valeur créée de manière générale... bref c'est un peu tout ça.

– Quelle est la place des parents dans le produit ?

LJ : Nous ce qu'on a refusé de faire c'est du *BtoC*, de vendre aux parents notre produit. On était très fortement poussés et du coup on dit à l'enseignant s'il le souhaite qu'il peut partager notre outil avec les parents pour du travail à la maison ou autre. Dans la pratique, on a quelques parents qui utilisent quand même notre outil sans passer par le prof même s'ils ne sont pas censés pouvoir le faire… mais on mesure bien ce genre de truc, on ne bloque pas non plus totalement. En tout cas, ce qu'on essaye de faire c'est de ne pas se concentrer sur les parents.

– D'accord, vous n'allez pas trop sur le parascolaire ?

LJ : Non on ne fait pas de parascolaire. On ne commercialise pas dessus. En fait il y a une fonctionnalité qui permet de partager avec les parents, ça a été hyper utile pendant le Covid, parce qu'il y avait un besoin d'usage à la maison assez fort donc on préfère que les parents soient… mais on le fait vraiment dans un cadre scolaire... [*la connexion devient mauvaise*].

– Question lexicale : quelle est votre perception de l'expression « intelligence artificielle » ? Est-ce qu'elle te paraît appropriée pour désigner l'activité de Lalilo ?

LJ : Moi ma perception c'est de parler d'une activité qui est habituellement faite par un être humain et qui est faite ou soulagée par l'usage d'un algorithme. Donc c'est plus par l'usage que je parle d'intelligence artificielle que par la technique qui est dessous. Est-ce qu'elle me semble adaptée pour Lalilo ? Oui sur nos deux briques en question. Le choix de l'activité d'un élève c'est vraiment quelque chose qu'un être humain fait d'ordinaire dans une salle de classe, et le fait de faire lire les élèves à voix haute et donner un feedback immédiat sur leur lecture c'est ce que fait un adulte qui apprend à lire à un enfant, donc dans ce sens-là je pense que c'est le cas. Ce qu'on essaye de faire c'est un bon produit pédagogique pour aider les enseignants, et franchement « IA » ce n'est pas... enfin on ne tient pas particulièrement à ce vocabulaire, c'est le Ministère qui y tient.

Entretien 5. Gérard Puimatto, le 30 mai 2024

– Je suis assez intéressé par la question du GAR, notamment du point de vue des entreprises de l'IA éducative, et la manière dont ce dispositif conditionne leur développement dans l'appareil éducatif français.

Gérard Puimatto : Peut-être qu'il vaut mieux que je vous redise un petit peu d'abord ce que c'est que le GAR, et comment il est arrivé. En gros le GAR aujourd'hui, ça fait 5-6 ans qu'il est en exploitation, mais son histoire remonte bien avant ça puisqu'avant il y a eu des préfigurations. Il y a eu des travaux conduits et cetera, et bien avant encore, j'ai participé à la création dans la région Paca d'un dispositif qui s'appelait Corrélyce qui ne faisait pas tout à fait la même chose mais qui a quand même en grande partie inspiré le GAR ensuite. Donc le principe étant qu'il fallait qu'on ait un processus d'identification unique et qu'il n'était pas question que chaque éditeur mette son propre processus d'identification pour les élèves, pour les enseignants aussi d'ailleurs. Alors pourquoi ? simplement parce que l'idée que veulent promouvoir les éditeurs en gros c'est que la ressource s'adresse à chaque enseignant et chaque élève. Et en fait dans le système scolaire français au moins et dans un certain nombre d'autres systèmes scolaires, ce n'est pas comme ça que ça se passe, parce qu'entre les deux il y a un établissement scolaire et la contractualisation avec l'éditeur etc. elle relève

de l'établissement scolaire et pas de l'élève ni de l'enseignant. Et du coup l'établissement scolaire a une responsabilité sur peut-être des contrats qui sont passés pour choisir des ressources, sur le choix des éditeurs, sur la relation à la responsabilité de l'éditeur qui attache l'établissement et en particulier sur l'utilisation des données personnelles. Donc le GAR, c'est un dispositif qui a été fait pour l'application du RGPD. Donc le RGPD nous dit, sommairement, qu'il y a des données personnelles qui sont celles de mineurs, en plus quand on a des élèves de collège, ils sont « mineurs numériques », donc, même si ce n'est pas des données sensibles au sens du RGPD, ces données peuvent être rapidement critiques et évidemment l'établissement scolaire a une responsabilité sur ces données. Alors il se trouve que l'analyse qu'il y eu au moment de la publication du RGPD était de dire « ok l'établissement scolaire a une responsabilité sur les données de ses élèves ». Juridiquement c'est le cas. Mais aucun établissement scolaire en France, ou à peu près, n'est capable de remplir tout seul les obligations RGPD pour ces contrats. Par exemple, il faudrait faire tenir à jour un registre de traitement, faire des analyses d'impact par ressource, etc. C'est à peu près irréalisable pour un établissement. Et du coup le ministère a envisagé de déporter cette responsabilité au niveau national pour soulager les établissements de ce bazar. Du coup le GAR a deux aspects : il y a un aspect technique de gestion de l'identification unique, et puis l'aspect juridique, extrêmement important, probablement plus important que le côté technique, parce qu'il déporte la responsabilité de la gestion des données personnelles au niveau du ministre. Donc le traitement de données GAR – au sein du RGPD un traitement de données, ça a une finalité – donc la finalité, c'est de permettre l'accès aux ressources dans des conditions de confidentialité correctes pour les données, de protection des données. Et c'est la gestion de ces données avec une responsabilité de traitement qui appartient au ministre (pas nominativement, mais par sa fonction). Le ministre a pris un arrêté définissant ce qu'on appelle le « traitement de données GAR », pour lequel donc il est dit que c'est le ministre qui est responsable du traitement. Ce qui implique que tout éditeur qui a besoin de données pour faire fonctionner ses ressources devient un sous-traitant du ministre pour les données. Ça c'est la structure du GAR. La structure juridique, qui définit le ministre comme responsable de traitement est beaucoup plus importante que le côté technique. Donc, si je suis éditeur, je passe un contrat de vente de ma ressource avec un établissement : le GAR va m'envoyer les identifiants des usagers (enseignants, élèves, personnels). Ok, mais en fait c'est le ministre qui me confie ces données pour faire fonctionner la ressource, en restant le responsable de traitement. Et, en tant qu'éditeur, je suis sous-traitant du ministre. Alors c'est une conception qui a eu du mal à s'imposer, parce qu'avant que le RGPD soit publié, il y a eu des tas d'ateliers avec des éditeurs, qui ont commencé par beaucoup contester cette position, en disant que ce qui les intéressait, c'était d'avoir une relation directe avec les enseignants et les élèves, qu'ils ne voulaient pas, au milieu, un acteur ministériel. Et puis finalement quand le RGPD est paru, il est bien apparu que ce n'était pas possible de faire autrement. Donc ça s'est imposé en fait. Alors sur le plan juridique, je pense que c'est extrêmement important, parce que c'est ça a changé complètement la conception des éditeurs. À commencer par les plus gros, qui sont les gros éditeurs scolaires bien sûr. Donc les Hachette, Editis, eux ils ont tout de suite reconstruit leur démarche en fonction de cette logique RGPD avec l'utilisation du GAR. Ce n'est toutefois pas systématique parce que le GAR n'est pas obligatoire : le GAR est un service qui est offert aux établissements via leur ENT donc il n'est pas obligatoire dans la loi. Il pourrait le devenir d'ailleurs, il y a une discussion qui est en cours en ce moment au niveau du ministère sur les nouveaux référentiels, où il pourrait peut-être devenir obligatoire. Pour l'instant, ce n'est pas le cas. Mais il est massivement adopté. Pour le second degré, en tout cas, on pourrait dire 95 % des établissements. Pour le 1er degré, c'est beaucoup plus compliqué parce que les ENT sont très rares. En tout cas, cette notion de sous-traitance, elle est vraiment très importante. Alors après, si je suis éditeur et que je veux commercialiser une ressource. Cette ressource doit faire l'objet d'une qualification. Non pas sur le plan fonctionnel, pour lequel

l'éditeur a une totale liberté, ni sur la qualité de la ressource. En revanche, l'évaluation qui est faite porte sur le traitement des données. D'abord pendant le développement on suit ce que fait l'éditeur, régulièrement, et à la fin, une fois qu'ils considèrent avoir fini leur développement, ils remplissent ce qu'on appelle une déclaration de conformité. Dans cette déclaration de conformité, on leur demande de décrire de quelles données dont ils ont besoin d'abord, ce qu'ils en font, comment ils les traitent et comment ils en assurent la sécurité. Donc il y a un certain nombre de règles qui ont été posées petit à petit avec des éditeurs, en collaboration avec eux d'ailleurs, pour savoir quelles étaient les conditions pour une conformité, pour une sécurisation suffisante et puis l'application des principes du RGPD. C'est à dire qu'on doit minimiser les données transmises, on ne transmet les données de l'élève ou de l'enseignant que si elles sont indispensables au fonctionnement de la ressource. Par exemple si j'ai une ressource qui est un manuel scolaire anglais 5e qui ne fait pas de suivi d'activités de l'élève, on transmet juste l'identifiant opaque qui est calculé par le GAR, et c'est à peu près tout. Si par contre on a des logiques de classe par exemple, de travail entre le prof et une classe, ou des choses comme ça, on peut transmettre la division, c'est à dire le code de la classe, par exemple 5e A, 5e B. On peut transmettre si nécessaire le groupe, on peut transmettre les options, etc. jusqu'aux noms/prénoms si nécessaire. Et toutes les fois qu'une ressource fait de l'assignation de tâches (si dans une ressource, je peux dire à tel élève « tu fais tel travail »), j'ai besoin d'avoir son nom et son prénom sinon ça ne peut pas fonctionner. Donc on peut fournir jusqu'aux nom et prénom uniquement si c'est nécessaire. C'est ce qu'on appelle dans le RGPD le principe de proportionnalité, c'est à dire qu'on fait des transferts de données de manière proportionnelle aux besoins de la ressource. Et cette démarche est validée par la CNIL bien-sûr, conforme au RGPD etc. Il y a eu un effet complémentaire aussi, c'est qu'on a vu se développer beaucoup de ressources qui utilisent des données personnelles pour les travaux d'élèves par exemple. Donc le prof commande un travail à l'élève, l'élève produit quelque chose dans la ressource etc. Et du coup il y a eu un deuxième arrêté qui a été pris avec l'avis de la CNIL, concernant ces données qui sont les données produites dans la ressource. Par exemple, je demande à un élève de faire une rédaction sur un sujet, en fait la rédaction au sens juridique, c'est une œuvre. Donc une œuvre qui a un auteur et dont la jurisprudence considère que c'est aussi une donnée à caractère personnel. Et du coup il y a eu une extension de l'arrêté pour dire que ces données produites à l'intérieur de la ressource étaient aussi sous la responsabilité du ministre. Cette évolution a été intégrée dans le GAR, avec un effet direct auprès de divers acteurs. Il y a un effet dans les établissements scolaires parce que les personnels de direction sont soulagés de ne pas avoir à gérer tout ce bazar qui est absolument énorme pour eux. Il y a un effet pour les collectivités parce que le GAR s'appuie sur les ENT. C'est un service qui est accessible à partir des ENT et donc toutes les collectivités qui ont un ENT aujourd'hui ont adhéré aux GAR, parce que c'est un complément fonctionnel qui leur convient bien et qui leur évite d'avoir à gérer tout ça. En région PACA, il y a à peu près 15 ans, ou un peu plus, presque 20 ans, on avait Corrélyce qui gérait les accès pour les lycées de la région PACA. Maintenant, ça n'a plus de sens puisque le GAR le gère de manière globale. Du coup toutes les collectivités pour le second degré ont adopté ce système-là, sans aucun problème. Ça veut dire que tous les éditeurs d'ENT l'ont adopté aussi. De toute façon, s'ils veulent vendre aux collectivités, il faut qu'ils répondent aux appels d'offres et dans les cahiers des charges des appels d'offres, il y a la conformité au GAR. Donc du coup tous les ENT ont basculé du côté du GAR quasi-naturellement. C'est important sur l'acteur collectivité, parce que c'est aussi lui qui paye les quantités de ressources en fait. Si on veut qu'une ressource soit présente dans des établissements de manière significative, et s'il n'y a pas de financement du côté de la collectivité, il n'y aura pas grand-chose. Du coup grâce à ça, un certain nombre (pas la totalité mais peut-être un sur deux) des acteurs collectivités se sont dits « complémentairement à notre marché ENT, il faut aussi qu'on prévoie un appel d'offres pour des ressources, enfin une politique de ressources en tout cas, pour savoir comment on met des ressources

à disposition des établissements scolaires ». La première région qui a fait ça c'est la région Grand Est, qui l'a fait de manière massive pour remplacer tous les manuels scolaires par du numérique. Et puis progressivement en quelques années toutes les régions et la moitié des départements ont basculé sur ce modèle-là. Du coup, cela veut dire que pour les ENT, le GAR a eu un effet très modélisant en fait. Pour les collectivités, ça a été très modélisant parce qu'aujourd'hui c'est très sécurisant pour une collectivité de faire un appel d'offres pour les ressources en disant « Voilà, on vous donne des crédits pour acheter des ressources. Vous achetez des ressources que vous voulez sous réserve qu'elles soient compatibles GAR ». Donc c'est une sécurité évidemment pour la gestion des données. C'est très modélisant aussi parce que ça structure le marché. Parce qu'à partir du moment où tous les marchés écrivent « compatibilité GAR » partout, c'est extrêmement structurant pour le marché et du coup pour les éditeurs aussi, parce que les éditeurs se disent « si je ne suis pas compatible GAR, je n'y arrive pas ». Du coup ils sont obligés de s'aligner sur cette logique-là. Alors parmi les éditeurs, il y en a qui ont très naturellement adhéré en disant « Bah voilà, c'est comme ça, le ministère a décidé ça, on y va ». Et il y en a qui traînent des pieds, clairement. De moins en moins quand même. Les premières années il y en a qui traînaient vraiment beaucoup des pieds. Maintenant je pense que c'est rentré dans les habitudes, ils savent faire techniquement, c'est passé donc voilà. Je pense que les choses se sont bien décantées, y compris les gros éditeurs scolaires qui ont basculé assez simplement dans ce modèle. Mais du coup je pense que ce qui est important pour vous probablement c'est le côté extrêmement modélisant. Parce que c'est modélisant pour l'offre, ça veut dire en gros c'est les collectivités qui achètent mais qui achètent établissement par établissement donc il y a une politique d'acquisition dans l'établissement scolaire. Et puis c'est très modélisant pour la manière dont on gère la relation aux données puisque du coup les éditeurs effectivement disposent de tas de données sur les élèves mais dans un cadre contractuel extrêmement contraint et on ne peut vraiment pas en faire n'importe quoi. Parce que si vous prenez Hachette par exemple, il a le portail KNE pour toutes les ressources numériques. KNE-Hachette suit un élève pendant toute sa scolarité, du CP jusqu'à la terminale, voire après. Et à partir du moment où il a un identifiant unique pour un élève, ça veut dire qu'il peut savoir tout ce que l'élève a fait pendant toute sa scolarité. Alors, je vous laisse imaginer tout l'intérêt que ce genre de données pourrait avoir dans des procédures marketing avec une utilisation du *Big Data* ou des choses de ce genre. Donc il fallait absolument borner ça, et c'est borné par le contrat que signent les éditeurs. Tous les éditeurs, pour adhérer au GAR, signent un contrat dans lequel ils s'engagent dans un certain nombre de choses, pour la protection des données et pour ne pas les utiliser en dehors du cadre de fonctionnement de la ressource. Voilà, sommairement c'est à peu près ça.

– D'accord, et du point de vue des plus petits acteurs et des start-ups qui portent des produits d'intelligence artificielle, on sait que pour elles les données d'apprentissage sont un peu le nerf de la guerre pour entraîner leurs algorithmes. Alors est-ce qu'il y a des critères ou des contraintes particulières pour ces acteurs-là ?

GP : Alors deux choses. D'abord des start-ups et la plupart des petits éditeurs d'ailleurs de ressources numériques, ils ont tous un comportement particulier. D'abord, ils sont tous, absolument tous, persuadés qu'ils ont la *killer application* qui va résoudre tous les problèmes de l'éducation. Pris individuellement ils ont tous la solution à tous les problèmes possibles de l'éducation. Du coup en gros ils ne comprennent même pas qu'on leur impose quelques contraintes que ce soit puisque, de toutes façons, eux, ils ont LA solution. Alors ça c'est au début quand ils arrivent en fait, il y a toujours une phase difficile au début, ils arrivent avec leur solution qui est extraordinaire et ils ne comprennent pas pourquoi le ministère ne saute pas dessus en fait. Donc au début, il peut y avoir des résistances, il y en a régulièrement pour beaucoup de start-ups qui ne sont pas passées par un canal

un peu structuré pour comprendre un peu le marché. Aujourd'hui, il y a les *EdTech* par exemple qui discutent beaucoup entre eux, ils sont au courant, il n'y a pas de problème. C'est une association, ils sont très opérationnels. L'AFINEF c'est pareil. Ces gens-là, ils ont des associations, ils savent, ils n'ont pas besoin de beaucoup d'infos complémentaires. Par contre pour tous ceux qui passent à côté, par exemple pour ceux qui passent par les incubateurs qu'il y a dans quelques collectivités, les débuts sont un peu compliqués parce qu'il faut qu'ils comprennent que l'éducation n'est pas un marché comme un autre. Donc ça c'est le premier problème. Alors après il y a aussi le problème de la relation à l'intelligence artificielle. Enfin l'*adaptive learning* d'abord, l'intelligence artificielle derrière. Le ministère a lancé, il y a deux ans je crois maintenant, une expérimentation sur l'intelligence artificielle qui n'est pas inintéressante mais qui ne va pas très très loin non plus. Bon auparavant la CNIL avait mis des bornes et nous avait demandé de ne pas accepter les ressources « intelligence artificielle ». Il y a deux ans à peu près, on a commencé les accepter. Je suppose que vous avez regardé le dossier technique CNIL sur l'intelligence artificielle… il y a tout un tas de critères, de risques et cetera. Il est plutôt bien fait, mais il fait un peu peur. Donc, il y a deux ans, on a commencé à accepter les ressources qui avaient de l'intelligence artificielle en regardant un petit peu comment fonctionnaient les données d'entraînement et cetera. Globalement, on s'est rendu compte qu'un grand nombre d'éditeurs, enfin ce n'est pas une offre monstrueuse non plus, les gens qui font de l'*adaptative learning,* il n'y en a pas des quantités non plus… mais globalement dans ceux qui en font, sur 100 qui nous disaient « Oui, j'ai une intelligence artificielle », quand on regardait de près l'intelligence artificielle, il y en avait dans trois ressources. La plupart utilisent des processus d'évaluation formative extrêmement traditionnels avec des données statistiques basiques et sans apprentissage d'intelligence artificielle. Après pour eux, l'intelligence artificielle c'est un argument marketing… pourquoi pas, mais bon. Après, on a quelques acteurs qui utilisent de l'intelligence artificielle, soit directement soit par des sous-traitants. Parce que je pense que, pour exploiter directement, il faut avoir de sérieuses compétences quand même. Donc souvent en passant par des sous-traitants. Un des premiers c'est Lalilo qui est un logiciel d'apprentissage de la lecture, très bien conçu d'ailleurs, et qui s'appuie exclusivement sur les sous-traitants. Quand on voit le nombre de sous-traitants de Lalilo, il y en a une vingtaine. Mais par contre là, c'est vraiment une intelligence artificielle, il y a vraiment un usage pertinent. Et l'autre gros acteur, plus gros que Lalilo d'ailleurs, qui fait de l'intelligence artificielle et qui le fait plutôt bien, c'est Evidence B. Après évidemment, il y a la question de données personnelles et je pense que si on prend le cas d'Evidence B, ils sont bien conscients que tout ça ne fonctionne qu'avec des données personnelles puisque les traces d'activité qu'on peut avoir, que laisse un élève pendant l'usage de la ressource, c'est toujours des données personnelles. Donc c'est contraint par des règles très précises, en particulier de non-utilisation en dehors du cadre de la ressource et cetera. Et si on regarde du côté de l'IA en fait, on a un processus qui est à la fois très simple et compliqué. C'est-à-dire que je suis élève, j'utilise la ressource donc je fais des traces d'activité, ces traces d'activité, elles sont liées à mon identifiant GAR et éventuellement mon nom et mon prénom si je les ai donnés. Dès que ces traces d'activité vont rentrer dans un modèle de données, en fait, elles sont pseudonymisées. On dit souvent anonymisées mais en fait c'est très rarement anonymisé, c'est plutôt pseudonymisé, parce qu'au terme du processus on est susceptible de rétablir le lien entre les données et l'utilisateur pour lui faire une recommandation. Donc c'est bien une pseudonymisation et une dé-pseudonymisation derrière. Et en fait, c'est ce processus de pseudonymisation qui aujourd'hui est le plus critique, parce qu'évidemment il peut y avoir des failles de sécurité monstrueuses. Alors les gens sérieux, professionnels de l'IA, connaissent et maîtrisent bien, mais les bricoleurs qui arrivent, il y en a quand même un petit peu dans les *EdTech*, il faut beaucoup plus se méfier parce qu'ils n'ont pas toujours des processus vraiment très fiables là-dessus. C'est vraiment le point critique je pense.

