

Rapport à monsieur le ministre de l'éducation nationale
et de la jeunesse
madame la ministre de l'Enseignement supérieur
et de la Recherche
madame la ministre des sports et des jeux
Olympiques et Paralympiques

**Les conventions d'échange de données entre les académies
et les collectivités territoriales :
état des lieux, pratiques et préconisations**

N° 20-21 059A - octobre 2022

*Inspection générale de l'éducation,
du sport et de la recherche*

**Les conventions d'échange de données entre les académies
et les collectivités territoriales :
état des lieux, pratiques et préconisations**

Octobre 2022

Michel QUÉRÉ
Nathalie COSTANTINI
Yves DELÉCLUSE

*Inspecteurs généraux de l'éducation,
du sport et de la recherche*

Une dédicace toute particulière pour le présent rapport est à formuler à l'encontre d'Alain Abécassis. Les points de vue argumentés tout comme la série de préconisations envisagées ont largement bénéficié de son expertise et de sa contribution au collectif de travail qui, au final, aura rédigé ce rapport. Pertinence dans la réflexion, acuité du jugement, malice dans l'analyse et souci de vraisemblance des avancées possibles sont des traits de caractère qu'il a déployés avec nous jusqu'au bout de ce qui lui a été possible de faire dans notre aventure collective. Ce fut un honneur et un bonheur d'avoir cheminé avec lui dans la construction de ce rapport.

SOMMAIRE

Synthèse	1
Liste des recommandations	3
Introduction.....	4
1. La donnée : une variable de plus en plus stratégique dans la mise en œuvre des politiques publiques.....	5
1.1. Une politique nationale générale qui évolue rapidement et est marquée par une accélération des étapes structurantes.....	6
1.1.1. <i>Le cadre général d'exercice de la donnée scolaire : la loi informatique et libertés.....</i>	<i>6</i>
1.1.2. <i>La loi pour la république numérique</i>	<i>7</i>
1.1.3. <i>Le règlement général sur la protection des données (RGPD) et ses incidences</i>	<i>8</i>
1.1.4. <i>Le rapport Bothorel, la circulaire du Premier ministre du 27 avril 2021 et les feuilles de route ministérielles</i>	<i>8</i>
1.1.5. <i>Une nouvelle norme européenne à prendre en compte.....</i>	<i>10</i>
1.2. Attributs et spécificités de la donnée : acteurs, outils et gouvernance (protection de l'utilisateur, sécurité d'usage, articulation finalité / destinataire).....	11
1.3. Les particularismes du champ de l'éducation en matière de gestion de la donnée.....	12
2. Un état des lieux de la gouvernance de la donnée scolaire	13
2.1. Le MENJ	13
2.1.1. <i>Gouvernance et instances.....</i>	<i>13</i>
2.1.2. <i>La feuille de route et la stratégie définie.....</i>	<i>15</i>
2.2. Une administration centrale intervenant en fonction du champ de responsabilité de chacune de ses composantes	16
2.2.1. <i>La DEPP, référent historique de la donnée scolaire.....</i>	<i>16</i>
2.2.2. <i>La DGESCO comme maître d'ouvrage de la politique du numérique éducatif au travers des remontées d'information nécessaires au pilotage du MENJ.....</i>	<i>17</i>
2.2.3. <i>La DNE, à l'interface des directions et en appui technique des acteurs.....</i>	<i>17</i>
2.2.4. <i>La DJEPVA, nouvelle entité du MENJ, autre culture et autres enjeux dans la gestion de la donnée..</i>	<i>19</i>
2.2.5. <i>La DAJ, une direction prudentielle dans la gouvernance de la donnée.....</i>	<i>20</i>
2.3. Le contexte de confinement a confirmé la nécessité du développement du numérique et a mis au premier plan de nouveaux besoins dans le champ de l'éducation	21
2.3.1. <i>Les interfaces écoles - ménages</i>	<i>21</i>
2.3.2. <i>Les ENTs comme outils de liaison et moyens d'enseignement.....</i>	<i>21</i>
2.3.3. <i>Le renforcement du lien entre les acteurs via les échanges numériques</i>	<i>22</i>
2.4. Des services déconcentrés marqués par l'absence de stratégies coordonnées dans la gestion de la donnée scolaire	22
2.4.1. <i>Un état des lieux éclairé par une enquête passée en académie</i>	<i>23</i>
2.4.2. <i>Des acteurs très variés et diversement coordonnés.....</i>	<i>24</i>

2.4.3.	<i>La régionalisation comme facteur d'hétérogénéité dans les échanges avec les collectivités territoriales</i>	25
2.4.4.	<i>Un besoin de doctrine nationale élaborée préalablement pour assurer des échanges avec les collectivités territoriales sécurisés.....</i>	25
2.5.	Une forte attente des collectivités territoriales.....	28
2.5.1.	<i>Une insatisfaction systématiquement relevée sur la disponibilité de la donnée scolaire et sa fiabilité.</i>	28
2.5.2.	<i>La complémentarité des préoccupations entre État et collectivités autour de la donnée scolaire impose un nouveau paradigme</i>	29
2.5.3.	<i>La volonté d'assortir les échanges de données à un projet éducatif partagé avec les services académiques au niveau régional</i>	31
2.5.4.	<i>Les ressources humaines pour produire et utiliser de la donnée</i>	34
3.	Une stratégie qui doit évoluer au bénéfice de l'enfant et du jeune : favoriser une gouvernance territoriale partagée de la donnée scolaire.....	35
3.1.	Définir les conditions de mobilisation des données « élèves »	35
3.1.1.	<i>Développer les apports réciproques entre État et collectivités territoriales</i>	36
3.1.2.	<i>Déterminer un équilibre entre une disponibilité de la donnée consolidée au niveau national et la possibilité de développer des usages spécifiques de la donnée au niveau régional.....</i>	36
3.2.	Une gouvernance à repenser au niveau national comme au niveau académique	37
3.2.1.	<i>D'une gouvernance formelle à une gouvernance opérationnelle au sein de l'administration centrale.</i>	37
3.2.2.	<i>Le rôle déterminant des SSA et leurs liens avec les services statistiques des collectivités territoriales..</i>	38
3.3.	Des actions à conduire pour garantir la réalisation d'évolutions souhaitables	39
3.3.1.	<i>Concilier l'exigence nationale relative à la qualité de la donnée avec un usage facilité au niveau académique, y compris avec des acteurs tiers (collectivités territoriales)</i>	39
3.3.2.	<i>Co-construire avec les collectivités territoriales un entrepôt partagé des données en éducation</i>	40
3.3.3.	<i>Maîtriser les droits d'accès et les conditions d'utilisation de la donnée en éducation pour sécuriser les échanges informels : le choix d'un modèle.....</i>	40
Annexe	49

SYNTHESE

Ce rapport est issu d'une sollicitation formulée par la DEPP en 2020 qui a été inscrite dans la feuille de route numérique ministérielle du MENJ datée de septembre 2021 (action 39-22 S1 : « initier une mission d'évaluation de l'inspection générale de l'éducation, du sport et de la recherche (IGÉSR) portant sur les conventions d'échange de données entre les académies et les collectivités territoriales (état des lieux, pratiques et préconisations) »).

