

# Groupe de travail Éducation, Recherche et Enseignement supérieur.

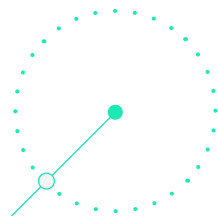
*Les données éducatives*

R  
E  
S  
E  
A  
R  
C  
H

sia-partners.com



@SiaPartners





# Sommaire

<b>5</b>	<b>En France, une ouverture des données inachevée</b>
5	Des données protégées sans cadre légal propre
5	<i>Un champ de définition très large</i>
6	<i>Une protection et un encadrement renforcés</i>
6	Une valorisation encore limitée
6	<i>Des ambitions freinées par un pilotage siloté</i>
7	<i>De premières avancées insuffisantes</i>
<b>8</b>	<b>À l'étranger, l'émergence de cas d'usages</b>
8	Un cadre éthique forgé par les pratiques
8	<i>Un cadre juridique faible</i>
8	<i>L'émergence de bonnes pratiques</i>
10	Le développement de cas d'usages
10	<i>Un service aux professeurs amélioré</i>
11	<i>Un outillage des prises de décisions</i>
11	<i>Un environnement pédagogique personnalisé</i>
<b>13</b>	<b>Les conditions à de nouveaux usages en France</b>
13	Un encadrement à deux échelles
13	<i>La nécessité d'un cadre légal précis</i>
13	<i>Une meilleure sensibilisation des personnels</i>
14	La création d'un Education Data Hub
14	<i>Un accès aisé, unifié, transparent et sécurisé</i>
15	<i>Plusieurs chantiers essentiels à mener</i>
16	Un appui sur les cas d'usages existants
16	<i>La généralisation des tableaux de bord</i>
17	<i>L'utilisation d'agents conversationnels</i>
17	<i>La mise en place de tableaux de gestion</i>
<b>19</b>	<b>Bibliographie</b>

# Préambule.

Le rapport sur la politique publique de la donnée du député Éric Bothorel, remis en décembre 2020 au Premier ministre<sup>(1)</sup>, a mis en lumière les défis de la France en termes d'ouverture et de partage des données. Aujourd'hui, à l'heure de la montée en charge du numérique et des formats hybrides alternant présentiel et distanciel, un autre domaine devrait faire l'objet d'une attention accrue : celui de l'éducation.

Le développement des services numériques dans le secteur de l'éducation modifie en profondeur ses processus administratifs (conception d'emplois du temps, gestion des absences, affectation des élèves) mais aussi les activités pédagogiques (accès à l'information, évaluation, individualisation des parcours d'apprentissage, organisation des enseignements).

Les données éducatives sont au cœur de ces évolutions. Alors que des craintes légitimes ont vu le jour concernant l'encadrement de ces données et les risques d'usages détournés (classement des élèves, etc.), leur analyse présente des opportunités dans le domaine de la pédagogie.

La présente étude tente d'éclaircir les enjeux liés à l'ouverture et au traitement des données recueillies dans le cadre scolaire, dans un contexte de développement du numérique éducatif et de l'exigence de protection des informations à caractère personnel. Dans une première partie, elle dresse un état des lieux de l'ouverture des données éducatives françaises en matière de collecte, de traitement et d'encadrement juridique. Elle tâche ensuite, à travers une analyse comparative, de rendre compte des avancées d'autres pays en la matière. Enfin, sur la base de ces éléments, elle propose des préconisations pour mettre ces données au service d'une amélioration de la qualité de l'enseignement et de la performance scolaire.

(1) Bothorel, Éric, Rapport au premier ministre, « Pour une politique publique de la donnée, des algorithmes et des codes sources », décembre 2020.

# En France, une ouverture des données inachevée.

*Des données protégées sans cadre légal propre.*

## Un champ de définition très large.

Les données éducatives renvoient à l'ensemble des données recueillies dans le cadre scolaire. Elles couvrent donc un cadre large : il peut notamment s'agir d'informations administratives sur les élèves, les enseignants et les parents ou encore de productions réalisées par les élèves ou leurs professeurs dans le cadre d'activités pédagogiques.

### Exemple

Effectifs d'étudiants inscrits dans les établissements et les formations de l'enseignement supérieur

Informations Tableau Export API

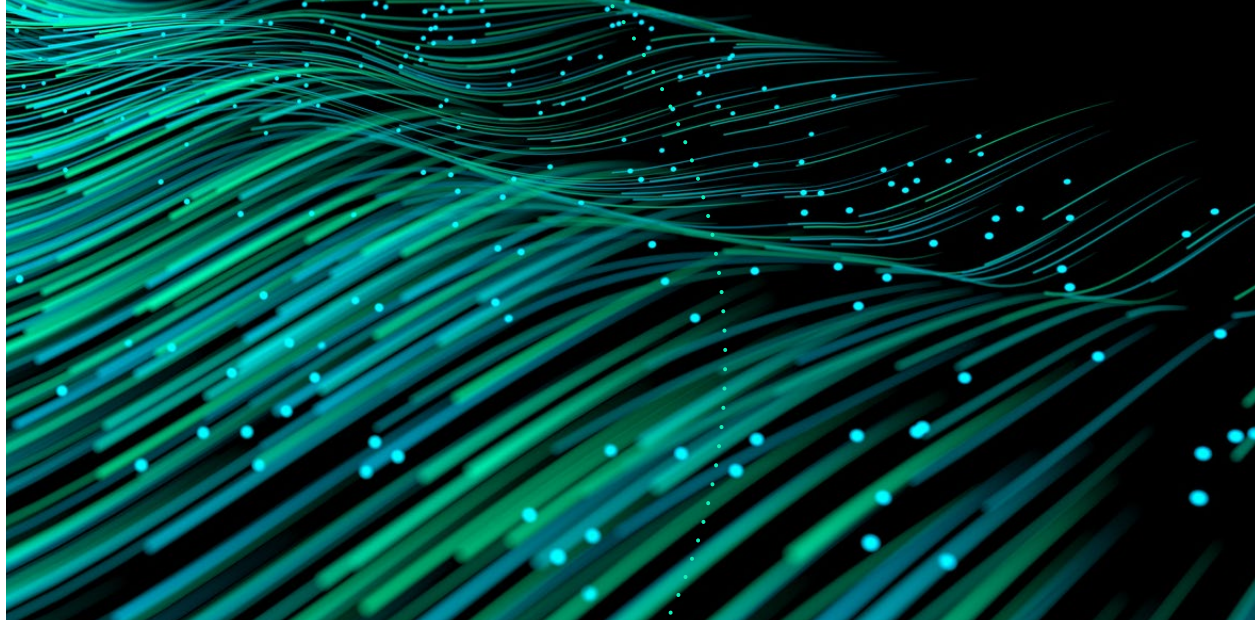
Rentrée universitaire	Niveau géographique	Unité géographique	regroupement	Regroupements de formations ou...	sec
2018-19	Commune	Hauteville-Lompnes	EC_PARAM	Écoles paramédicales et sociales	
2018-19	Commune	Misérieux	STS	Sections de techniciens supérieurs (STS)	
2018-19	Commune	Montluel	STS	Sections de techniciens supérieurs (STS)	
2018-19	Commune	Cyonnax	TOTAL	Total des formations d'enseignement su	
2018-19	Commune	Saint-Genis-Pouilly	EC_COM	Ecoles de commerce, gestion et compta	
2018-19	Commune	Bohain-en-Vermandois	STS	Sections de techniciens supérieurs (STS)	

Source : Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation

Avec la croissance de « l'école à distance », renforcée par la crise sanitaire, les établissements scolaires ont en effet recours à divers outils numériques (notamment les espaces numériques de travail, etc.). Or ces outils, dans leur fonctionnement même, génèrent, collectent et traitent des données scolaires. Ces données peuvent être relatives à un groupe (établissements, classes, générations de bacheliers, etc.) mais elles ont souvent un caractère personnel (noms, prénoms, coordonnées, résultats scolaires, données biométriques, identifiants de connexion, etc.).