– Alors est-ce qu'il y a également un travail de négociation ou d'accompagnement en matière de modèles économiques auprès de ces acteurs-là ?

GP : Pas du tout. C'était un refus massif de la profession. En 2015-2016 quand on a fait ces préfigurations, on a eu des quantités de réunions avec les éditeurs, évidemment leurs représentants, donc le Syndicat national de l'édition, l'association Savoir lire, l'AFINEF, à l'époque il n'y avait pas encore les *Edtech*. Et leur souci était de dire « Ok, on veut bien comprendre qu'on est sous-traitants sur les données mais on veut préserver notre autonomie éditoriale donc il est hors de question que le ministère donne un avis sur la qualité pédagogique et il est hors de question que le ministère participe à quoi que ce soit sur le plan de la commercialisation et de la diffusion économique ». À tel point qu'ils nous ont demandé à l'époque de ne pas faire de catalogue. Le catalogue, il y en a un interne au GAR évidemment, mais il n'est pas publié. Alors on est dans la grande tradition du manuel scolaire en fait, parce que cette ce modèle-là, c'est le modèle qui avait été inventé par Louis Hachette et François Guizot au XIX^{e} siècle. Quand ils ont inventé le manuel scolaire, sommairement, ils se sont répartis la tâche en considérant déjà que l'éditeur avait sa responsabilité pleine et entière et que le ministère n'avait pas de contrôle sur l'éditeur. Du coup les éditeurs, ils ont bien compris ça et campent là-dessus en considérant que leur liberté éditoriale est vraiment essentielle. Sinon ils n'existent plus en fait.

– Et le catalogue qui existe de manière interne porte sur quel aspect ?

GP : De manière, on a besoin de savoir quels sont les produits compatibles GAR, parce que mettons vous êtes dans un ENT, vous demandez à ouvrir Lalilo, il faut que derrière le GAR sache que pour activer Lalilo, il faut taper sur tel URL et cetera.

– D'accord mais il y a une garantie qu'il n'y ait pas de publication…

GP : Tout à fait. Alors c'est toujours en discussion parce qu'en fait d'un côté les éditeurs ne veulent pas de catalogue et d'un autre côté quand les régions passent un marché, mettons la région Grand-Est passe un marché pour les ressources GAR et la question c'est « Mais où est-ce qu'on trouve la liste des ressources ? ». Donc du coup il faut trouver un libraire qui va gérer un catalogue des ressources disponibles. Donc c'est comme ça que ça marche, d'ailleurs, c'est la librairie LDE, la Librairie de l'éducation, qui fait son propre catalogue. Ce qui est un peu aberrant, parce qu'on pourrait très bien publier le catalogue ou tout au moins le donner aux collectivités. Mais bon pour des raisons historiques, ce n'est pas possible. Du coup ça veut dire que le ministère n'a pas de… en dehors de quelques appels d'offres qu'il peut faire mais qui ne représentent vraiment pas grand-chose, le ministère n'a pas d'action économique du tout, ni d'accompagnement, ni de politique d'achat. En revanche les collectivités oui. Donc derrière il y a quand même une logique de marché qui s'installe. Mais ce n'est pas le ministère qui est l'acteur.

– D'accord. Alors j'ai vu ensuite dans mes recherches qu'il y avait un « Portail GAR », qui permettait aux éditeurs d'avoir des indicateurs de suivi…

GP : Oui tout à fait. Il y a deux dispositifs dans le GAR, il y a un premier dispositif qui est du *back office* complet, qui est derrière les ENT, et qui fait que quand un élève clique sur l'icône d'une ressource, le GAR se met en branle, en *back office* – personne ne le voit, on ne sait même pas ce qui

se passe derrière – et il établit la liaison et l'accès. Ça c'est le premier fonctionnement. Et ensuite il y a une application web qu'on appelle le Portail GAR qui en fait permet à tous les gens autorisés – les éditeurs, tous les cadres de l'éducation au niveau national, académique, départemental, les chefs d'établissement, etc – de se connecter sur le Portail GAR et avoir un certain nombre d'informations sur les statistiques d'usage par exemple. Il y a des statistiques d'usage par exemple. C'est toujours très compliqué les statistiques parce qu'on en collecte des quantités puis après il y a des droits très fins pour savoir qui a droit à quoi. Mais bon globalement, il y a toute une politique sur les statistiques d'usage. Et puis il y a ensuite un certain nombre de choses permettant justement de savoir comment marche la ressource que je veux utiliser etc. De façon complémentaire, il y a aussi un troisième service qu'on appelle la « Console d'affectation ». C'est un outil web aussi qui permet dans l'établissement scolaire… mettons j'ai acheté 25 exemplaires d'un manuel au niveau 5e dans l'établissement scolaire, il faut que je puisse attribuer ces 25 exemplaires aux 25 élèves qui sont concernés. Et ça c'est le boulot de la Console d'affectation, c'est-à-dire que j'affecte la ressource aux élèves et aux profs qui en ont besoin.

– D'accord et là-dessus est ce qu'il y a une négociation pour savoir au cas par cas, par outil, qui aura accès, par exemple dans le cas d'un projet d'IA où il y a un laboratoire de recherche derrière ?

GP : Bien-sûr. En fait, l'entreprise a signé un contrat, ça commence par là. La première chose, c'est de signer le contrat GAR avec le ministère, et à partir du moment où ils ont signé un contrat et s'ils nous demandent, on leur donne les accès dont ils ont besoin. S'ils ont besoin de 50 accès, ils auront les 50 donc il n'y a pas… mais par contre dans un cadre d'utilisation bien précis, c'est à dire qu'il n'est pas question par exemple de prendre les statistiques et de les mettre sur la place publique pour en faire un argumentaire commercial parce que ce serait contraire aux règles avec les autres éditeurs. Par contre, ils ont tout ce qu'ils veulent faire très simplement. Sachant que la première fois que l'éditeur arrive et installe une ressource sur le GAR – on dit « accrocher » une ressource au GAR –, il est accompagné, c'est à dire qu'on fait tout un travail avec lui pour l'accompagner, pour vérifier la conformité etc. Donc il y a toute une démarche très accompagnée pour les premières ressources en particulier. Après une fois qu'ils ont fait les premières, les autres ça va tout seul.

– D'accord et donc l'idée des indicateurs de suivi sur ce portail, c'est aussi de permettre une « montée » en qualité ?

GP : Alors les indicateurs ne portent pas sur le fonctionnement interne de la ressource parce que sur le fonctionnement interne de la ressource ils appartiennent directement à l'éditeur, c'est lui qui les collecte, c'est lui qui les gère. Par contre, c'est sur les accès. Pourquoi tant de réserves pour avoir ces indicateurs sur les accès ? Eh bien pour une raison toute bête : c'est que quand une collectivité achète 1000 abonnements à une ressource, si elle se rend compte après que sur l'année il y a eu douze accès évidemment l'année suivante, elle ne va peut-être pas racheter les 1000 abonnements. Or, en particulier pour les manuels scolaires, on a une très vieille habitude, c'est qu'on achète des manuels scolaires qui ne sont pratiquement jamais utilisés. Ça fait partie de la tradition du manuel scolaire... Mais aujourd'hui les collectivités disent « Mais tu te rends compte, j'ai acheté 1000 exemplaires personne ne les a utilisés, ça fait que chaque clic nous coûte 800 euros ». Bon d'accord, mais sauf que quand il y avait le manuel papier, la page lue elle coûtait aussi 800 euros. Mais du coup c'est quelque chose qui devient critique. Les éditeurs, les grands éditeurs scolaires en particulier, essaient de camoufler ça, ils ne veulent pas que les statistiques sortent. Alors que les statistiques sont flagrantes. Moi je les ai vues évidemment, elles sont flagrantes. Aujourd'hui, il y a des ressources qui

ne sont pas des manuels et qui sont largement plus utilisées que les manuels. Il y a des ressources comme Pearltrees par exemple qui sont énormément utilisés, qui sont des outils de production en fait où l'élève peut produire des choses avec, l'enseignant aussi. Et dans les ressources aussi extrêmement utilisées, il y a Sesamath, très utilisé par les profs de maths et les manuels scolaires arrivent loin derrière. Mais bon, ce n'est pas politiquement correct de le dire (rire). Bon l'IA c'est trop jeune. Ce n'est pas significatif encore. Ça fait plus l'objet de politiques volontaristes qu'autre chose en fait, c'est… Les produits d'Evidence B par exemple, commencent à être beaucoup achetés par les collectivités, ils font ce qu'ils peuvent derrière pour promouvoir l'usage en accompagnant les établissements etc., mais ça ne démarre pas comme une fusée non plus.

– Et ces indicateurs de suivi sont un enjeu important…

GP : Ils sont précieux pour l'acheteur/collectivité, ils sont aussi précieux pour l'établissement parce que quelque part même si ce n'est pas lui qui paye, il est comptable des deniers publics sur l'utilisation des ressources. Et puis ensuite, dans un établissement, c'est intéressant d'avoir ces outils d'analyse pour savoir qui fait quoi de quelle manière. Et après ce n'est pas pour autant qu'on va aller dans le fonctionnement de la ressource puisque ça pour le coup, ça appartient vraiment à l'éditeur.

– Ok je vois.

GP : *(après un silence)* Alors l'intelligence artificielle, évidemment, c'est très intéressant. Il y a un truc quand même qui est, je trouve, important c'est la politique du ministère qui n'arrête pas de dire « Ok pour l'intelligence artificielle, ça a un potentiel considérable et cetera, mais attention, il faut que ce soit simplement un outil d'aide à la décision et il n'est pas question que la machine se substitue à l'enseignant pour décider ». Sauf que quand on regarde de près, dès qu'on fait de l'intelligence artificielle, moi utilisateur, enseignant, je ne sais pas démonter l'algorithme pour savoir pourquoi j'ai une recommandation qui arrive en termes d'usage. Donc ce processus d'aide à la décision, c'est une boîte noire en fait, c'est lié au principe même d'intelligence artificielle d'ailleurs que ce soit une boîte noire. Donc si c'est une recommandation qui est issue d'une boîte noire, comment ça peut m'aider à décider quoi que ce soit puisque je ne peux pas calculer ou apprécier son niveau de pertinence ? Dans la théorie de l'acteur-réseau, certains auteurs disent qu'il y a deux types d'artefacts : il y a ceux qui sont à compétence explicitées ou à compétences embarquées. Pour les compétences embarquées, les produits d'intelligence artificielle, aujourd'hui, ils embarquent des algorithmes qu'on ne peut pas démonter facilement. L'inverse, la compétence explicitée, c'est quand on explicite auprès de l'utilisateur pour qu'il ait une représentation opérationnelle de comment fonctionne la ressource. Sauf qu'en IA on ne sait pas faire ça, parce que les algorithmes sont évolutifs, les bases de connaissances sont évolutives etc., donc on ne sait pas expliquer pourquoi on produit une recommandation. Enfin on sait expliquer vaguement, mais on ne sait pas démonter les choses dans le détail. Donc cette histoire d'aide à la décision, je pense que c'est quand même une sacrée ligne de tension pour l'avenir parce que l'aide à la décision, quand on ne te donne pas les moyens de décider, ça veut dire quoi ? Je pense que c'est un vrai problème qui va un jour ou l'autre péter à la figure de pas mal de gens…

– De nombreux chercheurs, en tout cas informaticiens, se sont déjà posé la question de l'explicabilité mais c'est effectivement tout un domaine de recherche qui est assez brûlant…

GP : Ma soeur est chercheure en astrophysique et, évidemment, ils ont vraiment de l'intelligence artificielle. Et en gros le fait d'apporter une preuve à partir de réseaux de neurones, d'intelligence artificielle etc., quelle est la valeur de cette preuve ? Elle s'appuie sur la qualité supposée du système qui la produit mais elle n'est pas démontable. Typiquement un réseau neuronal ce n'est pas démontable. Donc le niveau de preuve n'est plus du tout le même que ce qu'était la preuve scientifique qu'on avait avant. Je pense que conceptuellement c'est intéressant quand même. Et du coup sur le côté aide à la décision pour les enseignants, je pense que potentiellement c'est quelque chose de très critique. Le ministère avait fait un appel d'offres IA il y a trois ans maintenant, et en fait on s'était rendu compte que dans toutes les ressources, il y a des biais dans l'analyse des données. En fait, je pense que dans l'intelligence artificielle, il y a toujours des biais de données dans tous les cas, ces biais de données on les découvre par exemple, on a découvert avec Evidence B des biais de données sur l'histoire masculin/féminin auquel personne ne s'attendait, et on les découvre un peu par hasard finalement. Il y en a peut-être qu'on ignore et qui peuvent influencer énormément les résultats. Du coup ce côté aide à la décision à mon avis, c'est quand même un point focal qui un jour ou l'autre va émerger et va poser des tas de problèmes.

– Qui n'a pas encore…

GP : Non, les utilisateurs ne se rendent pas compte de ça.

– Et du point de vue des ministères ?

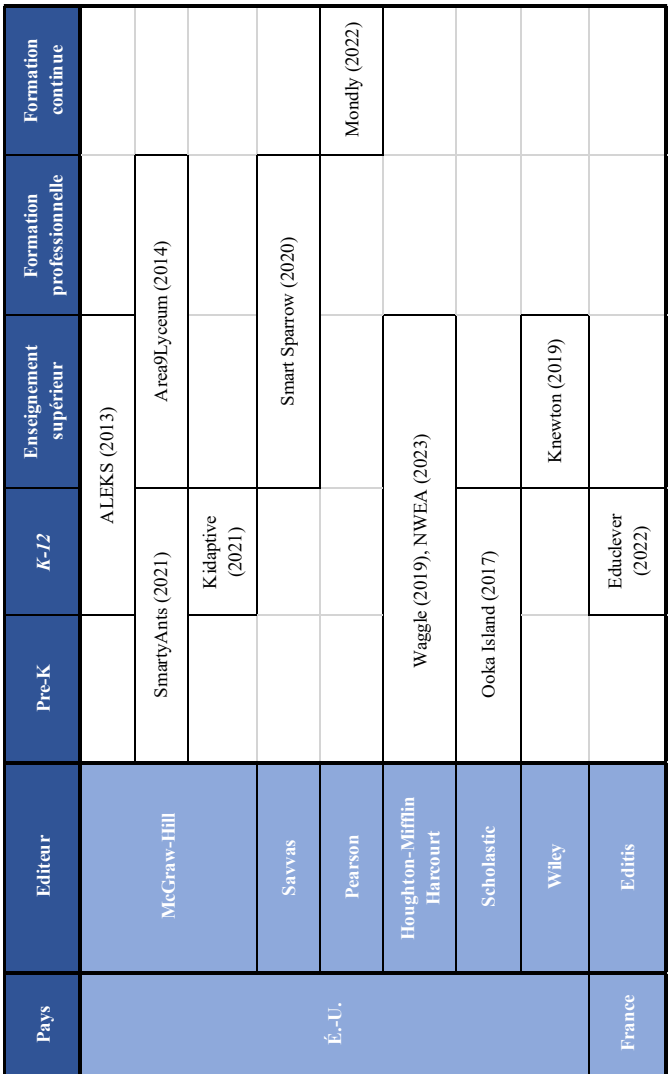
GP : Pas vraiment. Du point de vue du ministère, on est actuellement dans une phase de politique pro-IA à 100 %. Il y a le discours d'Emmanuel Macron, la commission qu'il a mise en place etc. On est pro-IA à un point… avec un niveau de réflexion qui est vraiment sommaire. Par exemple dans le dernier rapport de la Commission IA, en gros la CNIL leur pose des problèmes parce que la CNIL dit qu'il faut vérifier des tas de choses, et la proposition de la Commission, c'est de changer les missions de la CNIL. Soyons sérieux… ce n'est pas comme ça qu'on fera progresser l'IA de manière de manière pertinente. Mais bon voilà après c'est il y a des enjeux économiques, il y a tout ça donc… C'est toujours pareil : dans la théorie de l'acteur-réseau, on se dit qu'une technologie arrive à émerger parce qu'il y a des controverses qui se passent dans les réseaux technico-économiques, et que ces controverses, elles font avancer les choses jusqu'à ce qu'on arrive à trouver des alignements. Je pense que là il y a une controverse très sérieuse qui va s'établir sur ces questions-là.