Cette demande de la DEPP est à relier au constat d'une relative insatisfaction quant à l'encadrement des formes actuelles de relations en termes de données numériques entre les académies et les collectivités territoriales. Cette insatisfaction conduit à interroger principalement, soit la déclinaison académique d'un conventionnement établi entre la DEPP et l'ARF en 2015, soit la mise en œuvre des collaborations issues du « comité national des partenaires », comité animé par la DNE qui réunit des représentants du MENJ et des représentants des collectivités territoriales régionales.

Dans les deux cas, les médiations engagées n'épuisent pas le sujet d'une nécessaire relation d'échange rendue complexe de par la rapidité des évolutions technologiques, juridiques et organisationnelles qui affectent la production et les échanges de données dans le champ de l'éducation. C'est l'objet de la première partie de ce rapport que de rendre compte de cette complexité au travers une revue des évolutions contemporaines dans l'usage des données. Les rédacteurs ont souhaité insister tout particulièrement sur les enjeux de la mise en place progressive de la RGPD dans le champ de l'éducation et sur les décisions interministérielles récentes qui impactent fortement la gouvernance de la donnée.

Par ailleurs, les suites du rapport Bothorel (2020) amènent à conforter l'usage du terme de « donnée scolaire », même si ce terme n'est pas encore juridiquement fondé, pour mettre en avant le caractère spécifique du champ scolaire par rapport à d'autres champs sectoriels d'intervention publique (comme la santé par exemple). Les caractéristiques de la donnée tout comme les conditions de ses usages plaident pour reconnaître la montée en puissance des enjeux autour des échanges de la donnée scolaire, mais aussi à relever le caractère spécifique des solutions à mettre en œuvre pour venir répondre aux attentes de l'ensemble des acteurs publics concernés (ministère, services académiques, collectivités territoriales).

C'est en raison de la spécificité propre à la donnée scolaire que la deuxième partie du rapport est consacrée à l'analyse des attentes et des comportements des différents acteurs concernés par ce sujet. Les administrations centrales du MENJ, tout d'abord, investissent ce sujet avec des regards et des stratégies qui sont avant tout dictés par leurs champs fonctionnels d'intervention. Il en découle un sentiment de grande hétérogénéité dans les préoccupations et l'implication dans ce domaine de chacune d'entre elles (DGESCO / DEPP / DNE / DAJ / DJEPVA), ce qui légitime les efforts attendus de la mise en œuvre de la feuille de route numérique pour améliorer la coordination et renforcer une gouvernance commune à l'échelle du MENJ sur ces questions (par l'intermédiaire du comité stratégique de la donnée notamment).

Pour autant, de nombreux progrès restent encore à faire pour que le MENJ dispose d'une doctrine globale et collectivement partagée, que ce soit en administration centrale ou en services déconcentrés (académiques). À ce titre, l'adjonction récente de la DJEPVA dans le périmètre ministériel nécessite une attention forte à tous les échelons du ministère, en administration centrale tout comme en académie : de culture et de pratique d'intervention différentes, l'enjeu de complétude possiblement à trouver dans la compréhension des parcours individuels des élèves (dans et hors le temps scolaire) est un objectif puissamment utile qui doit s'accommoder aussi en matière de cohérence et d'échanges des données.

Cette réalité de l'administration centrale quant à la problématique du rôle, de l'importance et de l'usage de la donnée scolaire est encore plus fortement marquée en académie où une grande variété de situations prévaut. Une enquête passée en académie et accompagnée par les correspondants académiques de l'inspection générale conforte la grande diversité des modes d'échanges entre services académiques et collectivités (régionales le plus souvent mais aussi parfois infrarégionales). Cette enquête permet également de montrer l'intérêt pour le MENJ de disposer d'une meilleure connaissance des réalités territoriales locales. Un recensement systématique des différents modes de conventionnement avec les collectivités reste encore à l'agenda de la stratégie ministérielle. L'enquête permet aussi de constater que les principaux échanges de données concernent les données de gestion des élèves et que les avancées de l'« *open data* » ne suffisent pas pour répondre aux attentes des collectivités qui sont intéressées par des données de gestion en temps

réel (données régulièrement actualisées), ce afin de répondre aux exigences de l'exercice de leurs compétences propres. Enfin, ce travail a permis de mettre en évidence la problématique des espaces numériques de travail (ENTs) et des échanges qu'ils induisent ainsi que la richesse des données dont ni l'État, ni les collectivités ne maîtrisent pleinement les potentialités d'usage. Au final, l'analyse produite, corroborée par des observations de terrain en régions Auvergne-Rhône-Alpes, Corse et Île-de-France, font ressortir avec force la volonté politique des collectivités (particulièrement les régions) de participer aux côtés de l'Éducation nationale à la co-construction de projets éducatifs adaptés aux spécificités de leurs territoires et de mobiliser les données (scolaires et hors scolaires, le cas échéant) à cette fin.

Pour autant, force est de constater que la grande variété des échanges entre services académiques et collectivités territoriales est aussi le reflet d'une appétence et d'une expertise technique pour la mettre en œuvre très hétérogènes d'une région à l'autre.

Par le fruit de l'histoire et de relations interpersonnelles plus ou moins anciennes et consolidées à l'échelle des services techniques entre collectivités et académies, certaines régions sont plus avancées que d'autres dans la capacité à utiliser et à atteindre les finalités assignées aux échanges de données. Cette variété territoriale reflète aussi des contrastes sensibles d'investissements de la part des collectivités, certaines disposant d'une réelle expertise technique de traitement de données là où cette dernière reste sporadique dans d'autres contextes régionaux.

Dans tous les cas étudiés, les services statistiques académiques sont au cœur de la réflexion sur la donnée scolaire dans une configuration délicate selon le niveau d'élaboration de la stratégie numérique territoriale, ce qui pose le problème crucial de la responsabilité des échanges de données.

Les enseignements retirés de ces observations au niveaux national, régional et local amènent à proposer une série de préconisations qui visent à renforcer la prise de conscience amorcée par la feuille de route numérique du ministère et la montée en puissance du rôle stratégique qui sera joué par la donnée scolaire dans les années à venir.

Liste des recommandations

Recommandation n°1 : Tenir des « Assises nationales de la donnée scolaire » : ouvrir la réflexion MENJ - collectivités territoriales - Edtechs pour établir les termes d'un projet partagé autour de la donnée scolaire complétant ressources nationales et ressources régionales.

Recommandation n°2 : Élaborer et installer une doctrine nationale partagée de l'usage de la donnée scolaire permettant une plus grande homogénéité des comportements et des procédures utilisées en académies.

Recommandation n°3 : Renforcer l'acculturation du MENJ aux enjeux de la donnée scolaire : du producteur (écoles et EPLE), des « consolidateurs » (SSA - DEPP ; DGESCO - DNE) et du régulateur (DAJ) pour mieux maîtriser les effets induits de la RGPD, en particulier l'exercice du droit d'information et du droit d'opposition des usagers.

Recommandation n°4 : Revenir sur la convention DEPP ARF de 2015 et la transformer en convention cadre MENJ - collectivités territoriales régionales.

Recommandation n°5 : Élaborer un cadre de concertation mieux structuré entre MENJ et collectivités territoriales au national (comité des partenaires) tout comme en territoire par l'installation d'un comité régional de la donnée scolaire, déclinaison régionale du comité national des partenaires animé par la DNE.