Si les données scolaires ne sont pas considérées juridiquement comme « sensibles » (c'est-à-dire concernant l'origine, la religion, les opinions politiques, l'appartenance syndicale et la vie sexuelle<sup>(2)</sup>), elles peuvent toutefois le devenir après certains traitements :

(2) Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés (article 8).



● La direction des affaires juridiques du ministère de l'Éducation nationale a ainsi souligné, au sujet des signalements d'absences coïncidant avec des fêtes religieuses, que l'analyse de ces données peut indirectement informer sur la religion de la famille<sup>3</sup>.

● En éducation physique et sportive, une partie des données personnelles exploitées relève également de données sensibles (données de santé liées à la dispense de certaines activités, certificats de santé pouvant révéler certaines informations personnelles, données biométriques associées à des performances sportives, etc.).

Par ailleurs, l'augmentation des capacités de calcul et la masse des données pouvant être traitées grâce à de nouveaux algorithmes pourraient, par des croisements, permettre d'obtenir dans un futur proche des informations à caractère personnel à partir d'informations considérées initialement comme anonymes<sup>4</sup>.

### Une protection et un encadrement renforcés.

La collecte et l'utilisation des données éducatives est cadrée en France par le règlement général sur la protection des données (RGPD) et par les dernières dispositions de la loi du 6 janvier 1978 dite « Informatique et Libertés »<sup>5</sup>. Ces lois disposent notamment :

- L'obligation de déclarer auprès de la CNIL (Commission Nationale de l'Informatique et des Libertés) les fichiers contenant des données personnelles ;
- L'interdiction de collecter des données à caractère sensible (sauf exceptions) ;

● Le principe de collecte loyale de données (collecte directe, conservation limitée au minimum, etc.) ;

● L'obligation d'assurer la sécurité de l'ensemble des données collectées (pseudonymisation et chiffrement des données à caractère personnel, garantie de confidentialité, d'intégrité, de disponibilité et de résilience des systèmes et des services de traitement) ;

● L'obligation d'informer les individus concernés de la collecte de leurs données (identité du responsable du traitement, finalités du traitement, modalités d'exercice des droits), en des termes clairs, simples et facilement intelligibles ;

● Le droit à l'accès, la modification et la suppression des données en question.

La conformité à ce cadre légal est assurée dans chaque Académie par un délégué à la protection des données (DPD). Point de contact privilégié de la CNIL en interaction avec le DPD ministériel, ce dernier a la charge de veiller à l'application des règles en matière de traitement des données par les différents acteurs (responsables de traitements de données, établissements et éventuels prestataires ou sous-traitants). Il réalise notamment des audits pour identifier les éventuelles non-conformités à la loi, donne son avis sur la mise en œuvre d'un nouveau traitement ou la modification substantielle d'un traitement en cours et présente chaque année un rapport au rectorat<sup>6</sup>.

Cet encadrement, certes de plus en plus rigoureux, se heurte toutefois à l'absence de cadre juridique propre aux données éducatives.

### Une valorisation encore limitée.

#### Des ambitions freinées par un pilotage siloté.

Un objectif affiché du Ministère de l'Éducation nationale est de mieux valoriser les données éducatives, notamment pour :

● Personnaliser l'environnement d'apprentissage de l'élève (identification des forces et des besoins, accès à des parcours pédagogiques spécifiques, à des aides adaptées ou à des suggestions d'activités ou de ressources correspondant à ses marges de progression) ;

● Améliorer le service des professeurs en optimisant leur temps (moins d'activités chronophages) et en leur offrant une connaissance renforcée des besoins spécifiques de chacun des élèves ;

● Soutenir le travail des chercheurs en éducation et l'évolution des pratiques, en offrant une meilleure compréhension des interactions favorisant les apprentissages ;

● Mieux piloter le système éducatif, en utilisant les données à des fins statistiques pour évaluer les pratiques actuelles et prévoir leurs évolutions<sup>7</sup>.

Un comité d'expertise et un groupe de travail ont été mis en place au ministère pour travailler sur ces sujets, et plusieurs chantiers de recherches et d'expérimentations ont déjà été réalisés<sup>8</sup>.

Pour atteindre ces objectifs, plusieurs acteurs sont mobilisés :

● Au ministère de l'Éducation nationale, la Direction de l'évaluation, de

(3), (4) Inspection générale de l'Éducation nationale, Inspection générale de l'administration de l'Éducation nationale et de la Recherche, Rapport n°2018-016, « Données numériques à caractère personnel au sein de l'Éducation nationale », février 2018.

(5) Loi n°2018-493 du 20 juin 2018 relative à la protection des données personnelles.

(6) Ministère de l'Éducation nationale, de la Jeunesse et des Sports, Eduscol, « Délégués à la protection des données », octobre 2020.

(7), (8) Ministère de l'Éducation nationale, de la Jeunesse et des Sports, « Le numérique au service de l'école de la confiance », mars 2021.

la prospective et de la performance (DEPP) est en charge de concevoir et d'exploiter le SI statistique en matière d'enseignement, de publier des données statistiques et de mettre à disposition des acteurs des outils d'aide à l'évaluation, au pilotage et à la décision<sup>9</sup>.

- Au niveau national également, la Direction du numérique pour l'éducation (DNE) assure la mise en place et le déploiement du service public du numérique éducatif, notamment en coordonnant les actions en matière d'inventaire, de gouvernance, de production, de circulation et d'exploitation des données<sup>10</sup>.

- Dans chaque académie, une Délégation académique pour le numérique éducatif (DANE), pilotée par un délégué académique au numérique, est chargée de conseiller le Rectorat sur la conformité de la stratégie académique aux orientations nationales de développement et de formation aux usages du numérique. La DANE suit et anime également la mise en œuvre de cette stratégie, et en évalue les résultats<sup>11</sup>.

- De nombreux autres acteurs interviennent aux côtés de la DEPP et de la DANE (les DPD mentionnés plus haut mais également les inspections générales ainsi plusieurs directions du ministère, notamment la direction

des affaires juridiques et la direction du numérique de l'éducation), au point d'alourdir et de siloter le pilotage de cette politique publique.

### De premières avancées insuffisantes.

Le ministère de l'Éducation nationale s'est doté en 2016 d'une plateforme d'Open Data<sup>12</sup>, ce qui a permis d'ouvrir et de partager des données sur l'enseignement scolaire. Cette plateforme met à disposition 56 jeux de données (ensembles composant des bases d'informations) concernant notamment :

- Les effectifs (ex : nombre moyen d'élèves par classe par niveau de formation)
- Les résultats (ex : réussite au baccalauréat selon plusieurs critères)
- L'orientation des élèves (ex : vœux Parcoursup)
- La gestion des établissements (ex : données en équipements informatiques, temps scolaire)
- Les coûts et financements (ex : dépenses d'éducation par élève ou étudiant par niveau d'enseignement).

Malgré cette avancée, l'ouverture des données éducatives est encore limitée, pour plusieurs raisons :

- D'une part, les données disponibles sur la plateforme « data.education.gouv.fr » ne sont que partielles dans

la mesure où elles ne proviennent que du Ministère de l'Éducation nationale, alors que d'autres acteurs interviennent dans ce champ (les collectivités territoriales, certains opérateurs du ministère ainsi que des acteurs privés) ;

- D'autre part, les données publiées sont dispersées et peu lisibles (plus de deux tiers des jeux de données n'ont par exemple pas été actualisés depuis 2016), et la mise en œuvre de plateformes au niveau des régions académiques a été partielle.