ANNEXES

Annexe 1. Tableau récapitulatif des *AI Institutes* liés à l'éducation

Nom de l'*AI Institute*	National AI Institute for Student-AI Teaming (iSAT)	AI Institute for Engaged Learning (ENGAGE-AI Institute)	National AI Institute for Adult Learning and Online Education (AI-ALOE)	National AI Institute for Exceptional Education	National AI Institute for Inclusive Intelligent Technologies for Education (INVITE)
Financeur(s)	NSF	NSF	NSF, Accenture	NSF	NSF
Année de lancement	2020	2021	2022	2023	2023
Université fondatrice	University of Colorado	North Carolina State University	Georgia Institute of Technology	University at Buffalo	University of Illinois
Acteurs socio-économiques partenaires	**Huit universités** (Colorado State University; University of California, Santa Cruz, University of California, Berkeley; Brandeis University; Worcester Polytechnic Institute; Georgia Institute of Technology; University of Illinois at Urbana Champagne; University of Wisconsin-Madison), **deux districts** (Denver Public Schools and St. Vrain Valley Schools), **une agence de développement web** (Curve 10), **un pôle de compétitivité public-privé** (Pervasive Personalized Intelligence Center, et **une association** (Project VOYCE)	**Quatre universités** (North Carolina State University; Indiana University, University of North Carolina at Chapel Hill; Vanderbilt University), **une ONG** (Digital Promise), **écoles, musées**	**Six universités** (Harvard University, Vanderbilt University, Georgia Institute of Technology, Georgia State University, Technical College System of Georgia, UNC Queensboro), **trois entreprises** (Boeing, IBM, Wiley), et **une ONG** (1Edtech)	**Huit universités** (Cornell University, Penn State, Stanford University, University of Illinois Urbana-Champaign, University of Nevada, Reno, University of Oregon, University of Texas at El Paso, University of Washington).	**Huit universités** (University of Florida, Temple University, University of Oregon, University of Southern California, University of Michigan, Wright State University) et **une association** (ETS)
Thématique de recherche	Développer des "*AI partners*", systèmes d'assistance aux enseignants devant favoriser la travail collaboratif en petits groupes d'élèves	Développer des interfaces et des agents conversationnels exporant les valeurs pédagogiques de la narration, et approfondir le recours aux *Learning Analytics* pour les enseignants	Développer des assistants virtuels pour soutenir un apprentissage efficace des adultes et y intégrer des techniques de visualisation interactives pour aider les formateurs à intervenir.	Développer un système d'identification précoce des difficultés d'élocution et de langage des élèves, et un autre de recommandation auprès des orthophonistes en vue de mener des interventions personnalisées	Produire de nouvelles approches de modélisation de l'apprenant pour évaluer les facteurs cognitifs, émotionnels et sociaux ; approfondir l'aide au développement de compétences non-cognitives dans les environnements d'apprentissage STEM.
Niveau d'enseignement ciblé	Enseignement des sciences, technologies, ingénierie et mathématiques (STEM) aux niveaux primaire et secondaire	Enseignement formel ou informel des sciences, technologies, ingénierie et mathématiques (STEM) aux niveaux primaire et secondaire	Formation tout au long de la vie et formation professionnelle	*Kindergarten*/enseignement primaire	Enseignement des sciences, technologies, ingénierie et mathématiques (STEM) aux niveaux primaire et secondaire

Annexe 2. Tableau récapitulatif des acquisitions de start-ups par des groupes de l'édition scolaire et universitaire en France et aux États-Unis



Annexe 3. Tableau comparatif des deux GTnum liés à l'IA dans l'éducation

Nom du Gtnum	GTnum 8 – #IA_EO : "L'impact de l'IA à travers l'éducation ouverte".	Gtnum 9 – #Scol_IA : "Enseigner et apprendre à l'ère de l'intelligence artificielle"
Vague d'appel	2020-2022	2020-2023
Université (laboratoire) animateurs	Université de Nantes (Laboratoire des sciences du numérique de Nantes [LS2N]) et IMT-Atlantique Brest (Laboratoire des Sciences et Techniques de l'information de la Communication et de la Connaissance [LabSTICC])	Université Côte d'Azur (Laboratoire d'innovation et numérique pour l'éducation [LINE])
Objectif	"L'intelligence artificielle (IA) et les ressources éducatives libres (REL) sont deux phénomènes émergents depuis quelques années. Si chacun peut être étudié individuellement, observer les deux sujets ensemble permet de se poser de nouvelles questions : l'IA se nourrit de données ; est-il raisonnable, utile, de développer une politique ouverte en matière de contenus pédagogiques ? L'IA peut-elle aider à faciliter la navigation dans l'écosystème des REL ? à mieux trouver une ressource adaptée à un besoin particulier ? Inversement, peut-on proposer des REL pour apprendre l'IA ? et, plus particulièrement, pour aider les enseignants ?" (p.2)	"L'identification et le partage des pratiques et des usages créatifs du numérique pour une éducation adaptée aux enjeux des sociétés contemporaines liés à la diffusion des outils comportant des techniques d'IA" (p.2).
Disciplines associées	Informatique (3)*	Sciences de l'information et de la communication (3), Science de l'éducation et de la formation (3), Informatique (2)**
Acteurs socio-économiques associés	Académies de Rennes et Nantes	Académie de Nice, Inspé, Réseau Canopé, Teach on Mars (startup), Les petits Débrouillards (association)
	* Disciplines de rattachement des membres du comité de pilotage (Colin de la Higuera, Fahima Djelil et Serge Garlatti)	** Disciplines de rattachement des membres du comité de pilotage (Margarida Romero, Laurent Heiser, Maryna Rafalska) et des porteurs des 5 équipes de travail (Luis Galindo, Frédéric Alexandre, Magali Brunel, Solange Cartaut, Michel Durampart)

BIBLIOGRAPHIE

Introduction

Bachelard, G. (1938). *La Formation de l'esprit scientifique*, Paris, Vrin.

Bates, T. (2000). *Managing Technological Change. Strategies for College and University Leaders*. San Francisco, Jossey-Bass Publishers.

Bourdieu, P. (1987). *Choses dites.* Les éditions de Minuit, Collection Le sens commun.

Chaptal (2003). « Réflexions sur les technologies éducatives et les évolutions des usages : le dilemme constructiviste ». *Distances et savoirs*, volume 1, n°1, pp.121-147.

Descola P. (2008). « Sur Lévi-Strauss, le structuralisme et l'anthropologie de la nature », entretien avec Marcel Hénaff, *Philosophie* 98, juin 2008, pp. 8-36.

Ganascia, J. (2018). La révolution de l'Intelligence artificielle (IA) en autonomie. *Revue Défense Nationale*, 36-42. https://doi.org/10.3917/rdna.hs04.0036

Huet, A., Ion, J., Lefebvre A., Miège, B., Péron, R. (1978). *Capitalisme et industries culturelles*, Grenoble : Presses universitaires de Grenoble.

Katz, M. L. et Shapiro, C. (1985). « Network Externalities, Competition, and Compatibility ». *The American Economic Review*, 75(3), 424–440. <http://www.jstor.org/stable/1814809>

Kuhn, T. S. (2008/1962). *La structure des révolutions scientifiques*. Paris : Flammarion.

Jeanneret, Y. (2005). « Dispositif ». *In* Collectif (2005). *La 'Société de l'Information' : glossaire critique*. Paris, La Documentation Française : 50-51.

Maulini, O. & Perrenoud, P. (2005). La forme scolaire de l'éducation de base : tensions internes et évolutions. Dans : éd., *Les formes de l'éducation : variété et variations* (pp. 147-168). Louvain-la-Neuve : De Boeck Supérieur. https://doi.org/10.3917/dbu.mauli.2005.01.0147

Mœglin P. (1998) *L'industrialisation de la formation. État de la question.* Centre national de documentation pédagogique.

Mœglin, P. (2005). *Outils et médias éducatifs. Une approche communicationnelle*, Grenoble, Presses de l'Université de Grenoble.

Mœglin, P. (2010). *Les industries éducatives.* Paris. Presses Universitaires de France. Collection *Que sais-je ?*

Oppermann, R. (1994), « Adaptively supported adaptability », *Journal of Human-Computer Studies*, n° 40, p. 455-472

Peltier, C., Peraya, D., Bonfils P. et Heiser, L. (2022). « La forme universitaire à l'épreuve despratiques médiatiques personnelles », *Questions de communication*, 42 | 2022, mis en ligne le 01 février 2023, consulté le 03 mars 2023. <http://journals.openedition.org/questionsdecommunication/30184>

Petit, L. (2018). « Forme scolaire et évolutions paradigmatiques ». *Distances et Médiations des Savoirs*, Débat-discussion, 24, mis en ligne le 17 décembre 2018, consulté le 10 novembre 2024. URL : http://journals.openedition.org/dms/3233 ; DOI : https://doi.org/10.4000/dms.3233

Rawls, J. (1987/1971). *Théorie de la justice*. Traduction française de Catherine Audard, Paris, Seuil.

Rouvroy, A. (2013). « The end(s) of critique: data-behaviourism *vs.* Due process », in M. Hildebrandt, K. De Vries, eds, *Privacy, Due Process and the Computational Turn. Philosophers of Law Meet Philosophers of Technology*. Routledge.

Scholer, M. (1983). *La technologie de l'éducation. Concepts, bases et applications.* Montréal, Ministère de l'éducation du Québec/Presses de l'Université de Montréal, p.197.

Weber, M. (1963/1919). *Le Savant et le politique.* Paris, UGE, traduction française.

Young, M. D. (1994). *The Rise of the Meritocracy.* New Brunswick (USA) and London (UK) : Transaction Publishers.

Chapitre I

Anderson, J. R. (1987). Production systems, learning, and tutoring. In D. Klahr, P. Langley, & R. Neches (Eds.), *Self-Modifying Production Systems: Models of Learning and Development*. Bradford Books/MIT, Cambridge, MA.

Balacheff N. (1994) Didactique et intelligence artificielle. *Recherches en didactique des mathématiques*, 14 (1/2) 9-42

Balacheff, N. (1995). Conception, propriété du système sujet/milieu. Dans R. Noirfalise , M. J. Perrin-Glorian (dir.), Actes de la VII° Ecole d'été de didactique des mathématiques (p.215-229), Clermont-Ferrand : IREM de Clermont-Ferrand.

Baron, M., Gras, R. et Nicaud, J.-F. (1991). « Introduction », in Baron M., Nicaud J.-F. (eds.), *Deuxièmes journées EIAO de Cachan*, Editions de l'ENS de Cachan, pp.7-8

Berger, G. (1982). « Technologies et behaviorisme. Une rencontre essentielle et malencontreuse », in École Normale Supérieure de Saint-Cloud, Les Formes médiatisées de la communication éducative, actes du colloque du 9-10-11 novembre 1982, paris, pp.94-105.

Bloom, B. (1968). « Learning for Mastery ». Center for the Study of Evaluation of Instructional Programs, 1, 1-12.

Bloom, B. (1984). « The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring », *Educational Researcher*, 13: 6 (4-16).

Booms, A. (2024). « Norman A. Crowder et B. F. Skinner, frères ennemis de l'enseignement programmé ». *Albooms.fr*. Consulté le 2 août 2024. <https://albooms.fr/?p=1084>

Bonnet A., Kayser D., Cordier M.O. (1981). « An ICAI system for teaching derivatives in mathematics », *WCCE 1981*, Lausanne, juillet 1981.

Brown, J.S. et Burton, R.R. (1978). « Diagnostic Models of Procedural Skills in Basic Mathematical Skills ». Cognitive Science, 2, p. 155-192.

Bruillard, E. (1997). *Les machines à enseigner.* Paris, Editions Hermès.

Brusilovsky, P. (1996). « Methods and techniques of adaptive hypermedia ». *User Modeling and User Adapted Interaction*, 1996, v 6, n 2-3, pp 87-129

Brusilovsky, P., Schwarz, E., and Weber, G. (1996). « ELM-ART: An intelligent tutoring system on World Wide Web ». In: C. Frasson, G. Gauthier and A. Lesgold (eds.) Intelligent Tutoring Systems. Lecture Notes in Computer Science, Vol. 1086, (Proceedings of Third International Conference on Intelligent Tutoring Systems, ITS-96, Montreal, June 12-14,1996) Berlin: Springer Verlag, pp. 261-269

Burton, R.R et Brown, J.S. (1982). « An investigation of computer coaching for informal Learning activities » in Sleeman D., Brown J.S. (eds), *Intelligent Tutoring systems.* Academic Press, p. 157-183.

Carbonell, J. (1970). « AI in CAI : An Artificial Intelligence Approach to Computer-Assisted Instruction ». *IEEE transactions on man-machine systems.* Vol. MMS-11, no.4, Décembre 1970. <https://stacks.stanford.edu/file/druid:xr633ts6369/xr633ts6369.pdf>

Cardon, D., Cointet, J.-P. et Mazières, A. (2018). « La revanche des neurones : L'invention des machines inductives et la controverse de l'intelligence artificielle ». *Réseaux*, 211, 173-220. https://doi.org/10.3917/res.211.0173

Casas, M. (2002). « The use of Skinnerian teaching machines and programmed instruction in the United States 1960-1970 ». Harvard University, 1997, ERIC Document Reproduction Service No. ED 469942

Chaptal, A. (2003). *L'efficacité des technologies éducatives dans l'enseignement scolaire. Analyse critique des approches française et américaine*. Paris, Éd. L'Harmattan, coll. Savoir et formation.

Corbett, A. T., et Anderson, J. R. (1995). « Knowledge tracing: Modeling the acquisition of procedural knowledge ». *User Modeling & User-Adapted Interaction*, 4, pp. 253-278

Coulon, D. et Kayser D. (1975). « Aperçu sur les techniques d'éducation utilisant l'informatique ».

Revue française de pédagogie, volume 31, 1975. pp. 42-61. DOI : https://doi.org/10.3406/rfp.1975.1589.

Depover, C. & Quintin, J. (2011). Chapitre 1. Tutorat et modèles de formation à distance. Dans : Christian Depover éd., *Le tutorat en formation à distance* (pp. 15-27). Louvain-la-Neuve : De Boeck Supérieur. https://doi.org/10.3917/dbu.depov.2011.01.0015

Derry, S.J., Hawkes, L.W. (1993). « Local cognitive modeling of problem-solving behavior : an application of fuzzy theory » in Lajoie, S.P., Derry S.J. (eds), *Computers as cognitive tools,* LEA, Hillsdale, pp. 107-140

Dewey, J. (1968), *Expérience et éducation*, Paris, Armand Colin.

Dupuy, J.-P. (1994). *Aux origines des sciences cognitives.* Paris, La Découverte.

Elo, A (1978). *The Rating of Chessplayers, Past and Present*, New York, Arco Pub.

Etzwiler, D.D., Robb, J. R. (1972). « Evaluation of Programmed Education Among Juvenile Diabetics and Their Families ». *Diabetes*, Vol. 21, n°9. Septembre 1972.

Fleck, J. (1987), « Postscript: The Commercialisation of Artificial Intelligence », in B. Blomfield (ed.), *The Question of AI,* London, Croom-Helm, pp. 149-64.

Gardner, H. (1985), *The Mind's New Science*. *A History of Cognitive Revolution*. New York, Basic Books.

George, S. (2001). « Apprentissage collectif à distance. *SPLACH : un environnement informatique support d'une pédagogie de projet ».* Thèse de doctorat. Université du Maine, Lium.

Glikman, V. (2011). « Chapitre 7. Tuteur à distance : une fonction, un métier, une identité ? ». Dans : Christian Depover éd., *Le tutorat en formation à distance* (pp. 137-158). Louvain-la-Neuve : De Boeck Supérieur. https://doi.org/10.3917/dbu.depov.2011.01.0137

Goldstein, I., et Carr, B. (1977). « The computer as a coach: an athletic paradigm for intellectual education ». *Proc. of 1977 ACM annual conference*, Seattle, p. 227-233.

Goodman R. (1964). « L 'interprétation cybernétique des processus d'apprentissage ». *Hommes et Techniques,* 230, janvier 1964, p. 37-40.

Goody, J. (1986). *La logique de l'écriture. Aux origines de sociétés humaines,* Paris, Armand Colin, traduction française.

Héraud, J.-M. (2002). « PIXED : une approche collaborative de l'expérience et l'expertise pour guider l'adaptation des hypermédias en enseignement à distance », Doctorat de Université Lyon 1, Informatique.

Jones, M., Greer J., Mandinach, E, Du Boulay, B., Goodyear, P. (1992). « Synthesizing Instructional

and Computational Science ». In Jones M., Winne P. H. (eds.), *Adaptive Learning Environments*, NATO ASI Series, F85. Springer-Verlag, p. 383-401.

Keller, F. S. (1967). « Engineering personalized instruction in the classroom ». *Revista Interamericana de Psicologia/Interamerican Journal of Psychology*, 1(3)

Mauss, M. (1970). « Une catégorie de l'esprit humain : la notion de personne, celle de "moi" ». *Sociologie et Anthropologie, 1970*, 333-362.

Mœglin, P. (2005). *Outils et médias éducatifs. Une approche communicationnelle.* Grenoble, PUG.

Mœglin*,* P. (*2010*)*. Les Industries éducatives*. Paris, PUF, Que sais-je ?

Mœglin, P. (dir.). (2016). *Industrialiser l'éducation: anthologie commentée (1913-2012).* Saint-Denis, France : Presses universitaires de Vincennes.

Ohlsson, S. (1994). « Constraint-based student modeling ». *NATO ASI Series F Computer and Systems Sciences*, 125, 167–189.

Ohlsson, S. (1996). « Learning from performance errors ». *Psychological review*, 103, 241.

Oléron, P. (1965). « L'élève, le programme et la machine ». *L'éducation nationale, numéro spécial, L'enseignant et les machines,* n°15-16, p.39-41

Pirolli, F. et Crétin-Pirolli, R. (2011). « Web social et multimédia : propriétés d'une relation symbiotique ». *Les Enjeux de l'information et de la communication*, 12(2), 73-82. https://doi.org/10.3917/enic.hs03.0006

Ritter, S., Anderson, J.R., Koedinger, K.R*.,* Hadley, W., Corbett, A. T., et Lilly, M. (2004). « Classroom Integration of Intelligent Tutoring Systems for Algebra and Geometry ». Carnegie Learning.

Ritter, S., Anderson, J.R., Koedinger, K.R. et Corbett, A. T. (2007). « Cognitive Tutor: Applied research in mathematics education ». *Psychonomic Bulletin & Review* 14, 249–255. <https://doi.org/10.3758/BF03194060>

Saettler, P. (1968). *A History of Instructional Technology*. McGraw-Hill Book Company, 330 W.

Self, J. (1974) « Student models in Computer-Aided Instruction ». *International Journal of Man-Machine Studies*, 6, p. 261-276.