Recommandation n°6 : Renforcer la connivence INJEP-DEPP-SIES en académie par l'intermédiaire des SSA pour développer une compréhension adaptée des parcours scolaires et d'engagement des jeunes, ce qui suppose, au cas par cas, un renforcement de ces services au niveau régional.

Recommandation n°7 : Renforcer le rôle de l'AMDAC et, à travers lui, mieux opérationnaliser la gouvernance nationale et le suivi en territoire de l'usage de la donnée scolaire ;

Recommandation n°8 : Scénariser, dans la suite des assises nationales de la donnée scolaire, une infrastructure partagée de la donnée scolaire dans une complémentarité entre une alimentation nationale et des apports régionaux, et promouvoir l'installation d'une « plateforme » de la donnée scolaire, qui pourrait disposer d'une certaine latitude de différenciation dans ses contenus selon les régions considérées.

Introduction

Ce rapport de mission vient à l'appui de la réalisation de la feuille de route ministérielle du numérique du MENJ au sens où l'éclairage proposé vient concrétiser l'action 39-22S1 : « Initier une mission d'évaluation de l'inspection générale de l'éducation, du sport et de la recherche (IGÉSR) portant sur les conventions d'échanges de données entre les académies et les collectivités territoriales (état des lieux, pratiques et préconisations) » (cf. annexe 1 : *Pour une politique de la donnée ambitieuse au service de l'École, de la Jeunesse et des Sports*, feuille de route politique de la donnée du MENJS, page 27).

Au-delà de cette commande, la mission trouve racine dans une demande initiée par la direction de l'évaluation, de la prospective et de la performance (DEPP) et relative aux effets d'une convention passée en 2015 entre la DEPP¹ et l'association Régions de France (ARF) pour consolider les relations d'échanges de données entre services académiques et les collectivités territoriales régionales. Les évolutions réglementaires opérées depuis 2015² ainsi que les difficultés de mise en œuvre de cette convention, tout particulièrement en matière d'identification du périmètre des données et de prise de responsabilité dans la diffusion (ou non) de la donnée, ont infléchi sensiblement le contexte. Deux changements importants viennent repositionner les enjeux relatifs aux échanges de données : le RGDP d'une part, la feuille de route stratégique ministérielle d'autre part.

Dans ce rapport, le terme de « donnée scolaire » sera utilisé comme un terme générique, dans la continuité du rapport de l'inspection générale de 2018³. Si ce terme ne répond pas à une définition réglementaire précise, il permet d'englober les différents sujets relatifs à la réalité et au contexte de l'école. Ainsi, la notion de donnée scolaire peut se définir comme tout assemblage d'informations susceptible de venir éclairer et guider le fonctionnement du système éducatif. Il incorpore la notion de statistique dans le sens où cette dernière répond à une organisation définie, cadrée et régulière de la donnée scolaire. Il peut être valorisé en interne au MENJ tout comme dans les collaborations avec des acteurs tiers. Recourir au terme générique de « donnée scolaire » est aussi le moyen de se démarquer du MESR et de sa problématique propre dans la production, l'usage et les échanges de données (cf. partie 2 du rapport) pour poser l'analyse dans le champ d'action du MENJ et englober des préoccupations spécifiques et partagées avec les collectivités territoriales.

Les observations menées par la mission s'inscrivent ainsi en cohérence de ce rapport de 2018 qui pointait déjà le début d'un usage des données à grande échelle au sein du système éducatif, le caractère artisanal et interpersonnel du traitement des données en académies, la grande variété des pratiques d'information, d'échanges et de formation des acteurs à un moment particulier, celui de la mise en place de la RGPD, et enfin les attentes exprimées par le terrain pour disposer d'une doctrine partagée et de recommandations plus explicites sur les pratiques à mobiliser avec les acteurs extérieurs, que ce soient les collectivités territoriales ou les acteurs privés (les « *Ed-Techs* »). La réalité observée sur le terrain témoigne d'une attente, toujours présente dans les services académiques, d'une doctrine nationale mieux partagée dans la gestion et l'usage de la donnée scolaire.

Le cœur des difficultés rencontrées pour établir une telle doctrine nationale réside dans l'appréciation pouvant être portée sur l'usage des données à caractère personnel et la responsabilité engagée par l'administration. Malgré les efforts réalisés pour développer une politique d'« *open data* » en capacité de permettre la mobilisation des données, des évolutions sont encore nécessaires pour répondre aux attentes des collectivités territoriales et les aider à élaborer et infléchir leurs stratégies, tout particulièrement en matière d'investissement dans les infrastructures scolaires mais aussi dans la gestion de celles-ci (cf. partie 3 du rapport).

L'état des lieux des situations académiques doit beaucoup au travail spécifique mené sur les territoires par l'intermédiaire des correspondants académiques de l'IGÉSR. L'enquête quasi-exhaustive passée en académie permet de conforter l'hypothèse d'une forte variabilité des relations entre services académiques et collectivités territoriales au niveau régional (cf. partie 3 du rapport). Elle conforte surtout la perception de la mission que les collectivités territoriales attendent moins du MENJ une avancée dans un conventionnement

1 Une copie de cette convention, impulsée par la directrice de l'époque, est annexée au présent rapport (annexe 2).

2 Cf. partie une du document.

3 Gilles Braun, Jean-Marc Merriaux, Jean Aristide Cavaillès, Jean-Marc Moullet, François Paquis, Stéphane Pellet (2018). Données numériques à caractère personnel au sein de l'éducation nationale (rapport n° 2018-016 février 2018). IGEN-IGAENR.

technique d'échanges de données qu'un progrès réel dans le processus de traitement de ces dernières. Si la donnée scolaire apparaît dans son ensemble disponible, la volonté d'une co-construction selon un usage raisonné et en fonction d'un objectif commun poursuivi entre services académiques et services territoriaux reste majeure. L'usage de la donnée scolaire est ainsi l'expression d'une attente de la part des collectivités d'être plus et mieux impliquées dans un projet éducatif régional partagé et symbole de la complémentarité des compétences avec les rectorats et régions académiques.

Pour mener à bien une telle ambition, il est nécessaire de se doter d'outils appropriés et plus homogènes (ou normés) dont on peut considérer qu'ils puissent être organisés voire pilotés au national (une sorte d'« *education data hub* » concentrant toutes les informations nécessaires à la construction de projets éducatifs régionaux) tout en respectant la diversité mentionnée des réalités régionales. L'état des lieux auquel il est utile de se livrer consiste aussi à établir les conditions par lesquelles l'exercice des responsabilités attendu des collectivités peut être pleinement opérant. En lien avec les compétences que l'État leur a conférées, bien maîtriser les responsabilités des collectivités requiert la possibilité de multiplier les données partagées pour les inscrire dans une politique éducative globale, qui de surcroît puisse valoriser l'ensemble des apports des collectivités (exemple des financements de crédits pédagogiques, ou encore des dépenses de transports d'élèves). À cet égard, un espace commun pour travailler de concert est un enjeu déjà inscrit dans la feuille de route ministérielle (action 13-22S1 : « Ouvrir une plateforme *Open Data* à l'échelon de chaque région académique en articulation avec la plateforme *Open Data* nationale »). Depuis 2018, la direction du numérique pour l'éducation (DNE) a développé une démarche « *d'APIsation* »⁴ ayant pour objectifs « *de valoriser les données du MENJS et d'identifier de nouveaux usages* ». La mission considère cependant que l'histoire du MENJ et son mode de régulation en matière de production et d'usage de la donnée doivent être infléchis pour dépasser une approche nationale descendante. Ainsi, on pourrait concilier un savoir-faire national dont l'expertise est centralement ancrée dans la DEPP et la DNE avec des apports académiques et régionaux spécifiques pour lesquels les SSA sont les acteurs essentiels et les collectivités territoriales des contributeurs effectifs.