Plus largement, deux problématiques majeures sont identifiées<sup>13</sup> :

- Premièrement, seuls les éditeurs privés et les chefs d'établissement ont aujourd'hui accès aux « traces numériques » (c'est-à-dire les données générées par les élèves et les professeurs lors de l'utilisation de services numériques), alors que ces services sont parfois financés par l'État et les collectivités territoriales ;

- Deuxièmement, l'ouverture des données éducatives pose de vraies questions liées aux craintes d'un usage dévoyé des données mises à disposition du public (ex : risque d'utilisation des données de résultats scolaires à des fins de classements des établissements sur le seul critère de réussite au baccalauréat).



À condition que leur protection soit rigoureuse, les données éducatives peuvent être utilisées pour individualiser les parcours de formation, améliorer l'évaluation des acquis et mettre à disposition des enseignants des outils pédagogiques plus efficaces. Dans le cas français, qui a fait l'objet de cette première partie, le constat est celui d'une ouverture et d'un management de la donnée inachevés, et d'un pilotage siloté.



(9) Ministère de l'Éducation nationale, de la Jeunesse et des Sports, « Direction de l'évaluation, de la prospective et de la performance », mars 2021.

(10) Ministère de l'Éducation nationale, de la Jeunesse et des Sports, « Direction du numérique pour l'éducation », avril 2021.

(11) Ministère de l'Éducation nationale, de la Jeunesse et des Sports, Eduscol, « Délégés académiques au numérique », mars 2021.

(12) Ministère de l'Éducation nationale, de la Jeunesse et des Sports, plateforme Open Data (data.education.gouv.fr).

(13) Digital New Deal (think tank), « Pour un Educational Data Hub au service de l'ouverture des données éducatives », avril 2021.

# À l'étranger, l'émergence de cas d'usages.

## Un cadre éthique forgé par les pratiques.

Bien que plusieurs lois relatives à la protection des données aient été adoptées dans différents pays, il n'existe pour le moment pas de socle juridique précis appliqué aux données éducatives. Afin de combler ce manque, une dynamique de construction de bonnes pratiques relatives à la collecte et à l'utilisation de ces données est constatée ces dernières années.

## Un cadre juridique faible.

À ce jour, on constate l'absence d'un accord global relatif à la collecte, l'utilisation et la protection des données. Alors que certains pays ont commencé à mettre en œuvre des accords liés aux données éducatives – les Etats-Unis notamment, avec le *Student Privacy Pledge* de 2014<sup>(14)</sup> – il n'existe à ce jour aucun socle juridique précis et défini concernant ces données.

En revanche, de nombreux Etats ont adopté des lois relatives à la protection des données : l'Union européenne est régie par le RGPD, le Canada par la loi PIPEDA, les Etats-Unis par le *Privacy Act* de 1974, etc. – pour ne citer que les

plus connus. Cependant la protection des données reste encore très inégale et ce, à plusieurs niveaux :

- Tout d'abord, tous les pays ne sont pas dotés de lois régissant la protection des données.

En effet, seuls 66 % en sont dotés (soit 128 sur 194 pays), divers projets de lois sont en cours dans 20 États. On comptabilise actuellement un peu moins de 40 pays (soit 19 %) où il n'existe aucune législation en matière de protection des données. Les 5 % restants correspondent à des pays avec une quasi-absence de données.<sup>15</sup>

- Le degré et le type de protection varient selon les pays et les territoires. Par exemple, l'Union européenne considère que peu de pays sont en adéquation (totale ou partielle) avec le RGPD en matière de transfert et de protection des données. Parmi eux, on retrouve notamment le Japon, la Nouvelle-Zélande, l'Argentine (adéquation totale), le Canada (adéquation partielle) mais pas les Etats-Unis ou l'Australie, où les transferts de données sont encadrés par des outils de transfert afin de garantir que les données soient protégées conformément à la réglementation européenne.<sup>16</sup>

Par ailleurs, il existe également des disparités au sein même des Etats-Unis, chaque Etat ayant décliné le *Privacy Act* de 1974 à sa manière. Alors qu'une loi fédérale est en cours d'élaboration au Congrès américain, certains Etats n'hésitent pas à se doter de leurs propres lois en matière de protection des données (la Californie par exemple, avec le *California Consumer Privacy Act* de 2018)<sup>17</sup>.

## L'émergence de bonnes pratiques.

Au fil des ans, différents groupes se sont formés afin de réfléchir à l'évolution des données éducatives, aussi bien sur ses aspects techniques qu'éthiques. Parmi eux, on retrouve notamment :

- La SoLAR (Society for Learning Analytics Research), réseau international interdisciplinaire de chercheurs spécialisés dans les données éducatives ;
- Le consortium LACE (Learning Analytics Community Exchange), financé par l'Union Européenne entre 2014 et 2016 dans le but de soutenir le développement d'une communauté européenne et de partager les bonnes pratiques émergentes<sup>18</sup> ;

(14) Engagement volontaire mais juridiquement contraignant, le Student Privacy Pledge est un engagement public destiné à protéger les données privées des élèves et étudiants. Il a été signé par 370 entreprises en 2019.

(15) CNUCED, « Online Consumer Protection Legislation Worldwide ».

(16) CNIL, « La protection des données dans le monde », novembre 2020.

(17) Mediavilla L., « Loi sur les données personnelles: la Californie ouvre le bal aux Etats-Unis », Les Echos, janvier 2020.

(18) Consortium LACE, « LACE LAW Manifesto », 2015.



- Le JISC (Joint Information Systems Committee), organisme à but non lucratif spécialisé dans le numérique éducatif au Royaume-Uni ;
- L'ICDE (International Council for Open and Distance Education), principale organisation mondiale qui œuvre pour offrir à tous une éducation accessible et de qualité grâce à l'apprentissage en ligne, ouvert et à distance.

Dans la ligne directrice de leur action, ces organismes ont publié différents rapports et travaux de recherches centrés sur la question éthique : *The LACE LAW Manifesto* (Consortium LACE, 2015), *Code of practice for learning analytics* (JISC, 2016)<sup>19</sup>, *Global Guidelines regarding ethically-informed practices in learning analytics* (ICDE, 2019)<sup>20</sup>, etc.

Bien que publiées par différents organismes, toutes ces publications convergent quant à la nécessité de mettre en place de bonnes pratiques sur l'utilisation et la protection des données éducatives au niveau international.