Sleeman, D. et Brown, J. S. (1982). *Intelligent Tutoring Systems*. New York: Academic Press.

Solomon, C. (1986). « Computer Environments for Children. A rejection on Theories of Learning and Education », *MIT Press,* 183 p.

Smallwood, R. (1962). « A Decision Structure for Computer-Based Teaching Machines ». *Computers*

and automation, Vol. XI, n°2, février 1962. <https://bitsavers.computerhistory.org/magazines/Computers_And_Automation/196202.pdf>

Suppes, P. (1966). « The Uses of Computers in Education ». *Scientific American,* Volume 215, n°3, septembre 1966.

Thorndike, E. L. (1912). *Education.* New York, MacMillan Company.

Watters, A. (2015). « Gordon Pask's Adaptive Teaching Machines ». *Hack Education*. Publié le 28 Mars 2015, consulté le 3 mai 2024. <http://hackeducation.com/2015/03/28/pask>

Wauters, K., Desmet, P., Van Den Noortgate, W. (2011). « Monitoring learners' proficiency: Weight adaptation in the Elo rating system ». In: Pechenizkiy, M., Calders, T., Conati, C., Ventura, S., Romero, C., Stamper, J. (eds) *Proc. of Educational Data Mining*, pp 247–252.

Weber, G., et Brusilovsky, P. (2001). « ELM-ART: An Adaptive Versatile System for Web-based Instruction ». *International Journal of Artificial Intelligence in Education*, 12, 351-384

Chapitre II

Bachimont, B. (2015). « Le numérique comme milieu : enjeux épistémologiques et phénoménologiques. Principes pour une science des données », *Interfaces numériques*. 4(3). https://doi.org/10.25965/interfaces-numeriques.386

Bahdanau, D., Cho, K. H., & Bengio, Y. (2015). « Neural machine translation by jointly learning to align and translate ». *3rd International Conference on Learning Representations,* ICLR 2015, San Diego, United States.

Baker, R. et Yacef K. (2009). « The State of Educational Data Mining in 2009: A Review and Future Visions ». *JEDM*, 1(1), 3-17.

Bonomo, V. (2017). « Brain-based learning theory ». *Journal of Education and Human Development,* 6(1), 27–43.

Compagnon, A. (2014). « Moocs et vaches à lait ». *Le Débat*, 180, 170-178. https://doi.org/10.3917/deba.180.0170

Commission Nationale de l'Informatique et des Libertés (2024). « Apprentissage profond (*deep learning*) ». *CNIL.* Consulté le 14 mai 2024. <https://www.cnil.fr/fr/definition/apprentissage-profond-deep-learning#:~:text=L%27apprentissage%20profond%20est%20un,donn%C3%A9es%20afin%20d%27%C3%AAtre%20entra%C3%AEn%C3%A9s>.

Connac, S. (2021). « Pour différencier : individualiser ou personnaliser ? », *Éducation et socialisation* 59 | 2021, mis en ligne le 31 mars 2021, consulté le 06 mai 2024. <http://journals.openedition.org/edso/13683> ; DOI : https://doi.org/10.4000/edso.13683

Daston, L. et Galison, P. (2012). *Objectivité*. Traduction de Sophie Renaut & Hélène Quiniou, préface

de Bruno Latour, Les Presses du réel, collection Fabula.

Delianidi, M., & Diamantaras, K. (2023). « KT-Bi-GRU: Student Performance Prediction with a Bi-Directional Recurrent Knowledge Tracing Neural Network ». *Journal of Educational Data Mining*, *15*(2), 1–21. https://doi.org/10.5281/zenodo.7808087

Gartner (2014). « Gartner Says Advanced Analytics Is a Top Business Priority ». *Gartner*. Dossier de presse. Publié le 21 octobre 2014. Consulté le 14 août 2024. <https://www.gartner.com/en/newsroom/press-releases/2014-10-21-gartner-says-advanced-analytics-is-a-top-business-priority>

Goldie, J. G. S. (2016) « Connectivism: a knowledge learning theory for the digital age? » *Medical Teacher*, 38(10), pp. 1064-1069. doi:10.3109/0142159X.2016.1173661

Klinkenberg, S., Straatemeier, M., Van der Maas, H. (2011). « Computer adaptive practice of maths ability using a new item response model for on the flyability and difficulty estimation ». *Computers & Education*, 57(2), pp. 1813–1824.

Knowles, J. E. (2015). Of Needles and Haystacks: Building an Accurate Statewide Dropout Early Warning System in Wisconsin. *Journal of Educational Data Mining*, 7(3), pp.18–67. https://doi.org/10.5281/zenodo.3554725

Labarthe, H. et Luengo, V. (2016). *L'analytique des apprentissages numériques*. Université de Paris 6, Laboratoire d'Informatique de Paris 6 (LIP6). Consulté le 10 mai 2024. <https://hal.archives-ouvertes.fr/hal-01714229/document>

Liu, Q., Shen, S., Huang, Z., Chen, E., et Andzheng, Y. (2021). « A survey of knowledge tracing ». *IEEE Transactions on Learning Technologies*. 17(2024), 1898-1919 https://doi.org/10.48550/arXiv.2105.15106.

Long, P. D. et Siemens G. (2011). « Penetrating the Fog: Analytics in Learning and Education ». *Educause Review,* 46(5), 30-40. Consulté le 10 mai 2024. <https://er.educause.edu/articles/2011/9/penetrating-the-fog-analytics-in-learning-and-education>

Mille, A. (2013). « L'interaction comme inscription de connaissance pour l'apprentissage humain ». Journée EIAH & IA 2013, May 2013, Toulouse, France. pp.1. hal-00824256

Piech, C., Bassen, J., Huang, H., Ganguli, S., Sahami, M., Guibas, L., Andsohl-Dickstein, J. (2015). « Deep knowledge tracing ». In Systems, C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett (eds.) « Advances in Neural Information Processing » Vol. 28. Curran Associates, Inc., 505–513.

Siemens, G. (2005). « Connectivism: A learning theory for the digital age ». *International Journal of Instructional Technology & Distance Learning*, 2, 3-10.

Tang, Y.Y. (2017). *Brain-based learning and education: Principles and practice.* New York: Academic Press.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et Polosukhin, I. (2017). « Attention is all you need ». *Advances in neural information processing systems*, *30*.

Chapitre III

Bender, E.M., Gebru, T., McMillan-Major, A. et Shmitchell, S. (2021). « On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? ». In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21).* Association for Computing Machinery, New York, NY, USA, 610–623. https://doi.org/10.1145/3442188.3445922

Brusilovsky, P. (2024). « AI in Education, Learner Control, and Human-AI Collaboration ». *Int J Artif Intell Educ* 34, 122–135 (2024). https://doi.org/10.1007/s40593-023-00356-z

Doroudi, S. et Brunskill, E. (2019). « Fairer but not fair enough on the equitability of knowledge tracing ». In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge,* pages 335–339, 2019

Hofmann, V., Kalluri, P.R., Jurafsky, D. et al. (2024) « AI generates covertly racist decisions about people based on their dialect ». *Nature,* 633, 147–154 (2024). https://doi.org/10.1038/s41586-024-07856-5

Hu, Q. et Rangwala, H. (2020). « Towards Fair Educational Data Mining: A Case Study on Detecting At-risk Students ». In: *Proceedings of The 13th International Conference on Educational Data Mining (EDM 2020)*, Anna N. Rafferty, Jacob Whitehill, Violetta Cavalli-Sforza, and Cristobal Romero (eds.) 2020, pp. 431 - 437

Jordan, M. I. (2019). « Artificial Intelligence — The Revolution Hasn't Happened Yet ». *Harvard Data Science Review*, 1(1). https://doi.org/10.1162/99608f92.f06c6e61

Kay, J. (2006). « Scrutable Adaptation : Because We Can and Must ». In V. P. Wade, H. Ashman, & B. Smyth (Éds.), *Adaptive Hypermedia and Adaptive Web-Based Systems* (p. 11-19). Springer. https://doi.org/10.1007/11768012_2

Le Cun, Y. (2019). *Quand la machine apprend*, Paris, Odile Jacob.

Ouakrat, A. et Mésangeau, J. (2016). « Resocialiser les traces d'activités numériques : une proposition qualitative pour les SIC ». *Revue française des Sciences de l'information et de la communication*, *8*. Consulté le 10 mai 2024. <https://journals.openedition.org/rfsic/1795> ; DOI : 10.4000/rfsic.1795

Peraya, D. et Luengo, V. (2019). « Les *Learning Analytics* vus par Vanda Luengo », *Distances et médiations des savoirs*, 27 | 2019, mis en ligne le 13 octobre 2019, consulté le 07 mai 2024. <http://journals.openedition.org/dms/4096> ; DOI : https://doi.org/10.4000/dms.4096

Rocher, L., Hendrickx, J.M. et de Montjoye, (2019). « Estimating the success of re-identifications in

incomplete datasets using generative models ». *Nat Commun* 10, 3069. https://doi.org/10.1038/s41467-019-10933-3

Schwartz, R., Dodge, J., Smith, N. A., et Etzioni, O. (2019). « Green AI ». *Communications of the ACM*, 63(12), 54-63.

TOP500 (2024). « Statistics of the TOP500 ». *TOP500*. Publié en juin 2024, consulté le 18 novembre 2024. <https://www.top500.org/statistics/list/>

Chapitre IV

Berger, G. (1982). « Technologie et behaviorisme, une rencontre essentielle et malencontreuse » in École normale supérieure de Saint-Cloud (1982). *Les formes médiatisées de la communication éducative*. Actes du colloque du 09 au 11 novembre 1982, pp. 94-105.

Boissière, J. et Bruillard, E. (2021). *L'école digitale : une éducation à construire et à vivre*, éditions Armand Colin, collection Sociologia, Malakoff.

Boltanski, L. et Chiapello, E. (1999). *Le nouvel esprit du capitalisme.* Paris, Gallimard.

Bouquillion Ph. (2012). « Concentration, financiarisation et relations entre les industries de la culture et industries de la communication », *Revue française des sciences de l'information et de la communication,* 1 | 2012, <http://journals.openedition.org/rfsic/94>, consulté le 2 octobre 2023.

Calviac S. (2019). « Le financement des universités : évolutions et enjeux » *Revue française d'administration publique*, (N° 169) (1), 51-68. https://doi.org/10.3917/rfap.169.0051.

Cauthen, L. (2021). « Just how large is the K-12 EdTech market? » *EdNews Daily*. Consulté le 3 octobre 2024. <http://www.ednewsdaily.com/just-how-large-is-the-k12-ed-tech-market/#:~:text=The%20U.S.%20K%2D12%20sector,of%20%247.5%20Billion%20over%202019.&text=Digital%20curriculum%20spend%20in%202020,of%20%242.1%20Billion%20over%202019>

Cengage Group (2023). « Cengage Group fourth quarter and fiscal year 2023 results ». *Cengage Group.* Consulté le 2 octobre 2024. <https://www.cengagegroup.com/news/press-releases/2023/cengage-group-fourth-quarter-and-fiscal-year-2023-results/>

Connac, S. (2021). « Pour différencier : individualiser ou personnaliser ? », *Éducation et socialisation*, 59 | 2021, mis en ligne le 31 mars 2021, consulté le 25 avril 2023. <http://journals.openedition.org/edso/13683> ; DOI : https://doi.org/10.4000/edso.13683

Coursera (2024). « Coursera Reports Fourth Quarter and Full Year 2023 Financial Results ». *Investor.coursera.com*. Publié le 1er février 2024, Consulté le 28 août 2024. <https://investor.coursera.com/news/news-details/2024/Coursera-Reports-Fourth-Quarter-and-Full-Year-2023-Financial-Results/default.aspx>

Curcic, D. (2023). « K-12 Textbooks Sales Statistics ». *Wordsrated*. Publié le 13 juin 2023, consulté le 14 juin 2024. https://wordsrated.com/k-12-textbooks-sales-statistics/

Deterding, S., Dixon, D., Khaled, R. et Nacke, N.E. (2011). « From game design elements to gamefulness: defining gamification », *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, ACM, pp. 9-15.

Direction des Études, de la Prospective, des Statistiques et de la Documentation (2024). « Chiffres clés. Statistiques de la Culture et de la Consommations ». *Culture.gouv*. Publié le 14 mars 2024, consulté le 2 octobre 2024. <https://www.culture.gouv.fr/espace-documentation/statistiques-ministerielles-de-la-culture2/publications/Collections-d-ouvrages/chiffres-cles-statistiques-de-la-culture-et-de-la-communication-2012-2023/Chiffres-Cles-2023-de-la-culture-et-de-la-communication>

Duolingo (2024). « *Form 10-K for the fiscal year ended December 31, 2023* ». *Investors.duolingo.com.* Publié le 29 février 2024, consulté le 28 août 2024. <https://investors.duolingo.com/static-files/06078a85-5de8-4c86-9a6a-7036259b07a4>

EdX (2022). « Impact Report. 10 Years 10 Mantras ». *Impact.edx.org.* Consulté le 28 août 2024. https://impact.edx.org/hubfs/impact-report-2022.pdf?hsCtaTracking=8366bf42-9765-4a38-8d85-2e7a0829016d%7C7fbba536-5150-4ee1-8e51-f11c889f44f2

Edutechnica. (2021). « *LMS data: Spring 2021 updates* ». *Edutechnica*. Publié le 21 juin 2021, consulté le 9 septembre 2024. <https://edutechnica.com/2021/06/21/lms-data-spring-2021-updates/>

Flichy, P. (1980). *Les Industries de l'imaginaire. Pour une analyse économique des médias*, PUG, Grenoble.

Freitag, M. (1995). *Le naufrage de l'université et autres essais d'épistémologie politique*, Québec et Paris, Nuit blanche éditeur et Éditions La Découverte.

Fourgous, J.-M. (2010). « Réussir l'école numérique ». Rapport de la mission parlementaire sur la modernisation de l'école par le numérique. Remis le 15 février 2010, consulté le 23 février 2021. <https://www.vie-publique.fr/files/rapport/pdf/104000080.pdf>

Hachette (2024). « Le groupe Hachette Livre en bref ». *Hachette.com*. Consulté le 28 août 2024. <https://www.hachette.com/le-groupe-hachette-livre-en-bref/>

Huet, A., Ion, J., Lefebvre, A., Miège, B. et Péron, R. (1978). *Capitalisme et industries culturelles*, Grenoble : Presses universitaires de Grenoble.

Houghton Mifflin Harcourt (2022). « HMH announces fourth quarter and full year 2021 results ». *HMH.* Consulté le 2 octobre 2024 <https://www.hmhco.com/about-us/press-releases/hmh-announces-fourth-quarter-and-full-year-2021-results>

Lacroix, J-G et Tremblay, G. (1991) *Télévision. Deuxième dynastie*. Montréal, Presses de l'Université du Québec.

Levoin, X. (2017). « Mutations des filières de l'édition scolaire et du numérique éducatif à la lumière

des BRNE ». Rapport interne STEF. Consulté le 28 août 2024. <https://eda.u-paris.fr /wp-content/uploads/sites/51/2024/04/LEVOIN-2017-Rapport_final_BRNE.pdf>

Miège, B., Pajon, P. (*1990*). « La *syntaxe des réseaux* ». In *Miège,* B, (1990, *dir.*)*, Médias et communication en Europe*. Grenoble, PUG

Mœglin, P. (1998). « Industrialisation, crise, réindustrialisation » in Mœglin, P. (1998, dir.). *L'industrialisation de la formation. Etat de la question*. Paris, Centre National de la Documentation Pédagogique.

Mœglin, P. (2005). *Outils et médias éducatifs. Une approche communicationnelle*, Grenoble, Presses de l'Université de Grenoble.

Mœglin, P. (2010). *Les industries éducatives.* Paris. Presses Universitaires de France. Collection *Que sais-je ?*

Porcher, L. (1998). « Préface », in Delamotte, E. (1998). *Une introduction à la pensée économique en éducation.* Paris, PUF, pp.1-3.

Scholastic (2024). « Scholastic reports fourth quarter and fiscal 2024 results ». *Investor.Scholastic.* Consulté le 2 octobre 2024. <https://investor.scholastic.com/news-releases/news-release-details/scholastic-reports-fourth-quarter-and-fiscal-2024-results>

Seveyrat, E. (2020). EdTech : Orthodidacte fusionne avec Educlever ». *Mesinfos.* Publié le 18 novembre 2020, consulté le 28 août 2024. https://mesinfos.fr/auvergne-rhone-alpes/edtech-orthodidacte-fusionne-avec-educlever-72100.html

Syndicat National de l'Édition (2024). « Les Chiffres de l'édition du Syndicat national de l'édition. Synthèse. 2023-2024 ». *SNE.fr.* Publié le 4 juillet 2024, consulté le 28 août 2024. https://www.sne.fr/document/synthese-des-chiffres-de-ledition-2023-2024/

Tremblay, G. (2007). « De Marshall McLuhan à Harold Innis ou du village global à l'empire mondial », *tic&société*, Vol. 1, n°1 | 2007, mis en ligne le 15 février 2008, consulté le 28 août 2024. URL : http://journals.openedition.org/ticetsociete/222 ; DOI : https://doi.org/10.4000/ticetsociete.222

Vulser, N. (2024). « Scor envisage de céder le groupe d'édition Humensis ». *Le Monde.fr.* Publié le 11 juillet 2023, consulté le 28 août 2024. https://www.lemonde.fr/economie/article/2023/07/11/scor-envisage-de-ceder-le-groupe-d-edition-humensis_6181539_3234.html

Watson, A. (2022). « *McGraw Hill Education revenue worldwide from 2013 to 2021* ». *Statista.* Consulté le 2 octobre 2024. <https://www.statista.com/statistics/800771/mcgraw-hill-education-revenue/>

Chapitre V

1EdTech (2021). « *Learning Impact 2021 Conference* ». *IMS Global Learning Consortium.* Consulté le 12 septembre 2024. <https://www.imsglobal.org/event/learning-impact-2021-conference>

1EdTech (2024). « OneRoster 1.1 introduction ». *IMS Global Learning Consortium.* Consulté le 12 septembre 2024. <https://www.imsglobal.org/oneroster-11-introduction>

Alfonseca, K. (2024). « Education policy: How Harris and Trump differ on K-12, higher education and more ». *ABC7.com.* Publié le 23 août 2024, consulté le 29 août 2024. https://abc7.com/post/education-policy-how-kamala-harris-donald-trump-differ-12-higher-education-more/15222990/

Amazon Web Service (2024a). « AWS au service de l'EdTech ». *AWS.Amazon.* Consulté le 8 avril 2024. < https://aws.amazon.com/fr/education/ed-tech/>

Amazon Web Service (2024b). « Etude de cas AWS : Knewton ». *AWS.Amazon.* Consulté le 8 avril 2024. <https://aws.amazon.com/fr/solutions/case-studies/knewton/>