Dans cette perspective, on ne rappellera jamais assez combien la qualité des collaborations engagées entre services académiques et collectivités territoriales dépend très fortement de la nature des informations qui sont recueillies pour lesquelles le chef d'établissement est un acteur déterminant. La formation des personnels de direction à la collecte d'informations, à la sensibilisation ou à l'acculturation de la qualité de la donnée, ainsi qu'à la capacité de mobiliser les bases de données, en particulier pour apprécier les parcours des élèves, constitue un enjeu substantiel des stratégies académiques et régionales. Ce qui est vrai au niveau de l'EPL est aussi au niveau des services académiques où les médiations en vue de collaborations méritent de s'inscrire dans un processus de gouvernance mieux normé et partagé.

Cette ambivalence, entre centralisation, déconcentration et décentralisation, est au cœur d'un usage utile de la donnée scolaire ; les apports conjoints des services académiques et des collectivités territoriales doivent tout à la fois servir la mesure de la situation nationale du système éducatif et rendre compte d'un projet éducatif régional tenant compte des spécificités territoriales (éducatives mais aussi économiques et sociales). La recherche d'une bonne maîtrise de ce curseur est au cœur des préoccupations et des préconisations émises dans ce rapport.

1. La donnée : une variable de plus en plus stratégique dans la mise en œuvre des politiques publiques

Dans le champ éducatif, notre histoire contemporaine est rythmée par des impulsions fortes et structurantes de l'intégration du numérique dans le quotidien de chaque citoyen ; si la crise sanitaire récente constitue un changement profond dans les enjeux et les usages de la donnée scolaire, il n'en demeure pas moins que la réflexion sur l'usage de la donnée publique peut être repérée, retracée et organisée en référence à la promulgation de trois grandes lois qui ont rythmé cette progression : la loi informatique et libertés de 1978, la loi pour une république numérique de 2016 et la loi relative à la protection des données personnelles qualifiée de loi RGPD de 2018.

⁴ Cf. note de bas de page n°7

1.1. Une politique nationale générale qui évolue rapidement et est marquée par une accélération des étapes structurantes

Le numérique, apparu progressivement dans la vie publique et privée, est à l'origine d'une véritable transformation sociétale, qui s'est progressivement installée et dont la centralité relève de la production, de la diffusion et de l'échange de l'information. L'information requiert une infrastructure de gestion de la donnée qui a conduit le pouvoir politique non seulement à inscrire la donnée et ses usages dans un cadre juridique national qui s'est progressivement structuré et renforcé mais aussi à concilier la réglementation avec les attendus de la législation européenne. Plusieurs étapes et faits marquants jalonnent l'histoire de cette transformation sociétale ainsi que la conciliation de cette évolution numérique entre avancées nationales et intégration des normes à l'international. S'agissant de la puissance publique, les ministères de l'éducation nationale et de la jeunesse ainsi que le ministère de l'enseignement supérieur et de la recherche sont aussi engagés dans ce processus de transformation, faisant notamment évoluer leurs pratiques organisationnelles avec, depuis la crise sanitaire, une importance accrue de l'usage de la donnée et de l'information à des fins pédagogiques.

1.1.1. Le cadre général d'exercice de la donnée scolaire : la loi informatique et libertés

En 1978, la CNIL, autorité administrative française indépendante, est créée dans le cadre de la loi 78-17 intitulée *Loi informatique et libertés*. Son rôle de régulateur lui permet d'agir auprès des professionnels, pour assurer la conformité à la loi dans l'utilisation des données personnelles, et également auprès des particuliers dans la maîtrise et l'exercice de leurs droits.

Cette première loi d'envergure va subir différentes modifications au cours de l'année 2004 afin notamment d'intégrer la législation européenne. C'est ainsi que :

- la loi pour la confiance dans l'économie numérique dite « LCEN » n° 2004-575 du 21 juin 2004, sur le droit de l'Internet, transpose la directive européenne 2000/31/CE du 8 juin 2000 et certaines dispositions de la directive du 12 juillet 2002 sur la protection de la vie privée dans le secteur des communications électroniques ;
- la loi n° 2004-801 du 6 août 2004 prend des dispositions pour intégrer complètement la directive 1995/46/CE du Parlement européen et du Conseil du 24 octobre 1995 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données.

Parallèlement à ces évolutions supranationales, la législation française va diligenter différentes missions pour élaborer les outils nécessaires au développement de la communication numérique et organiser une transformation de la gouvernance de la politique de la donnée publique pour la placer au niveau interministériel. Ainsi, le décret du 21 février 2011 marque la création d'une mission « Etalab » placée sous l'autorité du premier ministre chargée d'élaborer un portail unique, interministériel, destiné à rassembler et mettre à disposition librement, l'ensemble des informations publiques de l'État, de ses établissements publics administratifs et, si elles le souhaitent, des collectivités territoriales et des personnes de droit public ou de droit privé chargées d'une mission de service public. La mission Etalab coordonne l'action des administrations de l'État et apporte son appui aux établissements publics administratifs pour faciliter la réutilisation la plus large possible de leurs informations publiques.

De cette mission naîtra :

- le 5 décembre 2011 le portail www.data.gouv.fr avec ses déclinaisons thématiques ;
- le site internet api.gouv.fr qui référencera des applications avec habilitation par *datapass* ;
- le système code.gouv.fr pour la mise à disposition des codes sources.

En 2012, le secrétariat d'état à la prospective, à l'évaluation des politiques publiques et au développement de l'économie numérique, porté par le premier ministre, formalisera le Plan France numérique 2012 dont les objectifs s'articulent avec ces avancées logistiques et structurelles, essentielles pour engager de nouvelles réformes.

1.1.2. La loi pour la république numérique

Il faut noter que ces évolutions s'inscrivent dans un mouvement qui dépasse le contexte national comme en témoigne la définition en 2011 d'un partenariat international « pour un gouvernement ouvert » (PGO) que la France rejoint en 2014 et dont les 79 pays membres ont pour but de transformer l'action publique selon des principes de transparence, de collaboration et de redevabilité pour lesquels la donnée est un ingrédient essentiel. Lieu de partage de bonnes pratiques, le PGO offre une plateforme pour mettre en relation, développer et stimuler la communauté des réformateurs de l'État à travers le monde au travers de plans d'actions nationaux renouvelés tous les deux ans.