Plus concrètement, voici les 10 champs d'action identifiés par l'ICDE afin d'utiliser les données éducatives dans un cadre éthique :

- Une plus grande transparence concernant la collecte des données éducatives, leur analyse et leur impact sur l'apprentissage.
  - La mise en place d'une transparence institutionnelle, qui communique sur les objectifs recherchés par l'étude et la collecte des données éducatives ;
  - En termes pratiques, cela consiste à informer les élèves et étudiants sur les processus de collecte et d'analyse de ces données, pour leur permettre de comprendre les objectifs de cette démarche et l'impact que cela aura sur leur apprentissage (i.e. quelles données sont collectées ? Comment sont-elles analysées ? Dans quel but ?).
- Les questions relatives à la propriété des données éducatives (*Data ownership and control*).
  - Bien que certaines données soient détenues par des institutions

publiques, celles-ci ne sont pas pour autant propriétaires de ces données ;

- Une autorisation temporaire d'utiliser et d'analyser ces données pourrait être envisagée afin de permettre l'avancée des chercheurs, à condition qu'un cadre juridique précis soit mis en place.
- L'accessibilité des données éducatives : l'accès aux données est à différencier selon les utilisateurs.
  - Instaurer différents niveaux de permissions pour accéder aux données ;
  - Permettre aux élèves et étudiants d'accéder à leurs données en cas de besoin.
- La validité et la fiabilité des données éducatives : les institutions devront s'assurer que les données sont correctes et fiables avant de procéder à leur analyse.
  - Dans le cadre d'analyses prédictives ou de calculs statistiques, une attention particulière devra être portée pour s'assurer que les bases de données sont complètes et suffisamment détaillées afin de permettre ces opérations ;
  - En outre, les modèles utilisés pour analyser, interpréter et communiquer sur l'analyse de l'apprentissage aux parties prenantes (personnel de soutien, conseillers, facultés, étudiants) doivent être solides, sans biais algorithmique ; transparents lorsque cela est possible et clairement compris par les utilisateurs finaux ;
  - Cet aspect implique également une expertise et une rapide montée en compétences du personnel concernant la compréhension de ces données et analyses.
- La responsabilité des institutions envers les différentes parties prenantes face à l'apparition d'une obligation morale d'agir.
  - Bien qu'il ne soit pas toujours possible d'intervenir (en raison d'un manque de ressources), il est important que les institutions mettent a minima en place des processus permettant l'identification et la réaction face à ces problèmes.

● Des communications plus régulières, directes et toujours avec un objectif de transparence sont à mettre en place afin de permettre à l'ensemble des parties prenantes de comprendre l'utilisation et l'analyse des données éducatives tout en permettant la progression de l'apprentissage.

- La dimension culturelle et les valeurs associées sont à ne pas négliger lors de l'analyse de données issues de pays avec des contextes culturels différents.
  - Dans des contextes multiculturels, la compréhension et l'interprétation des données sont d'autant plus importantes ;
  - De ce fait, des précautions doivent être prises lors de l'achat d'analyses auprès de développeurs afin de s'assurer que l'approche correspond bien aux attentes et objectifs identifiés et si, le cas échéant, une adaptation de l'analyse doit être effectuée.

● L'inclusion de toutes les parties prenantes.

- Les données éducatives sont avant tout pensées pour soutenir l'apprentissage de tous, c'est pourquoi il faudra veiller à ne pas exclure certains élèves et étudiants en les catégorisant trop directement à partir des données collectées, celles-ci ne reflétant qu'un aspect de la situation.

● L'importance du consentement des élèves et étudiants, concernant l'utilisation de leurs données.

- Si le consentement est demandé lors de l'inscription, il devra inclure la transparence, des options ultérieures concernant l'ajout de données supplémentaires ainsi que la suppression de données ;
- Une autre approche fait la distinction entre consentement initial (lors de la collecte de données) et consentement spécifique (lors de l'utilisation de ces données).

● L'engagement et la responsabilité des étudiants.

- Impliquer activement les élèves et étudiants dans l'analyse des données éducatives et donner à voir leurs effets sur l'apprentissage est

(19) Joint Information Systems Committee (JISC), « Code of practice for learning Analytics », 2016.

(20) International Council for Open and Distance Education, « Global Guidelines for ethically-informed practices in learning analytics », 2019.

recommandé (dans la mesure du possible) ;

- Cela permettrait notamment de parvenir à une interprétation plus précise des données relatives aux comportements et stratégies d'apprentissage des élèves et étudiants, tout en permettant le développement d'une forme de consentement plus « éclairé » de la part des élèves et des étudiants.

### **Le développement de cas d'usages.**

Actuellement, de plus en plus de solutions numériques ayant recours aux données éducatives sont mises en place afin d'améliorer les conditions d'apprentissage des élèves et des étudiants et des professeurs, les travaux des chercheurs et faire évoluer le pilotage des systèmes éducatifs.

Une étude menée par le *Center for Digital Education* auprès d'enseignants du supérieur aux Etats-Unis a permis d'identifier les bénéfices majeurs des données éducatives. Ces bénéfices

concernent notamment le suivi et prédiction des performances des élèves (selon 69% des participants à l'étude), l'augmentation du taux de diplômés (selon 61% des participants), l'ajustement en temps réel des programmes (selon 47% des participants), la mesure des performances institutionnelles des établissements (selon 44% des participants) et l'identification de failles dans l'administration (selon 22% des participants).

Nous avons identifié plusieurs exemples de cas d'usage à valeur ajoutée, en cohérence avec les objectifs du Ministère de l'Éducation nationale cités plus précédemment (cf. *Des ambitions freinées par un pilotage siloté - P6*).

### **Un service aux professeurs amélioré.**

Dans le domaine de l'amélioration des services aux professeurs grâce à l'utilisation des données éducatives, deux cas d'usages principaux se distinguent : l'optimisation du temps et l'accès à des connaissances renforcées sur les besoins spécifiques des élèves.

Afin d'optimiser le temps des professeurs, différents acteurs ont mis au point des algorithmes utilisant l'intelligence artificielle permettant de collecter, traiter et valoriser les données éducatives qui leur sont transmises. En voici deux exemples :

- Une intelligence artificielle<sup>21</sup> analyse les documents envoyés par les étudiants puis les classe selon différents critères et tendances. Le professeur ne note ensuite que les différents groupes identifiés par l'IA, gagnant un temps considérable lorsque les documents sont particulièrement nombreux, ce qui lui permet notamment de se concentrer sur l'analyse des résultats et la compréhension des comportements d'apprentissage des étudiants (que la plateforme offre également d'identifier).

Les avantages de cette IA sont cependant à nuancer et à mettre en perspective. En effet, plusieurs limites et dérives pourraient découler d'une utilisation trop systématique d'une telle solution : distanciation entre les élèves et le professeur, catégorisation des élèves et perte de l'aspect personnalisé du suivi du fait de la généralisation des profils, etc.

- L'Université de Murcia (Espagne) a lancé un agent conversationnel destiné à répondre aux questions des étudiants. Sur 38 000 questions, ce dernier a répondu correctement plus de 91% du temps, faisant gagner un temps considérable aux équipes éducatives. Il a aussi permis d'augmenter la motivation des étudiants, tout en collectant de nombreuses données éducatives (permettant *in fine* de proposer de nouveaux cours et services, correspondant à leurs demandes). Des solutions similaires sont par ailleurs en cours de tests dans d'autres universités (Royaume-Uni, Etats-Unis, etc.).<sup>22</sup>

Les données éducatives permettent également aux professeurs d'accéder aux connaissances plus détaillées sur les besoins spécifiques de leurs élèves grâce au développement de :

- Tableaux de bord, afin de suivre la progression de leurs élèves, de leurs fixer des objectifs, de les aider en cas de besoins, de déceler des cas



(21) Développée par Sense.

(22) Rouhiainen L., « How AI and Data Could Personalize Higher Education », Harvard Review, octobre 2019.

de décrochage... (exemples d'applications privées proposant ce type de solutions : Panopto, Knewton, Sown To Grow, LectureTools...);

- Plateformes internes collaboratives, qui permettent de mutualiser les informations que possèdent différents membres du corps enseignants (exemple d'applications privées proposant ce type de solutions : Emote).

### Un outillage des prises de décisions.

À un autre niveau, les données éducatives permettent également l'évaluation des performances des établissements vis-à-vis des objectifs fixés au préalable par les institutions tutelles et ainsi de les aider dans la prise de décision en vue d'améliorer leurs performances globales.