Amazon Web Service (2024c). « L'IA permet à Duolingo de personnaliser l'apprentissage des langues ». *AWS.Amazon.* Consulté le 8 avril 2024. <https://aws.amazon.com/fr/machine-learning/customers/innovators/duolingo/>

Amazon Web Service (2024d). « AWS Education Accelerator ». *AWS Startups.* Consulté le 8 avril 2024. <https://aws.amazon.com/startups/accelerators/education-accelerator>

Amazon Web Service (2024e). « *Frais de service ». AWS. Consulté le 4 octobre 2024.* <https://docs.aws.amazon.com/fr_fr/marketplace/latest/userguide/listing-fees.html>

Area9 Lyceum (2024). « Automotive industry: Adaptive learning solutions ». *Area9Lyceum.* Consulté le 4 octobre 2024. <https://area9lyceum.com/experience-adaptive/industries/automotive/>

Beisser, S. (2008). *Unintended Consequences of No Child Left Behind Mandates on Gifted Students*. The Forum on Public Policy. Consulté le 21 janvier 2024. <https://files.eric.ed.gov/fulltext/EJ1099437.pdf>

Berger, L. et Stevenson, D (2007) « K-12 Entrepreneurship: Slow Entry, Distant Exit ». American Enterprise Institute Conference « The Supply Side of School Reform and the Future of Educational Entrepreneurship », 25 octobre 2007. Consulté le 10 février 2024. <https://d3btwko586hcvj.cloudfront.net/uploads/pdf/file/15/K12_Entrepreneurship__-1523562447.pdf>

Bosmeny, T. (2017). « ASU GSV SUMMIT: Clever ». *Youtube,* chaîne de *Global Silicon Valley.* Publié le 17 mai 2017, consulté le 9 septembre 2024. <https://www.youtube.com/watch?v=c3fPDWCbUxA&t=481s>

Berger, L. et Stevenson, D. (2007). « K-12 Entrepreneurship: Slow Entry, Distant Exit ». American Enterprise Institute Conference « The Supply Side of School Reform and the Future of Educational Entrepreneurship », 25 octobre 2007, consulté le 1er avril 2024. <https://d3btwko586hcvj.cloudfront.net/uploads/pdf/file/15/K12_Entrepreneurship__1523562447.pdf>

Bouchigny, S., Mégard, C., Knibbe, C., Bécu-Robinault, K., Gelas, D. et Alibert, F. (2016). « Une application du projet Tactileo, la paillasse virtuelle sur table interactive ». Revue *Technologie* n°204, septembre-octobre 2016. *Eduscol*. Consulté le 2 avril 2024. <https://eduscol.education.fr/sti/sites/eduscol.education.fr.sti/files/ressources/techniques/11270/11270-204-p16.pdf>

Bouquillion, P. (2012). « Concentration, financiarisation et relations entre les industries de la culture et industries de la communication », *Revue française des sciences de l'information et de la communication*, 1 | 2012, mis en ligne le 01 septembre 2012, consulté le 28 mars 2024. URL : http://journals.openedition.org/rfsic/94 ; DOI : https://doi.org/10.4000/rfsic.94.

Bradley, B. (2021). « K-12 dealmaking: Kahoot acquires Clever in deal combining game-based platform and single sign-on provider ». *EdWeek Market Brief.* Publié le 11 mai 2021, consulté le 10 septembre 2024. https://marketbrief.edweek.org/strategy-operations/k-12-dealmaking-kahoot-acquires-clever-in-deal-combining-game-based-platform-and-single-sign-on-provider/2021/05/

BrightEye Ventures (2024). « The European Edtech Funding Report 2024 ». *Brighteyevc.* Consulté le 27 mars 2024. <https://www.brighteyevc.com/post/the-european-edtech-funding-report-2024>

Casas, M. (2002). « *The Use of Skinnerian Teaching Machines and Programmed Instruction in the United States, 1960-1970* ». ERIC Institute of Education Sciences. Consulté le 10 février 2024. <https://files.eric.ed.gov/fulltext/ED469942.pdf>

Cavanagh, S. (2019). « Pearson sells its K-12 instructional materials business to investment firm ». *EdWeek Market Brief.* Publié le 19 février 2019, consulté le 27 septembre 2024. <https://marketbrief.edweek.org/strategy-operations/pearson-sells-its-k-12-instructional-materials-business-to-investment-firm/2019/02>

Chaptal, A. (2003). *L'efficacité des technologies éducatives dans l'enseignement scolaire. Analyse critique des approches française et américaine*. Paris, Éd. L'Harmattan, coll. Savoir et formation, 384 p.

Charle, C. et Verger, J. (2012). *Histoire des universités, XIIe-XXIe siècle*. Paris, PUF.

Crawford, R. (2022). « Action Research as Evidence-based Practice : Enhancing Explicit Teaching and Learning Through Critical Reflection and Collegial Peer Observation ». *Australian Journal of Teacher Education*, 47(12). https://doi.org/10.14221/1835-517X.6065

Cuban, L. (1993). « Computers Meet Classroom: Classroom Wins ». *Teachers College Record: The Voice of Scholarship in Education*, 95, 185 - 210.

Duolingo (2021). « Form S-1 registration statement under the securities act of 1933 ». *Investors.Duolingo.com.* Publié le 18 juin 2021, consulté le 13 septembre 2024. <https://investors.duolingo.com/node/6516/html>

Duolingo (2022). « *Form 10-K for the fiscal year ended December 31, 2021 ». Investors.Duolingo.com.* Publié le 4 mars 2022. Consulté le 13 septembre 2024. <https://investors.duolingo.com/static-files/06078a85-5de8-4c86-9a6a-7036259b07a4>

Duolingo (2023). « Introducing Duolingo Max, a learning experience powered by GPT-4 ». Publié le 14 mars 2023. Consulté le 30 novembre 2024. <https://blog.duolingo.com/duolingo-max/>

Duolingo (2024a). « *Form 10-K for the fiscal year ended December 31, 2023 ». Investors.Duolingo.com. Publié le 29 février 2024.* Consulté le 13 septembre 2024. <https://investors.duolingo.com/static-files/06078a85-5de8-4c86-9a6a-7036259b07a4>

Duolingo (2024b). « À propos de nous ». *Duolingo.* Consulté le 13 septembre 2024. <https://www.duolingo.com/info>

Edsurge (2016). *Decoding Adaptive*, Pearson, Open Ideas.

Émile-Besse, L. (2004). « Le système éducatif américain », *Revue internationale d'éducation de Sèvres*, 35 | avril 2004, mis en ligne le 23 novembre 2011, consulté le 30 janvier 2024. <http://journals.openedition.org/ries/1760>. DOI : https://doi.org/10.4000/ries.1760

Enhancing Education Through Technology Act (2015). « H.R.4021 – Enhancing Education Through Technology Act of 2015 ». *Congress.gov*. Consulté le 10 février 2024. https://www.congress.gov/bill/114th-congress/house-bill/4021/text

GlobeNewswire. (2005). « ClassLink completes $3 million equity round for expansion throughout the United States ». *GlobeNewswire*. Publié le 1er août 2005, consulté le 10 septembre 2024. <https://www.globenewswire.com/news-release/2005/08/01/331018/5500/en/ClassLink-Completes-3-Million-Equity-Round-For-Expansion-Throughout-the-United-States.html>

Google Cloud (2024). « *Walden University case study ». Google Cloud*. Consulté le 12 septembre 2024. <https://cloud.google.com/customers/walden>

Google Cloud (2024a). « Google Cloud pour l'Edtech ». *Cloud.Google.* Consulté le 8 avril 2024. <https://cloud.google.com/edu/edtech?hl=fr#section-2>

Google Cloud (2024b). « Schoolytics: Democratizing data science for teachers, parents, and students with BigQuery ». *Cloud.Google.* Consulté le 8 avril 2024. <https://cloud.google.com/customers/schoolytics>

Google Cloud (2024c). « Accélérez le développement de votre start-up avec le programme Cloud Google for Startups ». *Cloud.Google.* Consulté le 8 avril 2024. < https://cloud.google.com/startup?hl=fr>

Google for Education (2024a). « Education Plus ». *Google for Education*. Consulté le 10 septembre 2024. <https://edu.google.com/intl/ALL_en/workspace-for-education/editions/education-plus/>

Google for Education (2024b). « Teaching and learning upgrade ». *Google for Education*. Consulté le 10 septembre 2024. <https://edu.google.com/intl/ALL_en/workspace-for-education/editions/teaching-and-learning-upgrade/>

The Governor's Office of Student Achievement (2024). « Personalized Learning Grant ». *Gosa.georgia.gov.* Consulté le 3 septembre 2024. <https://gosa.georgia.gov/programs-initiatives/impact-previous-programs-initiatives/personalized-learning-grant>

Grand View Research (2024). « Higher education technology market report ». *GrandView Research.* Consulté le 3 octobre 2024. <https://www.grandviewresearch.com/industry-analysis/higher-education-technology-market-report>

Herold, B (2016). "Personalized Learning: What Does the Research Say?" Education Week, October 18, 2016. https://www.edweek.org/technology/personalized-learning-what-does-the-research-say/2016/10

Grignon, T. (2016). Quand Google fait école. Une prétention pédagogique en question. Communication & langages, 188, 123-139. https://doi.org/10.3917/comla.188.0123

Institute of Education Sciences (2024). *Program and Applicant Information*. IES.ED.gov. Consulté le 10 février 2024 https://ies.ed.gov/sbir/applicant.asp

Kantar Worldpanel (2024). « Smartphone OS market share ». *Kantar Worldpanel*. Consulté le 13 septembre 2024. <https://www.kantarworldpanel.com/fr/smartphone-os-market-share/>

International Society for Technology in Education (2024a). « ISTE 2024 call for exhibitors (CFE) ». *International Society for Technology in Education*. Consulté le 15 septembre 2024. https://conference.iste.org/2024/exhibitors/pdfs/ISTE2024_CFE.pdf

International Society for Technology in Education (2024b). « ISTE Standards ». *ISTE.* Consulté le 15 septembre 2024 <https://iste.org/standards>

International Society for Technology in Education. (2024c). « *ISTE standards for students* ». *ISTE*. Consulté le 15 septembre 2024 <https://iste.org/standards/student>

International Society for Technology in Education (2024d). « ISTE Standards for educators ». *ISTE*. Consulté le 15 septembre 2024. <https://iste.org/standards/educators>

International Society for Technology in Education (2024e). « ISTE Standards for education leaders ». *ISTE*. Consulté le 15 septembre 2024 <https://iste.org/standards/education-leaders>

International Society for Technology in Education (2024f). « ISTE Standards for coaches ». *ISTE*. Consulté le 15 septembre 2024 <https://iste.org/standards/coaches>

Institute of Education Sciences (2024). « Program and Applicant Information ». *IES.ED.gov.* Consulté le 10 février 2024. <https://ies.ed.gov/sbir/applicant.asp>

Joly, P.B. (2010). *On the economics of techno-scientific promises.* In M. Akrich, Y. Barthe, F. Muniesa & P. Mustar (eds), *Débordements. Mélanges offerts à Michel Callon* (pp. 203–222). Paris: Presse des Mines.

Labelle S. (2001). « La société de l'information », à décrypter ! ». *Communication et langages*, n°128, 2ème trimestre 2001. Dossier : Les T.I.C., objets de valeur. pp. 65-79.

Lacroix, J-G et Tremblay, G. (1991) *Télévision. Deuxième dynastie*. Montréal, Presses de l'Université du Québec.

Lalilo (2017). « Dossier de *Presse 2017* ». Communiqué de presse. *AWS.* Publié en octobre 2017, consulté le 13 septembre 2024. <https://s3-us-west-2.amazonaws.com/assets.lalilo.com /presse.pdf>

Lazare, M. (2021). « A look at what's ahead for Google Classroom ». *Google Blog.* Publié le 17 février 2021, consulté le 10 septembre 2024. <https://blog.google/outreach-initiatives/education/classroom-roadmap/>

Lemke, C. et al. (2004). *National Trends: Enhancing Education Through Technology. No Child Left Behind Title II D – Year One in Review.* Rapport produit par METIRI Group, commandé par la *State Educational Technology Directors Association*, Février 2004. Consulté le 10 février 2024. <https://files.eric.ed.gov/fulltext/ED537536.pdf>

Loeb, S. (2014). « *Edtech app platform Clever raises $30 million ». VatorNews.* Publié le 16 décembre 2014, consulté le 10 septembre 2024. <https://vator.tv/news/2014-12-16-edtech-app-platform-clever-raises-30-million>

McGraw Hill (2023). « *Intervention and acceleration: 2023 catalog ». McGraw Hill Education*. Consulté le 10 septembre 2024. <https://www.mheducation.com/unitas/school/ catalogs/2023/intervention-acceleration.pdf>

McKenzie, Lindsay (2019). « End of the Line for Much-Hyped Tech Company ». *Inside Higher Ed*. Publié le 6 mai 2019, consulté le 26 mars 2024. https://www.insidehighered.com/digital-learning/article/2019/05/07/wiley-buys-knewton-adaptive-learning-technology-company

Meuret, D. (2015). « États-Unis : quelle suite pour « No Child Left Behind »? », *Revue internationale d'éducation de Sèvres*, 69 | septembre 2015. Consulté le 12 février 2024. <http://journals.openedition.org/ries/4405>. DOI: https://doi.org/10.4000/ries.4405

Metinko, C. (2024). « Artificial Buildup: AI Startups Were Hot In 2023, But This Year May Be Slightly Different ». *Crunchbase News*. Publié le 9 janvier 2024, consulté le 7 octobre 2024. <https://news.crunchbase.com/ai/hot-startups-2023-openai-anthropic-forecast-2024/>

Microsoft Education Team (2022) « Check out highlights from Microsoft's global EdTech summit ». *Microsoft Education Blog*. Publié le 12 mai 2022, consulté le 10 avril 2024. <https://educationblog.microsoft.com/en-us/2022/05/check-out-highlights-from-microsofts-global-edtech-summit>

Microsoft Learn (2024a). « SaaS pricing models ». *Microsoft Learn*. Publié le 14 juin 2024, consulté le 12 septembre 2024. <https://learn.microsoft.com/fr-fr/partner-center/marketplace-offers/plan-saas-offer#saas-pricing-models

Microsoft Learn (2024b). « IA pour l'éducation, opportunités de formation et ressources ». *Centre Microsoft Learn pour les enseignants.* Consulté le 9 avril 2024. <https://learn.microsoft.com/fr-fr/training/educator-center/topics/ai-for-education>

Microsoft 365 Education (2024). « Microsoft 365 plan comparison table ». *Microsoft 365 Education*. Consulté le 12 septembre 2024. <https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/final/en-us/microsoft-product-and-services/microsoft-education/downloadables/EDU-Microsoft-365-Plan-Comparison-Table.pdf>

Midwest Comprehensive Center (2023). « Transforming Systems for High Levels of Learning for All Students : Personalized Learning in Wisconsin ». *Eric.ed.gov*. Publié en septembre 2018, consulté le 3 septembre 2024. https://eric.ed.gov/?id=ED589979

Mœglin, P. (2005). *Outils et médias éducatifs. Une approche communicationnelle*, Grenoble, Presses de l'Université de Grenoble.

Mœglin, P. (2010). *Les industries éducatives.* Paris. Presses Universitaires de France. Collection *Que sais-je ?*

Mœglin, P. (dir.). (2016). *Industrialiser l'éducation : anthologie commentée (1913-2012).* Saint-Denis, France : Presses universitaires de Vincennes.

National Center for Education Statistics (2021). « Revenues and Expenditures for Public Elementary and Secondary Education: FY 19 ». U.S. Department of Education, Institute of Education Sciences. Publié en juin 2021, consulté le 10 février 2024. < https://nces.ed.gov/pubs2021/2021302.pdf>

National Center for Education Statistics (2022a). « Table 214.10. Number of public school districts and public and private elementary and secondary schools: Selected school years, 1869-70 through 2021-22 ». *Digest of Education Statistics*. U.S. Department of Education, Institute of Education Sciences. Tableau réalisé en octobre 2022, consulté le 11 février 2024. <https://nces.ed.gov/programs/digest/d22/tables/dt22_214.10.asp>

National Center for Education Statistics (2023a). « *Postsecondary Institution Revenues* ». *Condition of Education*. U.S. Department of Education, Institute of Education Sciences. Mis à jour en août 2023, consulté le 8 février 2024 <https://nces.ed.gov/programs/coe/indicator/cud/postsecondary-institution-revenue>

National Center for Education Statistics (2023b). « *PISA 2022 U.S. Results* ». U.S. Department of

Education, Institute of Education Sciences. Publié en décembre 2023, consulté le 10 février 2024. <https://nces.ed.gov/surveys/pisa/pisa2022/>

National Center for Education Statistics (2023c). « *Revenues and Expenditures for Public Elementary and Secondary Education: FY 21* ». U.S. Department of Education, Institute of Education Sciences. Publié en mai 2023, consulté le 10 février <https://nces.ed.gov/pubs2023/2023301.pdf>

National Conference of State Legislatures (2022). « *Elementary and secondary school emergency relief fund tracker ».* Ncsl.org, publié le 25 janvier 2022, consulté le 7 février 2024. <https://www.ncsl.org/in-dc/standing-committees/education/elementary-and-secondary-school-emergency-relief-fund-tracker>

National Science Foundation (2004). « NSF 05-509: Science of Learning Centers (SLC) Program Solicitation ». *NSF.org.* Publié le 15 octobre 2004, consulté le 12 février 2024. <https://new.nsf.gov/funding/opportunities/science-learning-sl/5567/nsf05-509/solicitation>

National Security Commission on Artificial Intelligence (2021). « Part II: Winning the technology competition, Chapter 10: The Talent Competition ». In *Final Report* (pp. 171-182). Rapport d'étude. Publié en mars 2021, consulté le 10 février 2024. <https://assets.foleon.com/eu-central-1/de-uploads-7e3kk3/48187/nscai_ch10_digital_02-26-21.691636d1c7c7.pdf>

Neill, M., Guisbond, L. et Schaeffer, B. (2004). « Failing Our Children. How "No Child Left Behind" Undermines Quality and Equity in Education. An Accountability Model That Supports School Improvement ». *National Center for Fair & Open Testing*. Publié en mai 2004. Consulté le 23 janvier 2024. <https://eric.ed.gov/?id=ED529802>.