La loi pour une République numérique promulguée le 7 octobre 2016, aussi appelée « loi Lemaire », est le résultat de diverses consultations engagées depuis 2014 sur les enjeux du numérique. Cette loi ambitieuse entend encourager l'innovation et l'économie numérique, promouvoir une société numérique protectrice et garantir l'accès de tous au numérique. Cette loi a engendré la production de plus de 40 décrets et présente trois orientations majeures :

- une plus grande circulation des données et du savoir en obligeant l'ouverture des données publiques et d'intérêt général (économique, social, sanitaire ou environnemental), y compris les codes sources et les bases de données pour toutes les administrations d'État, les collectivités locales de plus de 3 500 habitants, les établissements publics et les organismes privés chargés d'un service public, à l'exception des entités de moins de 50 agents ou salariés. Les algorithmes publics, de plus en plus fréquemment utilisés dans les décisions administratives (par exemple pour l'affectation des élèves dans les établissements scolaires ou dans l'enseignement supérieur avec Parcoursup), sont aussi concernés par cette obligation réglementaire ;
- la création d'un service public de la donnée (décret du 14 mars 2017) piloté par la mission Etalab qui offre une infrastructure nationale autour de bases de référence dont le répertoire national des associations (RNA) ou le répertoire national des métiers et des emplois (ROME) ;
- le libre accès aux écrits de la recherche publique dans le respect de la propriété intellectuelle.

De la difficulté du recours à la notion de donnée d'intérêt général

La loi Lemaire a introduit la notion de donnée d'intérêt général comme un espace de partage entre données publiques et données privées, notamment relevant de celles détenues par les entreprises. L'exemple de l'apport utile de données privées à la sphère de la statistique publique était relevé et peut faire sens dans la gouvernance de la donnée scolaire (cf. le contexte des ENTs). Cette notion de partage avait pour objet de déceler les perspectives de nouvelles opportunités économiques et de services par cet entrelacement. Pour autant, cette notion ne dispose ni d'une définition juridique, ni d'une consistance normative. Le rapport Bothorel (partie 5) est plus récemment revenu sur cette appellation pour souligner la difficulté d'une définition consensuelle et stabilisée des données d'intérêt général et pour indiquer la difficulté de dépasser une approche au cas par cas pour établir un consensus et une doctrine de portée générale applicable à cette notion. Pour autant, les conclusions du rapport Bothorel visent à délimiter juridiquement cette notion de donnée d'intérêt général sans préjudice pour les acteurs privés et en encourageant la puissance publique à sécuriser un cadre juridique protecteur des agents privés et qui leur permette le partage volontaire de données d'intérêt général, y compris pour ce qui concerne l'utilisation des données à caractère personnel. Au final, cette notion reste difficile à utiliser et sa portée normative n'est pas simple à mobiliser.

Pour autant, ces évolutions appellent des précautions, et la diffusion de certains documents ne peut se faire que sous réserve d'anonymisation ou d'occultation des mentions touchant notamment à la vie privée et à l'existence de secrets protégés. Cette diffusion est exclue si elle engage la « *sécurité des systèmes d'information des administrations* ».

La Commission d'accès aux documents administratifs (CADA) et la Commission nationale informatique et libertés (CNIL), dont les pouvoirs ont été renforcés, assurent un rôle essentiel dans l'appréciation de ces règles.

Par courrier daté du 4 juin 2015, une mission a été confiée au Conseil général de l'économie, au Conseil d'État et à l'inspection générale des finances sur la notion de « données d'intérêt général » dans le cadre de la

préparation du projet de loi relatif au numérique. Cette notion apparue dans la stratégie numérique du gouvernement présentée par le premier ministre le 18 juin 2015 désigne des données, d'acteurs publics ou privés, dont l'ouverture présente un fort enjeu d'intérêt général car pouvant être à l'origine de nombreux bénéfices sur le plan économique et social.

Si les usages de la donnée lui confèrent une valeur particulière (matière première, levier d'opportunité ou actif stratégique), ses caractéristiques (bien non rival, non fini, coût marginal très faible, valeur liée à l'usage) rappelées dans le rapport ci-dessus évoqué prônent pour une ouverture des données au bénéfice des différents agents économiques dans le but de générer par leurs croisements de nouveaux usages. Il s'agit en d'autres termes d'en faire un bien commun⁵. Des initiatives de bases de données produites et alimentées par les citoyens comme « *OpenStreetMap* » ou « *Openfoodfacts* » illustrent une telle approche.

Ainsi, toute donnée peut potentiellement à un moment ou un autre prendre le statut de bien commun et la France en la matière se positionne en leader en matière d'« *open data* » dans le groupe composé du Canada de l'Australie et encore plus en pointe, le Royaume uni et les USA⁶.

1.1.3. Le règlement général sur la protection des données (RGPD) et ses incidences

Une étape supplémentaire est franchie en 2016 par l'introduction du règlement général sur la protection des données (JO L. 119 du 4 mai 2016, page 1) issu du règlement 2016/679 du Parlement européen et du conseil du 27 avril 2016, avec effet de ce règlement au 25 mai 2018. Dès le 20 juin 2018 ont été définies de nouvelles obligations concernant les traitements algorithmiques automatisés.

La prise en compte du RGPD renforce l'importance de l'enjeu d'une protection des données personnelles en réaffirmant autour de 5 principes majeurs les droits fondamentaux des citoyens européens et de leur droit au respect de la vie privée :

- **le principe de finalité** : les données à caractère personnel ne peuvent être recueillies et traitées que pour un usage déterminé et légitime ;
- **le principe de proportionnalité et de pertinence des données** : seules doivent être traitées les informations pertinentes et nécessaires au regard des objectifs poursuivis ;
- **le principe d'une durée de conservation des données limitée** : les informations ne peuvent être conservées de façon indéfinie dans les fichiers informatiques. Une durée de conservation précise doit être déterminée en fonction de la finalité de chaque fichier ;
- **le principe de sécurité et de confidentialité des données** : le responsable de traitement est astreint à une obligation de sécurité : il doit prendre les mesures nécessaires pour garantir la confidentialité des données et éviter leur divulgation à des tiers non autorisés et éviter la perte d'intégrité des données ;
- **le principe du respect des droits des personnes** : information des personnes, droits d'accès et de rectification, droit d'opposition.

Le respect du RGPD impacte toutes les organisations, y compris les collectivités territoriales, qui sont amenées à effectuer des opérations de collecte, d'enregistrement, d'organisation, de conservation, d'adaptation, d'extraction, de consultation, d'utilisation, de communication, de diffusion, de rapprochement, de verrouillage, d'effacement ou de destruction de données personnelles.

Le RGPD a induit des adaptations profondes dans les modalités de fonctionnement de l'information et, pour fluidifier la démarche, la CNIL et la CADA ont co-rédigé en 2019 le *Guide pratique de la publication en ligne des données publiques*, en partenariat avec les services d'Etalab2.

1.1.4. Le rapport Bothorel, la circulaire du Premier ministre du 27 avril 2021 et les feuilles de route ministérielles

Au niveau national, toutes ces évolutions ont guidé les termes de la stratégie gouvernementale de la modernisation de l'État dont la dernière étape a été définie dans la circulaire n°6264/SG du 27 avril 2021.

⁵ Philippe Aigrain (2005). Cause commune, l'information entre bien commun et propriété. Fayard, Paris.

⁶ Cf. *The « open data » economy- unlocking economic value by opening government and public dat*, Capgemini consulting, 2013, février 7.

Élaborée au niveau du secrétariat général du gouvernement, elle impose à chaque ministère la rédaction d'une feuille de route avec les objectifs spécifiques retenus et les échéances à appliquer avant la fin de l'année civile. Le MENJS et le MESRI ont produit les leurs en septembre 2021.

La structuration du cadre de production et de gestion publiques de la donnée est régulièrement éclairée par de nombreux rapports (rapport Villani en 2018, rapport Bothorel en 2020). Cette structuration se poursuit dans une dynamique qui rend bien compte des évolutions technologiques et de la prise en compte des besoins de chaque société. La crise sanitaire a été, de ce point de vue, un accélérateur tant au niveau national qu'europpéen.