Des solutions numériques permettent aujourd'hui de réaliser les cas d'usage suivants :

- L'ensemble des jeux de données qu'un établissement a à sa disposition est regroupé au sein d'une même base de données : résultats académiques et performances des élèves, taux de présence et diversité des cours, besoins et objectifs du corps enseignant, etc.

- Ces données sont ensuite analysées au regard des objectifs stratégiques fixés par l'établissement et par ses institutions tutelles. Grâce à ses différentes analyses, mais également aux éventuelles prédictions de progression communiquées, la prise de décision est facilitée.

Cependant, afin de garantir le succès de telles opérations, il est nécessaire de veiller à plusieurs éléments, dont notamment l'interopérabilité et la qualité des données. Un travail en amont est donc nécessaire afin de permettre l'intégration de l'ensemble de ces données à la base de données qui les regroupe, les analyse et in fine aide à la prise de décision tout en évaluant les performances d'un établissement.

### Un environnement pédagogique personnalisé.

L'analyse des données éducatives permet également d'envisager de per-

sonnaliser l'environnement d'apprentissage de l'élève.

Personnaliser l'environnement et donc l'apprentissage des élèves, c'est s'adapter à chacun des profils de l'apprenant : son profil personnel, son profil d'apprentissage (stratégies cognitives, style et rythme d'apprentissage, etc.) son profil de connaissances et de compétences, et son profil technologique. Cela signifie également de permettre à tous les apprenants d'accéder et de conserver les traces de leurs apprentissages (parcours, réalisations, etc.) afin que, s'ils le souhaitent, ils puissent en avoir un suivi.

Lors de ses recherches sur le développement de dispositifs en ligne permettant de personnaliser l'apprentissage, L. Sauvé a retenu 3 façons de person-

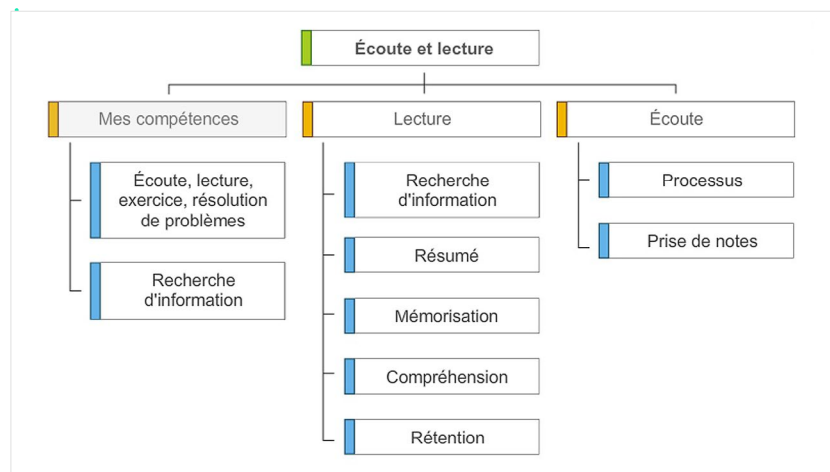
naliser la démarche d'apprentissage<sup>23</sup> :

- Par grille d'énoncés des difficultés à résoudre (l'apprenant liste les difficultés à résoudre et compétences à acquérir);
- Par carte conceptuelle (démarche plus visuelle au cours de laquelle l'apprenant regroupe les difficultés et compétences à acquérir par catégories);
- Par une démarche libre (où l'apprenant est libre de déterminer les mots-clés et ressources nécessaires à la création de sa démarche d'apprentissage).

Elle tire par ailleurs la conclusion que la présence de contenus interactifs et multimédias, utiles et adaptés aux attentes et aux besoins des participants permettent de maintenir leur intérêt, d'autant plus lorsqu'il s'agit de ressources d'apprentissage traitant d'un sujet précis et exigeant.

#### Démarche d'apprentissage par grille d'énoncés des difficultés à résoudre (en haut) et par cartes conceptuelles (en bas), d'après Sauvé L.<sup>24</sup>

Fait	Activité
<input type="checkbox"/>	Lire le Guide de navigation
<b>Communiquer en différé</b>	
<input type="checkbox"/>	Connaître les outils de communication en différé
<input type="checkbox"/>	Juger de la manière de se comporter dans des situations particulières liées à l'envoi et à la réception de courriels
<input type="checkbox"/>	Se former aux fonctions d'ordre technique qui sont propres aux gestionnaires de messages électroniques
<input type="checkbox"/>	Compléter les exercices autocorrectifs
<b>Communiquer en direct</b>	
<input type="checkbox"/>	Connaître les outils de communication en direct
<input type="checkbox"/>	Juger de la manière de se comporter dans des situations particulières liées à la téléprésence.
<input type="checkbox"/>	Se former aux fonctions d'ordre technique qui sont propres aux logiciels clients de messagerie instantanée ou aux applications de visioconférence.
<input type="checkbox"/>	Compléter les exercices autocorrectifs
<input type="checkbox"/>	Évaluer la convivialité de l'environnement
<b>Enregistrer du contenu</b>	
<input type="checkbox"/>	Connaître les outils d'enregistrement de contenu.
<input type="checkbox"/>	Juger de la manière de se comporter dans des situations particulières liées à l'enregistrement de contenu
<input type="checkbox"/>	Se former aux fonctions d'ordre technique qui sont propres aux logiciels d'enregistrement de contenu
<input type="checkbox"/>	Compléter les exercices autocorrectifs



(23) Sauvé L., « Des dispositifs en ligne pour personnaliser l'apprentissage tout au long de la vie : quelques recommandations », Distances et médiations des savoirs, 2014.

(24) *Ibid.*

En pratique, plusieurs solutions privées se basant sur les données éducatives ont été mises au point afin de personnaliser l'apprentissage. Elles proposent :

- La proposition de recommandations et exercices personnalisés ;
- L'accès à des tableaux de bord donnant à voir l'apprentissage (résultats, progression, objectifs à atteindre...);
- La possibilité de se fixer ses propres objectifs (les objectifs n'étant plus uniquement fixés par les professeurs) ;
- L'accès à des retours personnalisés (s'adaptant aux profils de l'élève, selon

sa performance mais aussi selon les objectifs fixés au préalable) ;

Plusieurs études sont par ailleurs venues confirmer les effets positifs (mais limités) de la personnalisation sur l'apprentissage des élèves. Elles ont notamment observé les progrès suivants :

- Le recours aux contenus personnalisés a permis de faire augmenter les notes aux évaluations et aux tests d'environ 1,4 points de pourcentage. A titre d'illustration, cela permet à un élève ayant un taux de réussite initial de 50% d'obtenir un taux de 51,4%,

soit une note de 10,28/20 (note initiale : 10/20).<sup>25</sup>

- Un établissement a dépassé ses objectifs de progression de plus de 200% au MAP® Language Usage Test pour 37% de ses élèves suite au déploiement de contenus et exercices personnalisés. A titre d'illustration, la progression annuelle d'un élève est passée de 10% (estimation) à 30% (après utilisation de contenus personnalisés), soit un passage de 12/20 à 15,6/20 (au lieu d'une estimation initiale à 13,2/20).<sup>26</sup>



(25) Johns Hopkins University, Center for Research and Reform in Education, « Study of Knewton Online Courses for Undergraduate Students : Examining the Relationships Among Usage, Assignment Completion, and Course Success », 2018.

(26) WestEd, « Correlation Between NoRedInk Performance and Achievement on the MAP Language Usage Assessment », 2017.