New York State Education Department (2020). « Building the Learning Environment: Educational Technology ». *Nysed.gov.* Mis à jour en décembre 2020, consulté le 9 février 2024. <https://www.nysed.gov/curriculum-instruction/building-learning-environment-educational-technology>

New Jersey Economic Development Authority (2007). « *Edison Innovation Fund investment assists Weehawken software and services provider ». NJEDA.* Publié le 24 avril 2007, consulté le 10 septembre 2024. <https://www.njeda.gov/edison-innovation-fund-investment-assists-weehawken-software-and-services-provider/>

Obama, B. (2015). « *Remarks by the President at Every Student Succeeds Act Signing Ceremony* ». Discours. The White House President Barack Obama. Publié le 10 décembre 2015. Consulté le 10 février 2024. <https://obamawhitehouse.archives.gov/the-press-office/2015/12/10/remarks-president-every-student-succeeds-act-signing-ceremony>

Pearson. (2024). « IMS Global Learning Consortium standards involvement ». *Pearson*. Consulté le 14 septembre 2024. https://accessibility.pearson.com/about/community-engagement/standards-involvement/ims-global-learning-consortium/index.php

Peter G. Peterson Foundation (2023). *«* How is K-12 Funded ? ». *Pgpf.org,* publié le 25 août 2023,

consulté le 10 février 2024. <https://www.pgpf.org/budget-basics/how-is-k-12-education-funded>

The Pew Charitable Trusts (2019). « “Lost Decade” Casts a Post-Recession Shadow on State Finances ». *Pewtrusts.org*. Publié le 4 juin 2019, consulté le 4 septembre 2024. <https://www.pewtrusts.org/en/research-and-analysis/issue-briefs/2019/06/lost-decade-casts-a-post-recession-shadow-on-state-finances>

Tools Competition. (2023). « Ed Tech innovators awarded more than $4 million for tools to aid 35 million learners ». *Tools-Competition.* Publié le 7 septembre 2023, consulté le 16 novembre 2024. <https://tools-competition.org/tools-competition-2022-23-press-release/>

Tools Competition. (2024). « Winners ». *Tools-Competition.* Consulté le 16 novembre 2024. https://tools-competition.org/winners/

Potts, J., Watzman, A., Zurawsky, C. et Lomando White P. (2004) « Carnegie Mellon, University of Pittsburgh Receive $25 Million NSF Grant To Create Center Focusing on the Science of Learning ». *Carnegie Mellon Today.* Publié le 5 octobre 2004, consulté le 10 février 2024. <https://www.cmu.edu/cmnews/extra/041005_nsf.html>

Raj Purohit, S. et Abowd, M. (2024). « How EdTech and workforce development companies drive business success through AWS Marketplace ». *Amazon Web Services*. Publié le 10 avril 2024, consulté le 4 octobre 2024. <https://aws.amazon.com/fr/blogs/publicsector/how-edtech-and-workforce-development-companies-drive-business-success-through-aws-marketplace/>

Rebillard, F. et Smyrnaios, N. (2010). « Les infomédiaires, au cœur de la filière de l'information en ligne Les cas de google, wikio et paperblog ». *Réseaux*, 2010/2 n° 160-161. pp. 163-194. https://doi.org/10.3917/res.160.0163.

Shemshack, A., Spector, J.M. (2020). « A systematic literature review of personalized learning terms ». *Smart Learning Environment*. 7, 33. Consulté le 10 février 2024. <https://doi.org/10.1186/s40561-020-00140-9>

Straight Talk on Evidence (2018). « An important – but fixable – flaw in the What Works Clearinghouse that can make ineffective programs appear effective (part three in a series) ». *Straighttalkonevidence.org*. Publié le 16 janvier 2018, consulté le 11 avril 2024. <https://www.straighttalkonevidence.org/2018/01/16/an-important-but-fixable-flaw-in-the-what-works-clearinghouse-that-can-make-ineffective-programs-appear-effective-part-three-in-a-series/>

Synergy Research Group (2024). « Cloud Market Gets its Mojo Back; AI Helps Push Q4 Increase in Cloud Spending to New Highs ». *Srgresearch*. Publié le 1er février 2024, consulté le 7 avril 2024. <https://www.srgresearch.com/articles/cloud-market-gets-its-mojo-back-q4-increase-in-cloud-spending-reaches-new-highs>

Texas Education Agency (2022). « State Approved Vendor List. Learning Acceleration Support Opportunities (LASO). Winter Cycle 2022-2023 ». *TEA.texas.gov*. Consulté le 4 septembre 2025. <https://tea.texas.gov/texas-schools/health-safety-discipline/covid/laso-state-approved-vendors-list.pdf>

Texas Education Agency (2024) « State-Adopted Instructional Materials ». TEA.org. consulté le 9 février 2024. < https://tea.texas.gov/academics/instructional-materials/state-adopted-instructional-materials>

Thuillas, O., Wiart, L. (2023). *Les plateformes à la conquête des industries culturelles*. Presses universitaires de Grenoble. https://doi.org/10.3917/pug.thuil.2023.01

Turner, F. (2012). *Aux sources de l'utopie numérique. De la contre-culture à la cyberculture, Stewart Brand un homme d'influence.* C&F Editions.

University System of Georgia (2024) « GeorgiaBEST ». usg.gov, consulté le 9 février 2024. https://www.usg.edu/georgia_best/

U.S. Department of Education (2023). « Applications for New Awards; Education Innovation and Research (EIR) Program-Early-Phase Grants ». *Federal Register*. Publié le 23 mai 2023, consulté le 27 mars 2024. <https://www.federalregister.gov/documents/2023/05/23/2023-10998/applications-for-new-awards-education-innovation-and-research-eir-program-early-phase-grants>

U.S. Department of Education (2024a). « About ED *». Ed.gov*, consulté le 10 février 2024. <https://www2.ed.gov/about/landing.jhtml>

U.S. Department of Education (2024b). « Race to the Top District (RTT-D) ». *Ed.gov*, consulté le 10 février 2024. <https://www2.ed.gov/programs/racetothetop-district/index.html>

U.S. Department of Education, Office of Educational Technology (2017a). « Reimagining the Role of Technology in Education: 2017 National Education Technology Plan Update ». Washington, D.C. Publié en janvier 2017, consulté le 10 février 2024. <https://tech.ed.gov/files/2017/01/NETP17.pdf>

U.S. Department of Education, Office of Educational Technology (2017b). « Reimagining the Role of Technology in Higher Education: A Supplement to the National Education Technology Plan ». Washington, D.C. Publié en janvier 2017, consulté le 10 février 2024. <https://tech.ed.gov/files/2017/01/Higher-Ed-NETP.pdf>

U.S. Department of Education, Office of Educational Technology (2023). « Artificial Intelligence and Future of Teaching and Learning: Insights and Recommendations *»*, Washington, DC. Publié le 1 mai 2023, consulté le 10 février 2024. <https://www2.ed.gov/documents/ai-report/ai-report.pdf>

U.S. Department of Education, Office of Educational Technology (2024). « National Educational Technology Plan *»*. Janvier 2024. Washington, DC. Publié le 22 janvier 2024, consulté le 10 février 2024. <https://tech.ed.gov/files/2024/01/NETP24.pdf>

Urban Institute (2022). *« Higher Education Expenditures* ». Urban.org. Données compilées le 21 octobre 2022. Consulté le 8 février 2024. <https://www.urban.org/policy-centers/cross-center-initiatives/state-and-local-finance-initiative/state-and-local-backgrounders/higher-education-expenditures#Question4Higher>

Utah State Board of Education (2023). « Utah's Personalized, Competency-Based Learning (PCBL) framework ». *Schools.utah.gov*. Publié en septembre 2023, consulté le 3 septembre 2024. <https://schools.utah.gov/curr/pcbl/_pcbl_/UtahPCBLFramework.pdf>

Utah STEM Action Center (2024). « *K-12 math personalized learning software grant* ». Stem.utah.gov. Consulté le 7 février 2024. <https://stem.utah.gov/educators/funding/k-12-math-personalized-learning-software-grant/>

Wan, T. (2017). « Pearson, an investor in Knewton, is phasing out partnership on adaptive products ». *EdSurge.* Publié le 12 mai 2017, consulté le 9 septembre 2024. <https://www.edsurge.com/news/2017-05-12-pearson-an-investor-in-knewton-is-phasing-out-partnership-on-adaptive-products>

Wan, T. (2019). « *Knewton was acquired for less than $17M, and former CEO Brian Kibby has left* ». *EdSurge.* Publié le 7 septembre 2019, consulté le 9 septembre 2024. https://www.edsurge.com/news/2019-09-07-knewton-was-acquired-for-less-than-17m-and-former-ceo-brian-kibby-has-left

Wan, T. (2024). « At $2.8B, 2023 U.S. Edtech Funding Is Back… to Pre-Pandemic Days. But in a Very Different World ». *Reachcapital*. Publié le 9 février 2024, consulté le 27 mars 2024. <https://www.reachcapital.com/2024/02/09/at-2-8b-2023-u-s-edtech-funding-is-back-to-pre-pandemic-days-but-in-a-very-different-world/>

Waters, T. (2012). *Schooling, Childhood, and Bureaucracy: Bureaucratizing the Child.* New York, Palgrave MacMillan.

Westervelt, E. (2015). « *Meet the mind-reading robo-tutor in the sky* ». *NPR*. Publié le 13 octobre 2015, consulté le 27 septembre 2024. <https://www.npr.org/sections/ed/2015/10/13/437265231/meet-the-mind-reading-robo-tutor-in-the-sky>

Weber, M. (1992). « L'objectivité de la connaissance dans les sciences et la politique sociales », in *Essais sur la théorie de la science* [1922], trad. franç., Paris, Plon, 1965, rééd. Presses-Pocket, « Agora ».

What Works Clearinghouse (2024). « Using the WWC to find ESSA tiers of evidence ». *IES. ED.gov.* Consulté le 10 avril 2024. <https://ies.ed.gov/ncee/wwc/Docs/multimedia/ESSA_508c_Infographic.pdf>

Yoder S. (2023). « School ed tech money mostly gets wasted. One state has a solution ». *The Hechinger Report*. Publié le 16 octobre 2023, consulté le 7 février 2023. <https://hechingerreport.org/school-ed-tech-money-mostly-gets-wasted-one-state-has-a-solution/>

Young, J. R. (2017) « Hitting reset: Knewton tries new strategy competing with textbook publishers ». *EdSurge*. Publié le 30 octobre 2017, consulté le 27 septembre 2024. <https://www.edsurge.com/news/2017-11-30-hitting-reset-knewton-tries-new-strategy-competing-with-textbook-publishers>

Chapitre VI

Académie d'Orléans-Tours. (2024). « Le recteur, chancelier des universités ». *Ac-Orléans-Tours*. Publié en juin 2024, consulté le 18 septembre 2024. <https://www.ac-orleans-tours.fr/le-recteur-chancelier-des-universites-121430>

Académie de Nancy-Metz. (2024). « Territoires numériques éducatifs ». *Ac-Nancy-Metz.* Publié en avril 2024, consulté le 18 septembre 2024. <https://www.ac-nancy-metz.fr/territoires-numeriques-educatifs-123804>

Adaptiv'Math (2021) « Notre équipe ». *Adaptivmath.* Consulté le 5 avril 2024. <https://www.adaptivmath.fr/team#nathan>

Agence Nationale de la Recherche (2020). « Appel à projets "Hybridation des formations de l'enseignement supérieur" ». *ANR.* Publié le 18 juin 2020, consulté le 11 avril 2024. <https://anr.fr/fileadmin/aap/2020/aap-ia-ncu3-2020-V2.pdf>

Agence Nationale de la Recherche (2021). « Appel à manifestation d'intérêt – Démonstrateurs numériques dans l'enseignement supérieur "DemoES" ». *ANR.* Publié le 9 mars 2021, consulté le 8 mars 2024. <https://anr.fr/fileadmin/aap/2021/ia-ami-demoes-2021-jo.pdf>

Afinef (2024a). « Présentation ». *Afinef.* Consulté le 30 septembre 2024. <https://afinef.net/a-propos/presentation/>

Afinef (2024b). « *Historique ». Afinef.* Consulté le 30 septembre 2024. <https://afinef.net/about-us/historique-2/>

Afinef (2024c). « Entreprises membres ». *Afinef.* Consulté le 30 septembre 2024. <https://afinef.net/a-propos/entreprises-membres/>

Apple. (2024). « Data & Marketing Association national client services terms ». *Apple.* Publié le 25 janvier 2024, consulté le 29 septembre 2024. <https://www.apple.com/legal/dma/dma-ncs.pdf>

Assemblée Nationale. (2023). « Question n°971 : Mise en œuvre des solutions numériques éducatives ». *Questions au gouvernement*. Publié le 15 novembre 2022, consulté le 29 septembre 2024. <https://questions.assemblee-nationale.fr/q16/16-971QE.htm>

Authemayou, C. (2017). « EdTech : Maskott joue la carte du tactile ». *L'Étudiant Educpros.* Publié le 13 février 2017, consulté le 2 avril 2024. <https://www.letudiant.fr/educpros/actualite/edtech-maskott-joue-la-carte-du-tactile.html>

Bazoge, M. (2024). « Les apps Android du Play Store s'ouvrent aux offres externes, mais Google veut une commission ». *01net.* Publié le 7 mars 2024, consulté le 30 septembre 2024. <https://www.01net.com/actualites/les-apps-android-du-play-store-souvrent-aux-offres-externes-mais-google-veut-une-commission.html>

Belkadi, N. (2022). « Prometheus-X défend une vision souveraine du numérique éducatif ». *EdTech Actu.* Publié le 20 janvier 2022, consulté le 1er décembre 2024. <https://edtechactu.com/outils-

collaboratifs/prometheus-x-defend-une-vision-souveraine-du-numerique-educatif/>

Bernigole V., Fernandez A., Loi M., Salles F. (2023). « PISA 2022 : la France ne fait pas exception à la baisse généralisée des performances en culture mathématique dans l'OCDE », *Note d'Information* n° 23.48, DEPP. https://doi.org/10.48464/ni-23-48

Béroud, F. (2024). « Le financement de la recherche et de l'enseignement supérieur par les collectivités territoriales entre 2020 et 2023 ». Note flash SIES. Avril 2024, n°7. Consulté le 18 septembre 2024. <https://www.enseignementsup-recherche.gouv.fr/sites/default/files/2024-04/nf-sies-2024-07-32703.pdf>

Berry G. (2008). *Pourquoi et comment le monde devient numérique*, Paris, Fayard, coll. « Collège de France ».

Bret A., Durand de Monestrol H., Hick M., Salles F., Fernandez A, Loi M. (2023). « PISA 2022 : culture scientifique, compréhension de l'écrit et vie de l'élève », *Note d'Information* n° 23.49, DEPP. https://doi.org/10.48464/ni-23-49

BrightEye Ventures (2024). « The European Edtech Funding Report 2024 ». *Brighteyevc.* Consulté le 27 mars 2024. <https://www.brighteyevc.com/post/the-european-edtech-funding-report-2024>

Caisse des Dépôts et Consignations (2015). « Territoires éducatifs d'innovation numérique. Espaces de formation, de recherche et d'animation numérique – e-FRAN ». Appel à projets permanent – Cahier des charges. *Association des Technologies de l'Information pour l'Education et la Formation*. Publié le 12 octobre 2015, consulté le 8 mars 2023. < http://www.atief.fr/sites/default/files/documents/AAP%20e-FRAN-VF.pdf>

Calviac, S. (2019). Le financement des universités : évolutions et enjeux. *Revue française d'administration publique*, 169, 51-68. https://doi.org/10.3917/rfap.169.0051

Camden, C. & Poncet, F. (2014). Recherche-action participative : nouvelles perspectives. Dans : Sylvie Tétreault éd., *Guide pratique de recherche en réadaptation* (pp. 383-422). Louvain-la-Neuve : De Boeck Supérieur. https://doi.org/10.3917/dbu.guill.2014.01.0383

Campus Matin (2023). « Service, accompagnement et évolution : les apports d'Amazon web service selon la edtech Domoscio ». *Campus Matin*. Publié le 8 mars 2022, consulté le 9 avril 2024. <https://www.campusmatin.com/numerique/edtechs/service-accompagnement-et-evolution-les-apports-d-amazon-web-service-selon-la-edtech-domoscio.html>

Conseil Scientifique de l'Éducation nationale (2019). « Présentation et activités 2019-2020 ». Réseau Canopé. Publié le 28 mai 2019, consulté le 15 mars 2024. <https://www.reseau-canope.fr/fileadmin/user_upload/Projets/conseil_scientifique_education_nationale/CSEN_Activites_2019-2020_WEB.pdf>

Commission Européenne (2018). « Communication de la Commission au Parlement européen, au Conseil, au Comité Économique et Social Européen Economic et au Comité des régions concernant le plan d'action en matière d'éducation numérique ». Publié le 17 janvier 2018, consulté le 10 avril

2024. <https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=CELEX:52018DC0022&from=FR>

Connac, S. (2021). « Pour différencier : individualiser ou personnaliser ? », Éducation et socialisation, 59 | 2021, mis en ligne le 31 mars 2021, consulté le 25 avril 2023. <http://journals.openedition.org/edso/13683> ; DOI : https://doi.org/10.4000/edso.13683

Cormaty, I. (2023). « Transformer des cours en contenus interactifs avec Nolej AI ». *Campus Matin.* Publié le jeudi 27 avril 2023, consulté le 3 avril 2024. <https://www.campusmatin.com/numerique/edtechs/transformer-des-cours-en-contenus-interactifs-avec-nolej-ai.html>

Cormaty, I. (2023). « Vers un même outil de visioconférence pour les établissements publics du supérieur ? ». *Campus Matin*. Publié le 29 juin 2023, consulté le 8 avril 2024. <https://www.campusmatin.com/numerique/equipements-systemes-informations/vers-un-meme-outil-de-visioconference-pour-les-etablissements-du-superieur-publics.html>

Cour des Comptes (2019). « Le service public numérique pour l'éducation. Un concept sans stratégie, un déploiement inachevé ». Rapport public thématique. *CComptes.* Publié en juillet 2019. Consulté le 5 mars 2024. <https://www.ccomptes.fr/sites/default/files/2023-10/20190708-rapport-service-public-numerique-education.pdf>

Clavey, M. (2024). « *MIA, la nouvelle IA de l'Éducation nationale, coûte 2,8 millions d'euros sur 3 ans ». Next INpact.* Publié le 10 avril 2024, consulté le 19 septembre 2024. https://next.ink/133921/mia-la-nouvelle-ia-de-leducation-nationale-coute-28-millions-deuros-sur-3-ans/

Commission Nationale de l'Informatique et des Libertés (2019). « *Les six grands principes du RGPD ». CNIL.* Publié le 23 août 2019, consulté le 28 septembre 2024. <https://www.cnil.fr/fr/comprendre-le-rgpd/les-six-grands-principes-du-rgpd>

Crédit Mutuel Arkéa. (2022). « Maskott renforce sa vocation d'entreprise EdTech ». *CM-Arkéa*. Communiqué de presse. Publié le 4 mai 2022, consulté le 25 septembre 2024. <https://www.cm-arkea.com/arkea/banque/upload/docs/application/pdf/2023-11/maskott-renforce-sa-vocation-dentreprise-edtech_cp_03052022.pdf>