La protection des données personnelles (extrait du Rapport Bothorel, page 61)

Les données personnelles, définies à l'article 4 du règlement général pour la protection des données (RGPD) comme « *toute information se rapportant à une personne physique identifiée ou identifiable* », font l'objet d'une protection spécifique et ne sont pas concernées par l'ouverture par défaut et la libre réutilisation des données et codes sources publics. Dans le cas des données personnelles qui relèvent de la vie privée, ces données ne sont pas communicables à des tiers, et *a fortiori* non publiables (84). Dans le cas des données personnelles non couvertes par le secret de la vie privée, la publication en ligne est proscrite, sauf s'il est possible d'anonymiser le document administratif qui contient des données personnelles. L'anonymisation ne doit pas permettre d'identifier un individu dans la base, de relier entre eux des ensembles de données distincts concernant un même individu, et de déduire de l'information sur un individu (85). La diffusion de données personnelles peut avoir lieu sans anonymisation seulement dans trois cas :

- si une disposition législative ou réglementaire l'autorise (86). Par exemple, l'article L. 127-10 du code de l'environnement permet la diffusion de certaines informations sur la base du découpage parcellaire et des adresses des parcelles au sein de bases de données de référence (87) ;
- si le consentement des personnes concernées a été recueilli (88) ;
- si les documents contenant des données personnelles sont listés à l'article D. 213-1-3 du CRPA et peuvent être rendus publics sans avoir été au préalable anonymisés (par exemple les permis de construire).

De plus, la réutilisation des informations contenues dans les données personnelles est encadrée. Cette réutilisation doit en particulier être licite, c'est-à-dire que toute réutilisation doit être fondée sur l'article 6-1 du RGPD qui exige le consentement des personnes pour un ou plusieurs traitements aux finalités définies et l'existence de fins légitimes poursuivies par le responsable du traitement. De plus, le responsable du traitement doit assurer la sécurisation des données.

84 Article L. 311-6 du code des relations entre le public et l'administration.

85 page 14 du Guide pratique de la publication en ligne et de la réutilisation des données publiques, CNIL et CADA.

86 Article L. 312-1-2 du code des relations entre le public et l'administration.

87 C'est sur ce principe qu'a pu être diffusée la base de données DVF des demandes de valeurs foncières.

88 Articles 4 et 7 du règlement général pour la protection des données. Mission Bothorel – pour une politique publique de la donnée 62. Par ailleurs, l'article 9 du RGPD énumère les données personnelles ayant un caractère particulier, à savoir celles qui révèlent l'origine raciale ou ethnique, les opinions politiques, les convictions religieuses ou philosophiques ou l'appartenance syndicale, ainsi que le traitement des données génétiques, des données biométriques aux fins d'identifier une personne physique de manière unique, des données concernant la santé ou des données concernant la vie sexuelle ou l'orientation sexuelle d'une personne physique. Le traitement de ces données est par principe interdit, sauf si cela entre dans les finalités listées par le même article. En particulier, en matière de santé, le traitement des données peut être autorisé s'il est nécessaire aux fins de la médecine préventive ou de la médecine du travail, de l'appréciation de la capacité de travail du travailleur, de diagnostics médicaux, de la prise en charge sanitaire ou sociale, ou de la gestion des systèmes et des services de soins de santé ou de protection sociale (...).

Le rapport Bothorel a engendré des ajustements significatifs dans les conditions d'ouverture et d'usage de la donnée publique et des codes sources. Il a notamment permis d'enrichir la réflexion sur la notion de donnée d'intérêt général introduite par la loi numérique et sur son usage. Il a, ce faisant, établi un état des lieux contemporain permettant de clarifier l'emploi de la donnée publique en problématisant l'équilibre à trouver

entre un principe d'ouverture qui est dans le sens de l'histoire et la nécessité de maintenir *a minima* un encadrement de son usage.

Outre le fait de souligner l'importance croissante de l'ouverture des données, le rapport Bothorel appelait de ses vœux un portage politique et administratif renforcé en faisant de l'ouverture et de l'usage des données une priorité gouvernementale pour laquelle la DINUM (Direction interministérielle du numérique) devait être responsable de la mise en œuvre et la DITP (Direction interministérielle de la transformation publique) en assurer le suivi de l'exécution. Parmi les recommandations émises figuraient l'édiction d'une circulaire interministérielle explicitant les principes de gouvernance, de missions et de responsabilité pour chaque administration ainsi que l'élargissement des fonctions des administrateurs ministériels de la donnée dans chaque ministère (passage de l'administrateur ministériel des données - AMD - à l'administrateur ministériel des données, des algorithmes et des codes sources - AMDAC).

Il est possible de présenter le décret du 27 avril 2021 du premier ministre comme une suite du rapport Bothorel dans la mesure où il est l'expression d'une politique volontariste pour établir les fondements d'une politique publique ambitieuse de la donnée, des algorithmes et des codes sources. Innovation dans les usages et évaluation des politiques publiques sont les deux piliers de l'ambition affichée. Celles-ci s'incarnent très concrètement dans la production de feuilles de route ministérielles, dans la nomination des AMDAC, et dans une attente de renforcement des échanges avec les collectivités territoriales par l'intermédiaire de l'agence nationale de la cohésion des territoires.

La mise en place de ces attentes a été réalisée dans le courant de l'année 2021. Elle a donné lieu à un plan d'actions pour chaque ministère, rendu public, conformément aux attentes exprimées par le premier ministre (cf. annexe 1).

1.1.5. Une nouvelle norme européenne à prendre en compte

Le dernier épisode de cet encadrement juridique de la donnée est fourni par la production du règlement du parlement européen et du conseil en date du 30 mai 2022 (UE) 2022/868 relatif à la gouvernance européenne des données et modifiant le règlement (UE) 2018/1724. Ce « *Data Government Act* » (DGA) européen vient renforcer l'approche européenne de l'« *open data* » et complète les règles du marché numérique et les outils disponibles en vue de faciliter l'accès aux données et leur réutilisation, notamment par le secteur public. Ce texte, applicable à compter du 24 septembre 2023, concerne les données détenues par des organismes du secteur public, protégées pour des motifs de confidentialité commerciale, y compris le secret d'affaires, le secret professionnel, le secret d'entreprise, le secret statistique, la protection des droits de propriété intellectuelle de tiers ou de protection des données à caractère personnel, dans la mesure où de telles données ne relèvent pas du champ d'application de la directive (UE) 2019/1024. Ce règlement encadre les conditions par lesquelles l'accès et l'usage des données détenues par des organismes publics pourront être organisés et régulés. Il met notamment l'accent sur la préservation de la protection des données personnelles par voie d'anonymisation, et sur l'accès aux données et leur réutilisation dans un environnement de traitement sécurisé qui est fourni ou contrôlé par l'organisme du secteur public, si possible à distance ou dans les locaux dédiés, si l'accès à distance pose difficulté ou préjudice (de sécurité par exemple).