# Les conditions à de nouveaux usages en France.

## **Un encadrement à deux échelles.**

### **La nécessité d'un cadre légal précis.**

Afin de permettre l'émergence de l'utilisation des données éducatives, il est nécessaire de définir un cadre légal précis pour leur collecte, leur gestion et leur exploitation. Ce cadre doit être conforme aux réglementations et recommandations existantes en termes de protection de la donnée, et doit les compléter en définissant clairement :

- Les modalités de collecte des données éducatives, c'est-à-dire les droits de collecte des différentes données personnelles (des enseignants, des élèves et de leurs familles), scolaires (résultant des interactions entre élèves et enseignants : relevés de notes, etc.), ou d'utilisation des équipements, mais aussi les modes et moyens de collecte : pour chaque type de donnée collectée, quel niveau d'information de l'utilisateur est requis ? comment le consentement est-il recueilli ? par qui est-il accordé lorsque l'utilisateur est un élève ? quels types d'outils (équipements, logiciels) sont autorisés à collecter les données ? sous quelles conditions ? qui les administre ?
- Les modalités de conservation des données éducatives, c'est-à-dire les infrastructures autorisées à les conserver, les modes de protection requis pour ces infrastructures et la durée

maximale de conservation des différents types de données.

- Les modalités d'exploitation des données éducatives, c'est-à-dire le niveau de confidentialité de chaque type de données (qui a le droit d'accéder à chaque type de donnée ? chaque type de donnée peut-il être partagé sur le modèle des données ouvertes ?), les conditions d'extraction et de croisement des données (sous quelles conditions peut-on par exemple croiser des données personnelles avec des données scolaires ?), et évidemment les conditions de leur utilisation (à quelle fin peut-on utiliser les données éducatives ? par quel moyen ?)
- Les modalités de gestion des réclamations, c'est-à-dire les droits des utilisateurs de demander l'accès à ou la suppression des données éducatives qui les concernent (quelles données sont considérées ou non comme personnelles ? quels droits l'utilisateur a vis-à-vis de ces données ?) et les devoirs de chaque organisme collectant, conservant ou exploitant ces données pour garantir l'exercice de ces droits.
- Les rôles et dispositifs à mettre en place dans chaque établissement pour garantir l'application de ce cadre.

Ce cadre légal doit être conçu en tenant compte des enjeux d'éthique liés au traitement. Il doit garantir le respect des libertés fondamentales en fournissant une définition claire des données d'éducation personnelle et de l'utilisation qui en est autorisée, en favori-

sant leur protection, et en évitant que puissent être déduites de ces données (par exemple par croisement) et/ou exploitées des données personnelles sensibles. Les modalités de recueil du consentement des élèves, familles et enseignants à la collecte doivent aussi être encadrées pour s'assurer que le refus de ce consentement ne défavorise pas un élève par rapport aux autres.

Les nouveaux usages pouvant émerger de l'exploitation de ces données posent aussi des questions d'éthique que le cadre légal doit anticiper. Par exemple, si l'utilisation des données pour automatiser certaines tâches simples permet de libérer du temps aux enseignants pour se consacrer à d'autres tâches, il ne faut pas que cette automatisation remplace totalement des tâches comme la correction des travaux d'élèves, pour lesquelles leur expertise, leur connaissance du parcours des élèves, et la nuance de leur jugement est nécessaire. De la même façon, les usages liés aux données éducatives ne doivent pas briser le lien humain entre les enseignants et les élèves.

### **Une meilleure sensibilisation des personnels.**

Pour appuyer l'application de ce cadre légal, il est important de sensibiliser les personnels des établissements aux bonnes pratiques d'utilisation des équipements numériques exploitant



ou donnant accès aux données collectées. Ces bonnes pratiques et les règles auxquelles elles sont associées doivent être accessibles facilement par toutes et tous via une documentation en ligne et papier complète, rédigée en des termes simples et compréhensibles. Un travail important de communication doit être mené par chaque académie pour permettre à tous les établissements d'identifier clairement les sources d'information disponibles.

Des formations ou ateliers de sensibilisation peuvent aussi être organisés par les académies pour permettre au personnel de chaque établissement de s'approprier ces règles et pratiques. Le/la chef(fe) d'établissement, en tant que responsable du traitement des données, et son/sa ou ses adjoint(e)(s) doivent être convié(e)(s) à ces sessions durant lesquelles ils/elles sont encouragé(e)s à partager clairement au sein de l'établissement :

- Les droits et devoirs de chacun(e) vis-à-vis des données, les obligations et responsabilités de l'établissement et des enseignants comme celles des élèves ;
- Les rôles de chacun(e), par exemple si un(e) référent(e) est désigné(e) en

appui au/à la chef(fe) d'établissement pour encadrer le traitement des données ;

- Les processus en cas de réclamation : à qui s'adresser pour accéder à ou demander la suppression de données, par quel biais, etc. ;
- Les bonnes pratiques d'utilisation des équipements numériques ou de consultation des données, en particulier des règles garantissant leur confidentialité et définissant les modalités de leur extraction.

### **La création d'un Education Data Hub.**

#### **Un accès aisé, unifié, transparent et sécurisé.**

Sur les bases de ce qui a été mis en place dans le domaine de la santé avec la création d'un *Health Data Hub* en 2019, le développement d'un *Education Data Hub* – plateforme de stockage virtuel de données éducatives provenant de sources multiples et facilitant leur analyse – pourrait être imaginé.

En effet, à date, les différentes ressources et bases de données éduca-

tives ne communiquent pas (silotage des acteurs, manque de transparence, dispersion des données...). La construction d'un *Education Data Hub* permettrait :

- La centralisation et le partage des données avec la création d'un guichet unique, dont l'accès aux données éducatives serait encadré par des chartes définies selon les différents profils des utilisateurs (exemples de profils : les producteurs de données, les utilisateurs de données et les citoyens, propriétaire des données) ;
- La protection des données, grâce à une plateforme technologique sécurisée. L'*Education Data Hub* pourrait notamment s'inscrire dans le projet de cloud GaïaX (infrastructures européennes de données sécurisées permettant d'éviter les inquiétudes créées par un hébergement auprès de sociétés américaines ne disposant pas des mêmes normes de sécurité et de partage de la donnée)<sup>27</sup> ;
- La mise en valeur et la qualité des données, l'*Education Data Hub* aurait ici un rôle de garant de la qualité des données partagées, de leur standardisation et de la valorisation de la collecte de ces données. Il jouerait également

(27) Digital New Deal (think tank), « Pour un Educational Data Hub au service de l'ouverture des données éducatives », avril 2021.



le rôle de vitrine nationale et internationale de notre patrimoine de données éducatives, mettant à disposition une véritable offre de service à destination d'acteurs publics et privés ;

- L'exploration de cas d'usage de données éducatives remontées par les métiers : un *Education Data Hub* créateur d'un environnement propice et promoteur de l'innovation et de la recherche ;
- La création d'un tiers de confiance, catalyseur et fédérateur d'une communauté des acteurs de la donnée éducative autour d'ambitions et d'objectifs communs (la *EdTech*). L'*Education Data Hub* pourrait également avoir le rôle de facilitateur de la mutualisation des ressources technologiques et humaines (circulation facilitée et partage des données entre les acteurs, mise à disposition d'expertises propres, mise en avant d'une culture de la donnée éducative...);
- *In fine* la relance d'une politique d'ouverture de la donnée en matière éducative.