CY Cergy Paris Université (2024) « *CY EdTech Labs »*. *CY Transfer*. Consulté le 24 septembre 2024. <https://cytransfer.cyu.fr/cy-edtech-labs>

Debes, F. (2024). « App Store, tiers, alternatives de paiement : Apple pas pressé d'ouvrir son écosystème ». *Les Echos.* Publié le 4 mars 2024, consulté le 29 septembre 2024. <https://www.lesechos.fr/tech-medias/hightech/app-store-tiers-alternatives-de-paiement-apple-pas-presse-douvrir-son-ecosysteme-2080091>

Direction des affaires juridiques (2014). « Le partenariat d'innovation ». *Economie.gouv*. Mis à jour le 29 septembre 2014, consulté le 15 mars 2024. <https://www.economie.gouv.fr/files/files/directions_services/daj/marches_publics/conseil_acheteurs/fiches-techniques/preparation-procedure/partenariat-innovation.pdf?v=1569591647>

Direction du numérique pour l'Éducation-TN2 (2020). « Groupes thématiques numériques. Appel à manifestation d'intérêt. Direction du numérique pour l'Éducation - 2020-2022 ». *Carnet de recherche Hypothèses « Éducation, numérique et recherche ».* Publié le 25 mai 2020, consulté le 7 mars 2024. <https://edunumrech.hypotheses.org/files/2020/07/AMI_GTnum_DNE_TN2_mai20_vf.pdf

Direction du numérique pour l'Éducation-TN2 (2024). Intelligence artificielle et éducation : Apports de la recherche et enjeux pour les politiques publiques (édition 2024) [Billet]. Éducation, numérique et recherche. <https://edunumrech.hypotheses.org/10764>

Domoscio (2024). « Action on Line ». *Domoscio*. Consulté le 30 mars 2024. <https://domoscio.com/fr/client/action-on-line-2/>

Les Éditeurs d'Éducation (2024). « Qui sommes-nous ? ». *Les Editeurs d'Education*. Consulté le 29 septembre 2024. <https://www.leseditеursdeducation.com/qui-sommes-nous/>

EdTech France. (2024a). « À propos ». *Edtech France*. Consulté le 29 septembre 2024. https://edtechfrance.fr/a-propos/

Educ@tech Expo. (2024). « À propos ». *Educatech Expo*. Consulté le 24 septembre 2024. <https://www.educatech-expo.com/salon>

Educlever (2014). « Vers la fin de l'échec scolaire ? Educlever présente « la Cartographie des Savoirs », première plateforme online d'enseignement personnalisé ». *MyMaxicours.fr*. Publié le 20 novembre 2014, consulté le 1er avril 2024. <https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://e.maxicours.com/P/kit-presse/2014_11_20_CP.pdf&ved=2ahUKEwjLjo3qq6GFAxWlcKQEHR_CCYkQFnoECBYQAQ&usg=AOvVaw2yfmcEA1Rhbb8IDgUZeT3E>

Eduscol (2023) « L'intelligence artificielle pour accompagner les apprentissages des fondamentaux au cycle 2 ». *Eduscol.education.* Mis à jour en décembre 2023, consulté le 1er avril 2024. <https://eduscol.education.fr/1911/l-intelligence-artificielle-pour-accompagner-les-apprentissages-des-fondamentaux-au-cycle-2>

Eduscol (2023a) « Délégués académiques au numérique (DAN) ». *Eduscol.education.* Mis à jour en novembre 2023, consulté le 15 février 2024. <https://eduscol.education.fr/1062/delegues-academiques-au-numerique-dan>

Eduscol (2023b). « Réseaux et interlocuteurs du premier degré pour le numérique ». *Eduscol.education.* Mis à jour en décembre 2023, consulté le 15 février 2024. <https://eduscol.education.fr/1057/reseaux-et-interlocuteurs-du-premier-degre-pour-le-numerique>

Eduscol (2024a). « Réseaux et interlocuteurs du second degré pour le numérique ». *Eduscol.education.* Mis à jour en février 2024, consulté le 15 février 2024. <https://eduscol.education.fr/1058/reseaux-et-interlocuteurs-du-second-degre-pour-le-numerique>

Eduscol (2024b). « Le dispositif Edu'up ». *Eduscol.education.* Mis à jour en février 2024, consulté

le 15 février 2024. <https://eduscol.education.fr/1603/le-dispositif-edu>

EM Lyon Business School (2024). « Ed.Job.Tech ». *Accelerator EM Lyon.* Consulté le 24 septembre 2024. <https://accelerator.em-lyon.com/edjobtech/>

Evidence B. (2024). « Evidence B lève 4 millions d'euros pour accélérer sa croissance à l'international ». *Evidence B*. Consulté le 27 septembre 2024. <https://evidenceb.fr/articles/evidenceb-leve-4-millions-d-euros-pour-accelerer-sa-croissance-a-l-international>

EY-Parthenon (2022). « La filière Edtech française. L'année du milliard, enjeux et perspectives d'une accélération inédite ». *EY*. Publié le 16 mars 2022, consulté le 26 mars 2024. <https://www.ey.com/fr_fr/strategy/la-filiere-edtech-francaise-l-annee-du-milliard>

Le Figaro Emploi (2023). « Maskott (43000) : siret, siren, TVA, bilan gratuit... ». *Entreprises.lefigaro.fr.* Publié le 11 décembre 2023, consulté le 1er avril 2024. <https://entreprises.lefigaro.fr/maskott-43/entreprise-479875718>

Fourgous, J.-M. (2010). « Réussir l'école numérique ». Rapport de la mission parlementaire sur la modernisation de l'école par le numérique. Remis le 15 février 2010, consulté le 23 février 2021. <https://www.vie-publique.fr/files/rapport/pdf/104000080.pdf>

Gani, D. (2024). « Démission au CSEN de trois chercheurs, des lanceurs d'alerte ? » *Le Café Pédagogique*. Publié le 4 mars 2024, consulté le 15 mars 2024. <https://www.cafepedagogique.net/2024/03/12/demission-au-csen-de-3-chercheurs-des-lanceurs-dalerte/>

Google Cloud (2024). « Accélérez le développement de votre start-up avec le programme Cloud Google for Startups ». *Google Cloud.* Consulté le 1er décembre 2024. <https://cloud.google.com/startup?hl=fr>

Gouvernement (2020). « Programme d'Investissements d'Avenir – Challenges éducation – Appel à projets – Vague 1 Lauréats des challenges – Cahier des charges ». Bpifrance. Consulté le 18 mars 2024. < https://www.bpifrance.fr/download/media-file/70038>

Gouvernement (2021). « L'école numérique ». Gouvernement.fr. Mis à jour le 3 décembre 2021, consulté le 16 février 2024. <https://www.gouvernement.fr/action/l-ecole-numerique>

GTnum LINE #Scol_IA (2023). « Enseigner et apprendre à l'ère de l'intelligence artificielle ». portfolio. Carnet de recherche Hypothèses « Éducation, numérique et recherche ». Publié le 10 juillet 2023, consulté le 8 mars 2024. <https://edunumrech.hypotheses.org/files/2023/07/GTnum_LINE_SCOL-IA_portfolio_juin2023.pdf >

GTnum LS2N #IA_EO (2023). « L'impact de l'IA à travers l'éducation ouverte ». Publié le 6 septembre 2023. Consulté le 8 mars 2024. <https://edunumrech.hypotheses.org/files/2023/09/GTnum_LS2N-IA_EO_portfolio_juin2023_vf.pdf>

Guerrier, P. (2024). « LMS : La reprise de Domoscio par Rise Up servira à l'apprentissage personnalisé ». *RH Matin.* Publié le 9 mai 2024, consulté le 7 octobre 2024. <https://www.rhmatin.com/formation/lms/lms-la-reprise-de-domoscio-par-rise-up-servira-a-l-apprentissage-personnalise.html>

Jean, A. (2022). « Intelligence artificielle & éducation, point d'étape en 2022 au cœur des transformations pédagogiques en cours ». Association française pour l'Intelligence artificielle, Journée « Enseignement et Formation en IA », le 6 janvier 2022. Présentation powerpoint. Consultée le 22 février 2024. <https://afia.asso.fr/wp-content/uploads/2022/01/IA-et-e%CC%81ducation_AFIA-TN2_janvier-2022v2.pdf>

Lalilo (2024). « *Politique de confidentialité – P2IA* ». *Lalilo.* Consulté le 24 septembre 2024. <https://ressources.lalilo.com/p2ia-confidentialite.pdf>

Langlois, G. (2021). « EvidenceB personnalise les exercices scolaires selon la manière d'apprendre des élèves ». *L'Usine Digitale.* Publié le 31 août 2021, consulté le 7 octobre 2024. <https://www.usine-digitale.fr/article/evidence-b-personnalise-les-exercices-scolaires-selon-la-maniere-d-apprendre-des-eleves.N1136134>

Livingstone-Wallace, S., Welter K. (2024). « Les technologies au service de l'éducation : une chance pour la France ». Livre blanc de l'Afinef, Edtech France, EducAzur, Edtech grand Ouest, Numeum. *Numeum*. Publié le 1er mars 2024, consulté le 27 mars 2024. <https://numeum.fr/actu-informatique/les-technologies-au-service-de-leducation-une-chance-pour-la-france>

Ludomag (2024). « Maskott rejoint la famille Cabri pour former un champion des STEAM ». *Ludomag.* Publié le 19 septembre 2024, consulté le 10 octobre 2024. <https://www.ludomag.com/2024/09/19/maskott-rejoint-la-famille-cabri-pour-former-un-champion-des-steam/>

MakeSense (2024). « *Passerelles* ». *Passerelles.makesense.* Consulté le 24 septembre 2024. <https://passerelles.makesense.org/>

Microsoft (2023). « L'adaptive learning change les codes de l'apprentissage ». *Microsoft Experiences*. Publié le 17 août 2023, consulté le 10 octobre 2024. <https://new-experience-2024.azurewebsites.net/articles/intelligence-artificielle/adaptive-learning/>

Miège, B. (1996). *La Société conquise par la communication. Tome I. Logiques sociales*. Presse Universitaires de Grenoble. Collection « Communication, médias et sociétés ».

Ministère de la Transformation et de la Fonction Publiques (2024). « DaSES ». *Entrepreneur(e)s d'intérêt général*. Consulté le 7 avril 2024. <https://www.eig.numerique.gouv.fr/defis/dases-1/>

Ministère de l'Éducation nationale et de la Jeunesse (2017). « Lancement du Partenariat d'Innovation – L'intelligence artificielle au service de l'apprentissage du français et des mathématiques ». Publié le 12 juillet 2017. Consulté le 12 mars 2024. <https://cache.media.eduscol.education.fr/file/Ia_et_education/13/6/Support_Presentation_PI_V5_796136.pdf>

Ministère de l'Éducation nationale et de la Jeunesse (2020). « Démontrer et tester le déploiement du numérique dans les territoires – Proposition n°03 ». Education.gouv. Consulté le 12 mars 2024. <https://www.education.gouv.fr/media/116386/download>

Ministère de l'Éducation nationale et de la Jeunesse (2021a). « E-FRAN : Des territoires éducatifs d'innovation numérique ». *Education.gouv.* Consulté le 18 septembre 2024. <https://www.education.gouv.fr/e-fran-des-territoires-educatifs-d-innovation-numerique-326083>

Ministère de l'Éducation nationale et de la Jeunesse (2021b). *«* Cahier des clauses techniques particulières ». *Next INpact*. Consulté le 19 avril 2024. <https://next.ink/wp-content/uploads/2024/04/MPA_21001_Service_Numerique_MathsFrancais2nd_CCTP_v2Next_pub.pdf>

Ministère de l'Éducation nationale et de la Jeunesse (2022). « 69 solutions numériques éducatives et 34 entreprises du numérique retenues par l'État pour les Territoires numériques éducatifs ». *Education.gouv.* Publié en juin 2022, consulté le 18 septembre 2024. <https://www.education.gouv.fr/69-solutions-numeriques-educatives-et-34-entreprises-du-numerique-retenues-par-l-etat-pour-les-341393>

Ministère de l'Éducation nationale et de la Jeunesse (2023a). *Numérique l'éducation 2023-2027 pour La vision stratégique d'une politique publique partagée*. Education.gouv. Publié en janvier 2023, consultée le 23 février 2024. <https://www.education.gouv.fr/media/120418/download>

Ministère de l'Éducation nationale et de la Jeunesse (2023b). « Choc des savoirs – Élever le niveau de notre école ». Dossier de presse. *Education.gouv*. Publié le 5 décembre 2023, consulté le 12 mars 2024. <https://www.education.gouv.fr/choc-des-savoirs-une-mobilisation-generale-pour-elever-le-niveau-de-notre-ecole-380226>

Ministère de l'Éducation nationale et de la Jeunesse (2023c). « Conférence presse : comment mettre en œuvre un choc des savoirs pour élever le niveau de l' École ? ». Vidéo YouTube, publiée le 5 décembre 2023, consultée le 29 février 2024. <https://www.youtube.com/watch?v=6fxx7m9pf5A>

Ministère de l'Éducation nationale et de la Jeunesse (2024a). « Les régions académiques, académies et services départementaux de l'Éducation nationale ». *Education.gouv.* Consulté le 18 septembre 2024. <https://www.education.gouv.fr/les-regions-academiques-academies-et-services-departementaux-de-l-education-nationale-6557>

Ministère de l'Éducation nationale et de la Jeunesse (2024b). « GAR – Gestionnaire d'Accès aux Ressources numériques ». *GAR.Education.* Page d'accueil. Consulté le 8 octobre 2024. <https://gar.education.fr/>

Mœglin, P. (2010). *Les industries éducatives.* Paris. Presses Universitaires de France. Collection *Que sais-je ?*

Mons, N. (2007). *Les nouvelles politiques éducatives: La France fait-elle les bons choix ?*. Presses Universitaires de France. https://doi.org/10.3917/puf.mons.2007.01

Monteil, J.-M. (2016). « L'appel à projets e-FRAN, un investissement d'avenir pour l'école de demain ». Dailymotion, vidéo publiée en 2016, consultée le 22 février 2024. <https://www.dailymotion.com/video/x38ngd2>

Navi (2024). « Hachette Éducation ». Navi. Consulté le 5 avril 2024. <https://navi.education/hachette-education/>

NEOMA Business School. (2024). « *NEOMA EdTech Accelerator ».* Consulté le 19 septembre 2024. <https://startuplab.neoma-bs.fr/fr/neoma-edtech-accelerator/>

Nolej (2024). « *Level up your campus with AI! ». Page d'accueil. Nolej. Consulté le 8 octobre 2024.* <https://nolej.io/>

Perreau, C. (2021). « Le scolaire, parent pauvre de l'Edtech française ». *Les Échos*. Publié le 21octobre 2021, consulté le 7 octobre 2024. <https://www.lesechos.fr/start-up/ecosysteme/le-scolaire-parent-pauvre-de-ledtech-francaise-1356976>

Radiofrance (2018). « Stanislas Dehaene : "Nous n'allons pas faire d'IRM à tous les écoliers" ». Radiofrance. Émission publiée le 11 janvier 2018, consultée le 29 février 2024. <https://www.radiofrance.fr/franceinter/podcasts/l-invite-de-8h20-le-grand-entretien/stanislas-dehaene-nous-n-allons-pas-faire-d-irm-a-tous-les-ecoliers-1766225>

Réseau Canopé. (2024a). « Présentation des groupes de travail du Conseil scientifique de l'Éducation nationale ». *Canopé.* Consulté le 19 septembre 2024. <https://www.reseau-canope.fr/conseil-scientifique-de-leducation-nationale-site-officiel/groupes-de-travail/presentation.html>

Réseau Canopé. (2024b). « Nos services – L'accélérateur d'innovations pédagogiques *». Canopé.* Consulté le 19 septembre 2024. <https://www.reseau-canope.fr/laccelerateur-dinnovations-pedagogiques/nos-services.html>

Romero, M., Heiser, L., Rafalska, M., Galindo, L., Durampart, M., Brunel, M., Cartaut, S., Alexandre, F. (2020). « GTnum Scol_IA : Pratiques numériques, formes scolaires, usages créatifs du numérique et IA. Publié le 18 septembre 2020, consulté le 7 mars 2024. <https://www.researchgate.net/publication/372314901_GTnum_Scol_IA_Pratiques_numeriques_formes_scolaires_usages_creatifs_du_numerique_et_IA>

Rosenblatt, B. (2019). « *Pearson's "digital first" strategy will change how students get textbooks ». Forbes.* Publié le 20 juillet 2019, consulté le 10 octobre 2024. <https://www.forbes.com/sites/billrosenblatt/2019/07/20/pearsons-digital-first-strategy-will-change-how-students-get-textbooks/>

Rosier, F. (2018) « Stanislas Dehaene, des neurosciences aux sciences de l'éducation ». *Le Monde.fr.* Publié le 09 janvier 2018, consulté le 29 février 2024. <https://www.lemonde.fr/sciences/article/2018/01/09/stanislas-dehaene-des-neurosciences-aux-sciences-de-l-education_5239132_1650684.html>

Secrétariat général du Gouvernement (2021). « Circulaire n°6282/SG : Doctrine d'utilisation de

l'informatique en nuage par l'État » Publié le 5 juillet 2021, consulté le 29 septembre 2024. <https://www.transformation.gouv.fr/files/ressource/Circulaire-n6282-SG-5072021-doctrine_utilisation-informatique-en-nuage-Etat.pdf>

Seveyrat, E. (2020). « EdTech : Orthodidacte fusionne avec Educlever ». *L'Essor Isère*. Publié le 18 novembre 2020, consulté le 1er avril 2024. <https://mesinfos.fr/auvergne-rhone-alpes/edtech-orthodidacte-fusionne-avec-educlever-72100.html>

Syndicat National de l'Édition (2024). « Les Chiffres de l'édition du Syndicat national de l'édition. Synthèse. 2023-2024 ». *SNE.fr.* Publié le 4 juillet 2024, consulté le 28 août 2024. <https://www.sne.fr/document/synthese-des-chiffres-de-ledition-2023-2024/>

Tactileo (2017). « L'Adaptive Learning (1/3) ». *Tactileo.com.* Publié le 2 janvier 2017, consulté le 1er avril 2024. <https://www.tactileo.com/formation-professionnelle/actus/ladaptive-learning-13/>

Unowhy (2024). « L'Accélérateur ». *Unowhy*. Consulté le 30 septembre 2024. < https://www.unowhy.com/en/accelerateur-en/>

Villani, C. (2018). *Donner un sens à l'intelligence artificielle. Pour une stratégie nationale et européenne*, Rapport de mission parlementaire.