L'ensemble de ces évolutions réglementaires témoigne en premier lieu du fait que la donnée et la démultiplication de son usage sont des sujets ouverts et dont la régulation se traite à l'échelle européenne avant d'être transposées dans les réglementations nationales des pays membres. Outre la complexification de ces dispositions réglementaires, on notera aussi la prise de conscience d'un fragile équilibre lié à la mobilisation de données à grande échelle (y compris des données à caractère individuel) porteuses des avancées d'une économie numérique et la responsabilité publique de garantir la protection nécessaire de chaque individu.

Dans ce contexte, il reste à apprécier comment la puissance publique peut s'adapter et assurer cet équilibre dans le respect de valeurs éthiques et des principes précédemment énoncés. Le récent colloque « *avenirs, innovations, révolutions* » (AIR 2021) réalisé par la CNIL nourrit cette réflexion et apporte des réponses sur la problématique suivante « *entre partage et protection, quelle éthique pour l'ouverture des données ?* ». Au vu des progrès continus des possibilités technologiques et d'usage des données, l'action publique est toujours en recherche d'adaptation pour fournir un équilibre adéquat entre l'exigence de la protection des données et celle de leur libre disponibilité. Ce qui vaut pour les données publiques en général vaut tout

particulièrement pour la donnée scolaire, de par la sensibilité relative aux données concernant des individus mineurs.

1.2. Attributs et spécificités de la donnée : acteurs, outils et gouvernance (protection de l'utilisateur, sécurité d'usage, articulation finalité / destinataire)

Le paysage de l'encadrement de l'action publique en matière de production et d'usage des données est comme précédemment souligné très largement dépendant de décisions supranationales (européennes), relayées dans le contexte national par une régulation interministérielle, sous la houlette du Premier ministre.

Ainsi, dans la sphère de l'action publique, la mise en place de ce cadre stratégique nécessite pour être utile, une gouvernance claire, accompagnée d'une comitologie opérationnelle, précise dans ses échéances, et d'un réseau d'acteurs explicite et garant de la mise en application de ce cadre stratégique à différentes échelles territoriales.

Si la CNIL et la CADA sont toujours restées des institutions satellitaires de référence, la DINUM est devenue en 2019 le chef de file de ce grand projet de transformation numérique par le décret n° 2019-1088 du 25 octobre 2019. La DINUM se substitue ainsi à la Direction interministérielle du numérique et du système d'information et de communication de l'État (DINSIC). Son directeur devient de plein droit l'administrateur général de la donnée, des algorithmes et des codes sources (AGDAC).

Il est ici utile d'opérer une distinction entre l'utilisation des données de gestion qui permettent au quotidien le suivi de la situation statistique du système éducatif, des données que l'on peut qualifier de publiques au sens de leur disponibilité au plus grand nombre, par l'intermédiaire des jeux de données mis en « *open data* » notamment sur le site interministériel « data.gouv.fr ». Si les jeux de données disponibles en « *open data* » offrent une qualification de la donnée scolaire déjà importante, ils ne répondent pas totalement à l'attente des collectivités qui, elles, recherchent soit une précision (granularité) de l'information qui n'est pas accessible par l'« *open data* », soit une disponibilité dans le temps (une immédiateté dans la valeur de la donnée) et une projection pluriannuelle de celle-ci qui soit la plus fiable possible, ce dont la fréquence d'alimentation des jeux d'« *open data* » ne permet pas de couvrir. De plus, les données de gestion attendues par les collectivités relèvent souvent de données à caractère personnel qui ne sont, par définition, pas disponibles dans les jeux d'« *open data* ».

Ce paysage est aussi structuré par d'autres acteurs essentiels : la direction interministérielle de la transformation publique (DITP) qui a fait de la transformation numérique l'un de ses chantiers, l'Agence nationale pour la sécurité des systèmes d'information (ANSSI) qui assure la mission d'autorité nationale en matière de sécurité des systèmes d'information et est donc chargée d'une part de proposer les règles à appliquer pour la protection des systèmes d'information de l'État, d'autre part de vérifier l'application des mesures adoptées ; l'INSEE, enfin, qui apporte son expertise aux différents ministères grâce à sa filière professionnelle dédiée à la statistique et à la donnée.

Un comité interministériel de l'administration de la donnée (CIAD), rassemblant les administrateurs ministériels des données (AMDAC), et autant que de besoin, les référents placés auprès des préfets de région se réunit tous les semestres autour de l'AGDAC. Un tableau de suivi des ouvertures annoncées et effectives des jeux de données, API et codes sources est publié et tenu à jour sur une base trimestrielle dans data.gouv.fr.

Au niveau ministériel, chaque administrateur ministériel des données (AMDAC) est chargé de constituer et d'animer un réseau de correspondants au sein des directions de son périmètre ministériel. Il est notamment responsable du référencement et de la tenue à jour des organisations et données relatives à son périmètre ministériel sur les plateformes data.gouv.fr, api.gouv.fr et code.gouv.fr. Distinct du délégué à la protection des données (DPD) institué par le RGPD, il est un point de contact chargé de la mise en œuvre et de la promotion des actions ministérielles en matière de circulation de la donnée des algorithmes et des codes sources.

1.3. Les particularismes du champ de l'éducation en matière de gestion de la donnée

Si le périmètre de la mission s'inscrivait initialement de l'école à l'université, deux éléments ont conduit la mission à centrer son analyse sur la situation au sein du MENJ :

- le contraste de situation, voire l'inversion structurelle qui caractérisent les deux ministères concernés par la formation initiale : là où le MENJ dispose d'une culture de collecte, de production et de mise à disposition de la donnée aux tiers par des voies structurées et dont la DEPP est le pilier, l'enseignement supérieur (MESR) se situe dans un contexte beaucoup plus complexe, entre l'administration centrale, la direction générale de l'enseignement supérieur et de l'insertion professionnelle (DGESIP) et la direction générale de la recherche et de l'innovation (DGRI) et leur interface dédiée que constitue la sous-direction des systèmes d'information et des études statistiques (SIES), et les établissements juridiquement autonomes qui font remonter à leur convenance leurs informations dans des formats non nécessairement similaires ;
- pour l'enseignement supérieur, le travail de mise en cohérence de la donnée (vérification et harmonisation de la donnée) devient alors beaucoup plus difficile, prend davantage de temps et conduit souvent à des incompréhensions quant aux définitions de champs des données recueillies ainsi que des délais de mise à disposition des données qui peuvent parfois aller jusqu'à un an de décalage.

Dans la suite du rapport, l'analyse est donc centrée sur la situation du MENJ et sur l'intérêt des collectivités territoriales pour les données de ce ministère au regard de leurs prérogatives.

Pour autant, mobiliser la réflexion engagée dans l'enseignement supérieur est également utile pour apprécier et capitaliser sur des pratiques de construction et d'échanges de données. Si l'accent du rapport est centré sur l'enseignement scolaire, des incises méthodologiques tirées de la situation du MESRI pourront trouver une place judicieuse dans les développements ultérieurs du rapport.