### Plusieurs chantiers essentiels à mener.

Afin de répondre aux enjeux techniques, organisationnels et culturels soulevés par la construction d'un *Education Data Hub* et de garantir sa montée en puissance, plusieurs chantiers (non exhaustifs) sont essentiels :

- Identifier les cas d'usage à valeur ajoutée dans une démarche collective impliquant les utilisateurs « métier » : corps enseignant, direction d'école, parents d'élèves, etc. ;
- Identifier les données associées à ces cas d'usage à mettre dans ce *Education Data Hub*, les prioriser selon leur nature (données administratives/métier) ;
- Amorcer un travail d'harmonisation des données au niveau national, ainsi qu'une réflexion sur les travaux attendus en termes de complétude, de qualité et de fraîcheur de la donnée, pour garantir l'interopérabilité des données et leur capacité à dialoguer. Par ailleurs, la mise à jour du Référentiel Général d'Interopérabilité (2016) facilite l'exploitation et l'échange de données éducatives au niveau national<sup>28</sup> ;
- Insuffler une démarche collective de mise en qualité et de circulation de la donnée éducative en termes d'organisation, de culture mais également

(28) NB. La mise à jour du Référentiel Général d'Interopérabilité (2016) facilite l'exploitation et l'échange de données éducatives au niveau national : elle introduit la notion de profils d'interopérabilité, aborde la question des enjeux sémantiques en introduisant la notion d'identité pivot et recommande enfin de moins utiliser certains formats de données.



de processus afin de faciliter le travail d'harmonisation mais également pour que l'ensemble des acteurs de la communauté éducative soit en capacité de se saisir de ces données ;

- Structurer une gouvernance de la donnée, choisir la forme juridique adaptée à un *Education Data Hub* et disposer de compétences dédiées<sup>29</sup> afin de garantir la valorisation de la donnée et la prévenir des dérives tout en garantissant la maîtrise de l'Etat sur ces données ;
- Construire une architecture technique répondant à l'hétérogénéité des données, des acteurs et des outils SI utilisés.

### **Un appui sur les cas d'usages existants.**

Sur la base de notre étude des données éducatives à l'étranger, trois cas d'usage ont particulièrement retenu notre attention :

- La mise en place de tableaux de bord pédagogiques pour gérer le suivi des

élèves de manière précise et individualisée (cf. *La généralisation des tableaux de bord*) ;

- L'utilisation de chatbots pour faire gagner du temps aux enseignants et améliorer les pratiques pédagogiques (cf. *L'utilisation d'agents conversationnels*) ;
- L'optimisation du pilotage et de la performance des établissements scolaires via la mise en place de tableaux de gestion agrégeant les données disponibles (cf. *La mise en place de tableaux de gestion*).

### **La généralisation des tableaux de bord.**

Le tableau de bord est un outil constitué d'un ensemble d'indicateurs qui, appliqué dans le milieu scolaire, peut permettre d'assurer un suivi plus précis et plus individualisé élèves. À partir des données générées par l'utilisation d'outils numériques dans le cadre pédagogique, les enseignants pourraient en effet suivre et adapter les activités en fonction de chaque profil d'élève<sup>30</sup>.

De plus en plus d'enseignants expriment la nécessité d'utiliser ce type d'outil tant pendant qu'en dehors de la classe pour mieux accompagner leurs élèves, notamment à travers :

- La présentation des statistiques de progression au niveau du groupe et/ou de l'élève ;
- La mise en place de retours individualisés construits sur la base des résultats de l'élève et du type de fautes qu'il a commises ;
- La recommandation personnalisée d'exercices pour répondre aux besoins de chaque élève.

Parvenir à ce niveau de personnalisation de l'enseignement implique toutefois que les données traitées soient interopérables, c'est-à-dire la capacité à agréger des données issues de sources différentes. Pour pouvoir proposer des conseils et des exercices individualisés, il faut en effet accéder à des masses de données recueillies tout au long du parcours de chaque élève. Or, ces données sont actuellement

(29) Compétences à mobiliser dans le cadre de la montée en puissance rapide d'un *Education Data Hub* (liste non exhaustive) : data management, data science, sécurité informatique, expertise juridique, pilotage de programme, conduite du changement, méthodologie agile, UX design, expertise en bases de données, etc.

(30) Lecuyer, G., « Quels indicateurs pour personnaliser les apprentissages dans une école numérique », Université Pierre et Marie Curie, 2016.



kées dans des espaces numériques fermés, souvent par souci de sécurité), ce qui rend leur accès et leur exploitation difficile<sup>31</sup>.

Le constat fait par le rapport pour une politique publique de la donnée de décembre 2020<sup>32</sup> montre que les marges de progression de la France en termes d'interopérabilité des données sont importantes, y compris dans le champ éducatif :

- D'une part, la conversion des données vers un format standard commun et normalisé peut et doit s'inscrire dans les objectifs des pouvoirs publics ;
- D'autre part, dans la mesure où il est fréquent que plusieurs jeux de données existent sur un même sujet (du fait du nombre croissant des producteurs de données), le référencement et la validation de chacun d'entre eux sont devenus essentiels pour faciliter le croisement des informations.

### L'utilisation d'agents conversationnels.

À mesure que se développe la technologie des agents conversationnels dits « chatbots » (logiciels-robots capables de dialoguer avec un utilisateur via une plateforme de messagerie), des nouveaux cas d'usage apparaissent dans le domaine de l'éducation.

Les enseignants manquant généralement de temps et de ressources, les agents conversationnels présentent en effet de nombreux avantages :

- Ils peuvent répondre à des questions sur les dates de rendu, les plans de cours et les programmes ;
- Ils sont capables d'assister les enseignants en termes de veille (informations sur les nouvelles normes ou sur les nouveaux modes de notation par exemple) ;
- Ils peuvent aider les élèves à comprendre des notions difficiles, comme si elles étaient expliquées par une autre personne.

L'agent conversationnel n'a donc pas vocation à remplacer l'enseignant mais à bien l'assister dans ses tâches. Il permet également de différencier l'enseignement pour chaque élève car la conversation virtuelle s'adapte aux réponses et aux demandes de celui-ci

(en langue en fournissant du vocabulaire, en français pour apporter des définitions, en mathématiques pour aider à construire une démarche de résolution, etc.).

L'agent conversationnel peut aussi permettre d'accompagner les élèves sur des sujets qui ne concernent pas directement les travaux de classe (informations sur l'établissement, informations sur le règlement intérieur, réponses à des questions d'orientation, etc.)<sup>34</sup>.

Les enseignants ont donc tout intérêt à se saisir de cette technologie, non seulement pour enrichir leurs pratiques pédagogiques mais aussi parce que chaque tâche qu'un agent conversationnel réalise à sa place lui donne plus de temps pour se concentrer sur son rôle principal : instruire.



### La mise en place de tableaux de gestion.

La volonté d'améliorer la qualité des prestations scolaires est partagée par l'ensemble des acteurs de cette politique publique. Les établissements scolaires sont en effet de plus en plus amenés à formuler des objectifs de résultats et à définir des stratégies permettant de les atteindre<sup>35</sup>.

Dans cette perspective, l'agrégation des données collectées par les établissements scolaires au sein d'une même base de données permettrait de faciliter la prise de décision tout en renforçant la pertinence des objectifs fixés. Les données éducatives pourraient par exemple être rassemblées au sein d'un tableau de pilotage et utilisées comme indicateurs statistiques livrant des informations :

- Sur les élèves (nombre, taux de maintien dans le cycle, résultats, etc.)
- Sur la vie scolaire (taux de participation aux réunions, qualité des relations avec les parents, évolution des effectifs, etc.)
- Sur les autres établissements scolaires (taux de réussite national, régional, etc.)
- Sur les moyens et les conditions matérielles (parc informatique, locaux, infrastructures sportives, etc.)
- Sur l'équipe pédagogique (ancienneté des enseignants, compétences spécifiques, etc.)
- Sur le fonctionnement pédagogique (circulation de l'information, actions d'ouverture culturelle, etc.)