Vitard, A. (2022). « *Google se conforme au DMA en ouvrant le système de paiement de son Store aux développeurs ». L'Usine Digitale.* Publié le 20 juillet 2022, consulté le 30 septembre 2024. <https://www.usine-digitale.fr/editorial/google-se-conforme-au-dma-en-ouvrant-le-systeme-de-paiement-de-son-store-aux-developpeurs.N2028042>

Chapitre VII

Mœglin, P. (dir.). (2016). *Industrialiser l'éducation : anthologie commentée (1913-2012).* Saint-Denis, France : Presses universitaires de Vincennes.

Chapitre VIII

Babbel (2021, 14 septembre). *Prospectus for the public offering*.

Baumol, W.J. et Bowen, W.G. (1966). *Performing Arts. The Economic Dilemma. A study of Problems common to Theater, Opera, Music and Dance.* New York, The Twentieth Century Fund

Bicknell, K., Brust, C., et Settles, B. (2023). « How Duolingo's AI learns what you need to learn ». *IEEE Spectrum.* Publié le 5 février 2023, consulté le 21 octobre 2024. <https://spectrum.ieee.org/duolingo>

Blanchard, T.C., Hayden, B.Y., et Bromberg-Martin, E.S. (2015). « Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity ». *Neuron* 85 (3), 602–614

Boët, H. (2023). « L'intelligence artificielle dans la stratégie de croissance de Duolingo : entre argument financier et levier d'industrialisation « par l'engagement » », *Distances et médiations des savoirs,* 44 | 2023, mis en ligne le 20 décembre 2023, consulté le 28 octobre 2024. URL : http://journals.openedition.org/dms/9571 ; DOI : https://doi.org/10.4000/dms.9571

Boët, H. (2024). « L'analytique des apprentissages numériques dans les produits d'alphabétisation pour enfants aux États-Unis : enjeux socio-économiques autour des « algorithmes de réquisition éducative ». *Actes du 23e congrès de la SFSIC*, 14-16 juin 2023.

Bouquillion Ph. (2012). « Concentration, financiarisation et relations entre les industries de la culture et industries de la communication », *Revue française des sciences de l'information et de la communication*, 1 | 2012, mis en ligne le 01 septembre 2012, consulté le 10 juillet 2023. <http://journals.openedition.org/rfsic/94>

Bonenfant, M. et Philippette, T. (2018). « Rhétorique de l'engagement ludique dans des dispositifs de ludification », *Sciences du jeu*, 10 | 2018, mis en ligne le 30 octobre 2018, consulté le 17 mai 2024. URL : http://journals.openedition.org/sdj/1422 ; DOI : https://doi.org/10.4000/sdj.1422

Brault-Labbé, A. et Dubé, L. (2012), « Mieux comprendre l'engagement psychologique : revue théorique et proposition d'un modèle intégratif », *Les Cahiers Internationaux de Psychologie Sociale*, n° 81, pp. 115131. DOI : 10.3917/cips.081.0115

Bullich, V. (2018) « La « plateformisation » de la formation », *Distances et médiations des savoirs,* 21 | 2018, mis en ligne le 07 mars 2018, consulté le 17 mai 2024. URL : http://journals.openedition.org/dms/2096 ; DOI : https://doi.org/10.4000/dms.2096

Carnegie Learning (2024a). « Why CL ». *Carnegielearning.com.* Consulté le 23 mai 2024. https://www.carnegielearning.com/why-cl/

Carnegie Learning. (2024b). « Our Approach ». *Help Center Carnegielearning.com.* Consulté le 23 mai 2024.<https://support.carnegielearning.com/help-center/math/home-connection/article/about-carnegie-learning/>

Challenges (2024). « Netflix concentré sur la fidélisation de ses 277 millions d'abonnés ». *Challenges.* Publié le 18 juillet 2024, consulté le 28 octobre 2024. https://www.challenges.fr/economie/netflix-continue-de-gagner-des-millions-d-abonnes_900050

Chaptal, A. (2003). *L'Efficacité des technologies éducatives dans l'enseignement secondaire. Analyse critique des approches française et américaine. Paris*, L'Harmattan.

Citton, Y. (2014). *L'économie de l'attention: Nouvel horizon du capitalisme ?* La Découverte. https://doi.org/10.3917/dec.citto.2014.01

Clément, B., Roy, D., Oudeyer, P.-Y., Lopes, M. (2015). « Multi-armed bandits for intelligent tutoring systems ». *J. Educ. Data Mining* 7 (2), 20–48

Conway, C. M., et Christiansen, M. H. (2006). « Statistical Learning Within and Between Modalities: Pitting Abstract Against Stimulus-Specific Representations ». *Psychological Science*, 17(10), 905-912. https://doi.org/10.1111/j.1467-9280.2006.01801.x

Cuban, L. (1993). « Computers Meet Classroom: Classroom Wins ». *Teachers College Record: The Voice of Scholarship in Education*, 95, 185 - 210.

Deci, E. L., et Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum.

Dehaene, S. (2018), *Apprendre ! Les talents du cerveau, le défi des machines*. Odile Jacob.

Domoscio. (2024) « L'avenir de l'apprentissage est adaptatif ». *Domoscio Blog*. Publié le 2 mai 2024. Consulté le 17 octobre 2024. https://domoscio.com/fr/blog/lavenir-de-lapprentissage-est-adaptatif/

Duolingo (2021). « Form S-1 registration statement under the securities act of 1933 ». <https://investors.duolingo.com/node/6516/html>

Duolingo (2024). « Confidentialité des données et Duolingo : bienvenue dans le Centre de données de Duolingo ». *Support.Duolingo.com.* consulté le 24 mai 2024. <https://support.duolingo.com/hc/fr/articles/360004256711-Confidentialit%C3%A9-des-donn%C3%A9es-et-Duolingo-bienvenue-dans-le-Centre-de-donn%C3%A9es-de-Duolingo>

EdSurge (2016). *Decoding Adaptive.* London: Pearson.

Fancsali, S.E., Ritter, S. (2020). « Data-Intensive Learning Engineering & Applied Education Research with Carnegie Learning's MATHia Platform ». In V. Rus & S.E. Fancsali (Eds.) *Proceedings of the The First Workshop of the Learner Data Institute (@EDM2020): Big Data, Research Challenges & Science Convergence in Educational Data Science.* Consulté le 20 octobre 2024. <https://drive.google.com/file/d/1Xb_vlVVPelt-2PghejOc9aPQQUdRTOI4/view>

Fancsali, S. E., Murphy, A., et Mathieu, K. (2022). « Carnegie Learning UpGrade /MATHia digital learning platform - Example: DataShop documentation ». *Carnegie Learning*. Publié le 12 septembre 2022, consulté le 20 octobre 2024. <https://seernet.org/wp-content/uploads/2022/10/Carnegie-Learning-UpGrade-MATHia-DLP-Example-DataShop-Documentation-202200912.pdf>

Freeman, C., Kittredge, A., Wilson, H. et Pajak, B. (2023). « The Duolingo Method for App-based Teaching and Learning. Duolingo Research Report ». Livre blanc. *Duolingo*. Publié le 11 janvier 2023, consulté le 1er octobre 2024. <https://duolingo-papers.s3.amazonaws.com/reports/duolingo-method-whitepaper.pdf>

Jeanneret, Y. (2014). *Critique de la trivialité. Les médiations de la communication, enjeu de pouvoir*. Paris, Non Standard.

Jeanneret, Y. et Souchier, E. (1999). « Pour une poétique de l'écrit d'écran », *Xoana*, 6, 1999, p. 97-107.

Joly, P.B. (2010). « On the economics of techno-scientific promises ». In M. Akrich, Y. Barthe, F. Muniesa et P. Mustar (eds), *Débordements. Mélanges offerts à Michel Callon* (pp. 203–222). Paris : Presse des Mines.

Kang, M.J., Hsu, M., Krajbich, I.M., Loewenstein, G., McClure, S.M., Wang, J.T., Camerer, C.F. (2009). « The wick in the candle of learning: epistemic curiosity activates re-ward circuitry and enhances memory ». *Psychol. Sci*. 20 (8), 963–973

Katz, M. L. et Shapiro, C. (1985). « Network Externalities, Competition, and Compatibility ». *The American Economic Review*, 75(3), 424–440. <http://www.jstor.org/stable/1814809>

Labelle, S. et Seurrat, A. (2015) « Pour une approche communicationnelle du serious games », Brougères, G. (dir.). *Penser le jeu*, Paris : Éditions du Nouveau Monde.

Lalilo (2024a). « Pédagogie ». *P2IA.Lalilo*. Consulté le 16 octobre 2024. <https://p2ia.lalilo.com/pedagogie>

Lalilo. (2024b). « Politique de confidentialité ». *Lalilo.* Consulté le 17 octobre 2024. <https://ressources.lalilo.com/p2ia-confidentialite.pdf>

Loderer, K., Pekrun, R., et Lester, J. C. (2020). « Beyond cold technology: A systematic review and meta-analysis on emotions in technology-based learning environments ». *Learning and Instruction*, 70, 101162. https://doi.org/10. 1016/j.learninstruc.2018.08.002

Malone, T.W., (1980). « What Makes Things Fun to Learn? A Study of Intrinsically Motivating Computer Games: Technical report ». Xerox Palo Alto Research Center, Palo Alto, CA.

Marion, P. (1997). « Narratologie médiatique et médiagénie des récits », *Recherches en Communication*, n° 7, p. 83

Miège, B., Bouquillion, P., Mœglin, P. (2013). *L'industrialisation des biens symboliques : Les industries créatives en regard des industries culturelles.* Presses universitaires de Grenoble.

Mœglin P. (1998). « L'industrialisation de la formation. État de la question ». Centre national de documentation pédagogique.

Mœglin, P. (2005). *Outils et médias éducatifs. Une approche communicationnelle*. Grenoble, PUG.

Mœglin, P., (2015). « Pour une économie politique de la création. De la trivialité à la créativité », *Communication & Langages*, n° 185, p. 60

Mœglin P. (dir.) (2016). *Industrialiser l'éducation : anthologie commentée (1913-2012)*. Saint-Denis, France, Presses universitaires de Vincennes.

Oudeyer, P. Y., Gottlieb, J., et Lopes, M. (2016). « Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies ». *Progress in brain research,* 229, 257–284. https://doi.org/10.1016/bs.pbr.2016.05.005

Pajak, B. et Bicknell, K. (2022). « At Duolingo, humans and AI work together to create a high-quality learning experience ». *Duolingo Blog*. Publié le 14 septembre 2022. Consulté le 22 octobre 2024. <https://blog.duolingo.com/how-duolingo-experts-work-with-ai/>

Pasquinelli, M. (2014). « Chapitre 9. Google PageRank : une machine de valorisation et d'exploitation de l'attention ». Traduction Blachier, C., Lockwood, V., Zuo, X. Dans Citton, Y. (dir.), *L'économie de l'attention. Nouvel horizon du capitalisme ?* (pp. 161-178). La Découverte. https://doi.org/10.3917/dec.citto.2014.01.0161.

Pidoux, J. (2019). « Toi et moi, une distance calculée. Les pratiques de quantification algorithmiques sur Tinder. » in Yann Calbérac, Olivier Lazzarotti, Jacques Lévy et Michel Lussault (dir.). *Carte d'identités. L'espace au singulier*, Paris, Hermann, 2019.

Plass, J. L., et Hovey, C. (2021). « The Emotional Design Principle in Multimedia Learning ». In R. E. Mayer et L. Fiorella (Eds.), The Cambridge Handbook of Multimedia Learning (3rd ed., pp. 324–336). Cambridge University Press. https://doi.org/10.1017/9781108894333.034

Rey, A. et al. (2016). *Dictionnaire historique de la langue française*, Paris, Le Robert.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., et Gulin, A. (2018). « CatBoost: unbiased boosting with categorical features ». *Advances in neural information processing systems*, 31.

Renaissance Lalilo (2024). « Implementation guide ». *Lalilo*. Publié en mars 2024, consulté le 26 octobre 2024. <https://resources.lalilo.com/Implementation_Guide_Lalilo.pdf>

Renaissance Learning (2021). « Research Foundation for Lalilo ». White paper. *Lalilo*. Paru en avril 2021, consulté le 23 mai 2024. <https://resources.lalilo.com/resources/Lalilo+_+Efficacy+Study+.pdf>

Riphagen, M., Hout, M.V., Krijnen, D., et Gootjes, G. (2013). « Learning tomorrow: visualising student and staff's daily activities and reflect on it ». ICERIE2013.

Ryan, R. M. et Deci, E. L. (2000), « Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being », The American Psychologist, vol. 55, n° 1, pp. 6878. DOI : 10.1037/0003-066X.55.1.68

Sauzéon, H., Clément, B., Mazon, C., Roy, D. et Oudeyer, P.-Y. (2021). « Conception d'un Système Tutoriel Intelligent (STI) basé sur les progrès d'apprentissage : des étapes formelles aux étapes d'expérimentations chez les enfants ». 10e Conférence sur les Environnements Informatiques pour l'Apprentissage Humain, Marie Lefevre, Christine Michel, Jun 2021, Fribourg / Virtual, Suisse. pp. 20-27. hal-03293081

Scholastic (2019). « Scholastic F.I.R.S.T. Research Foundation ». New York, NY: Scholastic Inc.

Seurrat, A. (2018). « Les médiations des savoirs pratiques sur les médias et la communication ». Habilitation à diriger les recherches. Volume 3 : Mémoire de synthèse. Université Paris Sorbonne.

Skinner, B. F. (1971). « Operant conditioning ». *The encyclopedia of education*, Vol. 7, pp. 29– 33. Macmillan

Souchier, E. (2012). « La « lettrure » à l'écran. Lire & écrire au regard des médias informatisés ». *Communication & langages,* N° 174(4), 85-108. https://doi.org/10.4074/S0336150012014068.

Sperber, D., et Wilson, D. (1989). *La Pertinence. Communication et cognition*. Éditions de Minuit.

Spotify. (2024). « L'engagement publicitaire vu par la neuroscience ». *Ads Spotify*. Consulté le 27 octobre 2024. <https://ads.spotify.com/fr-FR/news-et-insights/engagement-publicitaire-vu-par-la-neuroscience/>

Turner, F. (2012). *Aux sources de l'utopie numérique. De la contre-culture à la cyberculture, Stewart Brand un homme d'influence*. C&F Editions.

Wiley (2024a). « Knewton releases new insights into student usage, engagement and performance with Alta ». *Wiley Knewton Alta.* Publié le 15 janvier 2019, consulté le 16 octobre 2024. <https://www.wiley.com/en-us/education/alta/resources/blog-knewton-releases-new-insights-with-alta>

Wiley. (2024b). « Privacy policy ». *Wiley Knewton Alta*. Consulté le 17 octobre 2024. <https://www.wiley.com/en-us/privacy>

Wilson, K., et Nichols, Z. (2015). « The Knewton platform: A general-purpose adaptive learning infrastructure » Livre blanc. *Knewton*. Publié en janvier 2015.

Zosh, J. M., Hirsh-Pasek, K., Hopkins, E. J., Jensen, H., Liu, C., Neale, D., Solis, S. L., & Whitebread, D. (2018). « Accessing the Inaccessible: Redefining Play as a Spectrum ». *Frontiers in Psychology*, 9, 1124. https://doi.org/10.3389/fpsyg.2018. 01124

Chapitre IX

Académie de Clermont-Ferrand (2022). « Adaptiv'Math dans le GAR de l'ENT Beneylu Cantal ». *IEN Aurillac Les 3 Vallées*. Publié le 20 décembre 2022, consulté le 25 octobre 2024. <https://mediascol.ac-clermont.fr/ien-aurillac3/2022/12/20/adaptivmath-dans-le-gar-de-lent-beneylu-cantal/>

Adaptiv'Math. (2024a). « Playlist et gestion des groupes ». *Adaptiv'Math Blog*. Publié le 5 mars 2024, consulté le 26 octobre 2024. <http://www.adaptivmath.fr/blog/playlist-et-gestion-des-groupes>

Adaptiv'Math. (2024b). « C'est la rentrée : Découvrez les nouvelles fonctionnalités ». *Adaptiv'Math Blog*. Publié le 1er septembre 2024, consulté le 26 octobre 2024. <http://www.adaptivmath.fr/blog/c-est-la-rentree-decouvrez-les-nouvelles-fonctionnalites>

Bouchet, F. et Roy, D. (2021). « L'apport combiné de deux algorithmes d'IA à l'optimisation des parcours d'apprentissage dans le projet Adaptiv'Math ». *47ème colloque de la COPIRELEM*, Jun 2021, Grenoble, France. pp.855-868. hal-03625573

Carnegie Learning. (2024). « MATHia Adventure: A gamified math learning experience ». *Carnegie Learning.* Consulté le 26 octobre 2024. <https://www.carnegielearning.com/solutions/math/mathia-adventure/>

Conway, C. M. (2020). « How does the brain learn environmental structure? Ten core principles for understanding the neurocognitive mechanisms of statistical learning ». *Neuroscience & Biobehavioral Reviews*, 112, 279–299. https://doi.org/10.1016/j.neubiorev.2020.01.032

Domoscio (2024). « Domoscio Hub: Adaptive learning & analytics platform. *Domoscio Blog*. Consulté le 25 octobre 2024. <https://domoscio.com/en/products/domoscio-hub/>

Duolingo (2023). « Introducing Duolingo Max, a learning experience powered by GPT-4 ». *Duolingo Blog.* Publié le 14 mars 2023. Consulté le 30 novembre 2024. <https://blog.duolingo.com/duolingo-max/>

Freeman, C., Kittredge, A., Wilson, H. et Pajak, B. (2023). « The Duolingo Method for App-based Teaching and Learning. Duolingo Research Report ». Livre blanc. *Duolingo*. Publié le 11 janvier 2023, consulté le 1er octobre 2024. <https://duolingo-papers.s3.amazonaws.com/reports/duolingo-method-whitepaper.pdf>

Lyotard J.-F. (1979). *La Condition post-moderne*, Paris, Éditions de Minuit.

Ng, C. F., et Ng, P. K. (2015). « A Review of Intrinsic and Extrinsic Motivations of ESL Learners ». *International Journal of Languages, Literature and Linguistics,* 1(2), 98–105. https://doi.org/10.7763/IJLLL.2015.V1.20

Projet Voltaire. (2024). « L'Ancrage Mémoriel® La garantie d'une mémorisation efficace et durable ». *Projet Voltaire.* Consulté le 25 octobre 2024. <https://www.projet-voltaire.fr/presentation/woonoz/>

Protalinski, E. (2020, 18 août). « How Duolingo uses AI in every part of its app ». *Venture Beat.* Consulté le 11 décembre 2023. <https://venturebeat.com/ai/how-duolingo-uses-ai-in-every-part-of-its-app/>.

Sloan, K. (2023). « 5 ways game-based learning fosters student success ». *Carnegie Learning Blog*. Publié le 27 juin 2023, consulté le 26 octobre 2023. <https://www.carnegielearning.com/blog/student-success-in-gamified-classrooms/>