La disponibilité et l'analyse de ces informations favoriserait le pilotage des établissements scolaires tant dans une dimension stratégique (en permettant d'élaborer des objectifs à long terme avec les stratégies qui s'y rapportent) que dans une dimension opérationnelle (en améliorant la gestion de l'établissement sur les plans de l'organisation, de l'administration, des finances, des ressources humaines et des infrastructures).

(31) Open Assessment Technologies (société), « L'interopérabilité : Le chaînon manquant de l'EdTech », décembre 2019.

(32) Bothorel, Éric, Rapport au premier ministre, « Pour une politique publique de la donnée, des algorithmes et des codes sources », décembre 2020.

(33) Snatchbot (solution numérique), « Améliorer la salle de classe avec les chatbots », 2021.

(34) Geschwine, L., Région académique Bourgogne-Franche-Comté, « Les chatbots en éducation », juillet 2020.

(35) Di Natale, J., « Contenu et structuration d'un tableau de pilotage des établissements scolaires », Institut des hautes études en administration publique, 2011.



# Conclusion.

En mettant à jour l'ampleur des inégalités d'accès à l'éducation, la crise sanitaire a accru la nécessité d'accélérer la transformation numérique du secteur éducatif. Dans ce contexte, le développement du numérique éducatif doit s'appréhender non comme une fin en soi mais comme un moyen destiné à garantir non seulement l'égalité des chances mais aussi l'adaptation du service public de l'éducation aux nouvelles pratiques liées au numérique ; pratiques que les élèves devront maîtriser dans leur vie professionnelle et sociale.

Les données éducatives ont une place essentielle dans cette évolution. À condition que leur exploitation soit juridiquement encadrée, elles peuvent être utilisées pour enrichir les pratiques pédagogiques, personnaliser l'enseignement et accroître les performances du système éducatif. Or, ce potentiel des données éducatives n'est pas encore réellement exploité en France. C'est la raison pour laquelle il est essentiel de relancer la politique d'ouverture des données en matière éducative : leur disponibilité et leur valorisation offriront de véritables opportunités dans le domaine de la pédagogie et de la gestion administrative de la scolarité.

*Mikaël Aït Si Ahmed,  
Justine Buvat,  
Mathilde Manson,  
Simon Rumel-Sixdenier*

# Bibliographie.

## Rapports.

- .Inspection générale de l'Éducation nationale, Inspection générale de l'administration de l'Éducation nationale et de la Recherche, Rapport n°2018-016, « Données numériques à caractère personnel au sein de l'Éducation nationale », février 2018.
- .Bothorel, Éric, Rapport au premier ministre, « Pour une politique publique de la donnée, des algorithmes et des codes sources », décembre 2020.
- .Consortium LACE, « LACE LAW Manifesto », 2015
- .International Council for Open and Distance Education (ICDE), « Global Guidelines for ethically-informed practices in learning analytics », 2019.
- .JISC, « Code of practice for learning Analytics », 2016
- .Johns Hopkins University, Center for Research and Reform in Education, « Study of Knewton Online Courses for Undergraduate Students: Examining the Relationships Among Usage, Assignment Completion, and Course Success », 2018
- .WestEd, « Correlation Between NoRedInk Performance and Achievement on the MAP Language Usage Assessment », 2017

## Articles.

- .Di Natale, J., « Contenu et structuration d'un tableau de pilotage des établissements scolaires », Institut des hautes études en administration publique, 2011.
- .Digital New Deal (think tank), « Pour un Educational Data Hub au service de l'ouverture des données éducatives », avril 2021.
- .Geschwine, L., Région académique Bourgogne-Franche-Comté, « Les chatbots en éducation », juillet 2020.
- .Lecuyer, G., « Quels indicateurs pour personnaliser les apprentissages dans une école numérique », Université Pierre et Marie Curie, 2016.
- .Mediavilla L., Loi sur les données personnelles : la Californie ouvre le bal aux Etats-Unis, Les Echos, janvier 2020
- .Open Assessment Technologies (société), « L'interopérabilité : Le chaînon manquant de l'EdTech », décembre 2019.
- .Rouhiainen L., "How AI and Data Could Personalize Higher Education", Harvard Review, octobre 2019.
- .Sauvé L., « Des dispositifs en ligne pour personnaliser l'apprentissage tout au long de la vie : quelques recommandations », Distances et médiations des savoirs, 2014.
- .Snatchbot (solution numérique), « Améliorer la salle de classe avec les chatbots », 2021.

## Documents ministériels.

- .Ministère de l'Éducation nationale, de la Jeunesse et des Sports, plateforme Open Data ([data.education.gouv.fr](https://data.education.gouv.fr)).
- .Ministère de l'Éducation nationale, de la Jeunesse et des Sports, « Les enjeux de la protection des données au sein de l'éducation », février 2020.
- .Ministère de l'Éducation nationale, de la Jeunesse et des Sports, Comité d'éthique pour les données d'éducation, « Enjeux d'éthique des usages des données numériques d'éducation dans le contexte de la pandémie », juillet 2020.
- .Ministère de l'Éducation nationale, de la Jeunesse et des Sports, Eduscol, « Délégués à la protection des données », octobre 2020.
- .Ministère de l'Éducation nationale, de la Jeunesse et des Sports, Eduscol, « Délégués académiques au numérique », mars 2021.
- .Ministère de l'Éducation nationale, de la Jeunesse et des Sports, « Direction de l'évaluation, de la prospective et de la performance », mars 2021.
- .Ministère de l'Éducation nationale, de la Jeunesse et des Sports, « Direction du numérique pour l'éducation », avril 2021.
- .Ministère de l'Éducation nationale, de la Jeunesse et des Sports, « Le numérique au service de l'école de la confiance », mars 2021.

## Documents législatifs.

- .Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés.
- .Loi n°2018-493 du 20 juin 2018 relative à la protection des données personnelles.

## Cartographies.

- .CNUCED, Online Consumer Protection Legislation Worldwide
- .CNIL, « La protection des données dans le monde », novembre 2020

# Vos contacts.

## Bertrand Le Moigne

Directeur Associé

Secteur Public

+33 6 20 84 44 24

[bertrand.lemoigne@sia-partners.com](mailto:bertrand.lemoigne@sia-partners.com)

## Pierre Artaud

Directeur Associé

Secteur Public

+33 6 59 48 51 93

[pierre.artaud@sia-partners.com](mailto:pierre.artaud@sia-partners.com)

- Abou Dabi
- Allemagne
- Arabie Saoudite
- Belgique
- Canada
- Émirats Arabes Unis
- États-Unis
- France
- Hong Kong
- Italie
- Irlande
- Japon
- Luxembourg
- Maroc
- Panama\*
- Pays-Bas
- Qatar
- Royaume-Uni
- Singapour

\*Sia Partners Panama, membre du réseau Sia Partners

A thick teal diagonal bar that tapers from left to right, positioned above the main title.

# À propos de Sia Partners.

Pionnier du Consulting 4.0, Sia Partners réinvente le métier du conseil et apporte un regard innovant et des résultats concrets à ses clients. Nous avons développé des solutions basées sur l'Intelligence Artificielle et le design pour augmenter l'impact de nos missions de conseil. Notre présence globale et notre expertise dans plus de 30 secteurs et services nous permettent d'accompagner nos clients dans le monde entier. A travers notre démarche «Consulting for Good», nous mettons notre expertise au service des objectifs RSE de nos clients et faisons du développement durable un levier de performance pour nos clients.

[www.sia-partners.com](http://www.sia-partners.com)