La convention en date du 27 juillet 2015 pose la première étape d'une collaboration institutionnelle entre les collectivités représentées par l'ARF et à l'époque le MENESR par le biais de la DEPP. Elle répond à un besoin d'informations au moment où sont renforcées les compétences des régions en matière éducative notamment en matière de planification des formations, de gestion des lycées, et de maîtrise du cadre de la formation professionnelle et de l'apprentissage. Elle a pour objectif de normaliser les échanges d'informations en volume et en qualité de façon à permettre aux différents décideurs de mieux dialoguer aux fins de construire des politiques cohérentes et efficaces à partir de données sélectionnées documentées et communément reconnues. Elle est donc une base de contractualisation dans le respect des règles du secret statistique et de l'engagement exprès de non cession des données à des tiers, amenée à être déclinée dans les territoires par des conventions locales sur un volume, soit minimal d'échanges mais non limitatif, selon les besoins locaux. Dans ce premier acte, des précisions sont apportées sur la nature des données qui peuvent être transmises par chaque niveau de production, les usages qui peuvent en être faits et les personnes habilitées à recevoir ce matériel :

- par la DEPP : la BCN (banque centrale de nomenclature), le RAMSESE (répertoire académique et ministériel sur les établissements du système éducatif), l'ACCE (application de consultation et cartographie des établissements) sur inscription, la BCP (banque centrale de pilotage) en accès réservé, APAE (aide au pilotage et à l'auto évaluation des établissements) en accès réservé, IVAL (indicateurs d'évaluation des lycées), les prévisions d'effectifs ;
- par toutes les académies : l'annuaire restreint des personnels en accès réservé, le fichier individuel anonyme des élèves et des formations suivies du second degré à la date du constat de rentrée, le fichier de données locales sur les établissements publics du second degré, le fichier de comptage du nombre de boursiers, le fichier des comptes financiers, les MEF académiques, les résultats aux examens, les flux d'élèves, les effectifs, les capacités d'accueil ;
- par certaines académies selon des normes très strictes des données relatives à l'orientation ;
- par les régions : le fichier des ressources humaines et des données immobilières, le fichier des capacités d'accueil dans les formations sanitaires et sociales, le fichier de l'enquête détaillée sur les apprentis, le parc des logements de fonction et, selon les académies, les subventions apportées aux établissements, le parc numérique et les aides apportées aux lycéens.

Si l'idée de mise en cohérence reste première dans la construction de cette convention, celle d'interopérabilité, dans le sens d'une alimentation du MENESR en données émanant des collectivités, n'est que partielle, l'État restant garant de la cohérence des systèmes d'informations. Cette première étape engage d'ores et déjà à aller plus loin dans la démarche partenariale pour définir une véritable stratégie autour de l'usage de la donnée et penser outils de stockage, garantie sur les règles de protection inhérentes à la donnée et engagements de responsabilité des acteurs.

2. Un état des lieux de la gouvernance de la donnée scolaire

2.1. Le MENJ

2.1.1. Gouvernance et instances

Depuis 2015, le MENJ a mis en place un comité de pilotage ministériel « *open data* ». Suite aux premiers travaux menés par l'AMD, ce comité a élargi son périmètre à une représentation des académies et des opérateurs du MENJ en 2020. Ce comité de pilotage « *open data* » vient à l'appui du suivi de la politique interministérielle de la donnée prônée par les services du premier ministre et de la DINUM en particulier. En 2019, le ministère a installé un comité d'éthique pour les données de l'éducation. Rassemblant, au-delà des ressources internes au MENJ, des parlementaires, des chercheurs et des responsables d'association, ce comité réfléchit à l'usage des données de l'éducation dans l'optique de préserver un juste équilibre entre protection et valorisation de la donnée. Ce comité d'éthique de la donnée vient en appui d'observations et de recommandations stratégiques du fonctionnement numérique du ministère. La circulaire du premier ministre du 27 avril 2021 a obligé chaque ministère à réfléchir à une stratégie de la donnée incarnée dans une feuille de route ministérielle dotée notamment d'un plan d'actions phasé, ce qui a conduit à la mise en place d'un comité stratégique de la donnée.

Enfin, un comité des partenaires, émanation de la DNE et animé par l'AMDAC, finalise la liste des instances concernées par la gestion de la donnée scolaire. C'est cette dernière structure qui est en interface avec les collectivités et qui recueille les besoins, les attentes et propose des améliorations quant aux modes de fonctionnement actuels.

Dans cette gouvernance, l'AMDAC joue un rôle central. Il est tout d'abord un coordonnateur des directions d'administrations centrales qui, elles, s'impliquent spécifiquement dans des projets relevant de leur domaine propre. Les échanges concernant les traitements de données sont principalement menés au niveau national tout comme les orientations prises qui relèvent des positions des directions centrales. En revanche, le niveau opérationnel reste largement celui des académies (ou des régions académiques) et la mission a pu relever que la connaissance de ce qui se pratique sur le terrain reste faible et parcellaire au niveau central. L'exemple de la crise sanitaire est à ce titre significatif si l'on en juge par les difficultés rencontrées pour faire remonter au national des données coordonnées et harmonisées, au jour le jour. Il n'existe pas de consolidation au national de l'ensemble des collaborations, y compris formelles (conventions), passées entre les académies et les collectivités. L'AMDAC dispose de quelques éléments mais pas d'un recensement explicite qui reste encore un objectif à atteindre.

Ainsi, même si l'infrastructure de pilotage de la donnée est bien structurée, les interfaces entre ces structures de coordination auraient peut-être intérêt à être mieux précisées, voire simplifiées. Mais la question première est : quelle gouvernance pour mener à bien la sécurisation technique et juridique précédemment posée comme nécessaire et répondre concrètement aux besoins exprimés infra dans les territoires ?

Il serait tout d'abord souhaitable de privilégier une vision stratégique de gouvernance partagée entre le MENJ et les collectivités territoriales. Le scénario d'une plateforme des données éducatives englobant une partie des données de l'enseignement supérieur (les données Parcoursup en particulier), qui par ailleurs donnerait à voir une vision commune des parcours et des destins scolaires, présente aussi outre l'avantage du lieu unique de stockage, la possibilité de bien distinguer entre données statistiques (soumises au secret statistique et non facilement réutilisables) et données de gestion mobilisables à des fins de pilotage. Cette stratégie plus claire de positionnement entre les données qui relèvent de la production statistique de celles extraites à des fins de pilotage (et non couvertes par le secret statistique) mérite d'être partagée entre MENJ et MESR.

La situation du MESR

Tout comme le MENJ, le MESRI a répondu aux attentes de la circulaire du 21 juin 2021 en produisant, lui aussi, sa feuille de route ministérielle. Au national, les bases de données majeures du MESR d'un intérêt majeur pour les collectivités territoriales sont SISE et Parcoursup ; elles sont suivies et valorisées par le SIES.

Ces données donnent lieu à valorisation sous la forme d'indicateurs de pilotage (*data* ESR) et d'extractions à façon permettant de dégager des indicateurs territoriaux (régionaux) qui forment le commun dénominateur quantitatif des documents STRATER. Les lignes directrices de la feuille de route ministérielle sont en conséquence plus centrées sur des objets à promouvoir tels que *data* ESR, les données Parcoursup, ou encore, dans le champ de la recherche, la promotion de la problématique de la science ouverte via l'essor d'une approche « *FAIR by design* » (*FAIR* : facile à trouver, accessible, interopérable et réutilisable).

Les échanges et entretiens réalisés témoignent là encore de relations locales parfois établies entre collectivités territoriales et établissements d'enseignement supérieur pour pallier ce qui est considéré comme un manque réel pour piloter en territoire et prendre des décisions particulièrement étayées. Le MESR est ainsi un environnement où des médiations locales ponctuelles, spécifiques et variées se construisent au cas par cas en matière de production et d'utilisation de la donnée, sans que cette situation ne fasse l'objet ni de formalisation connue au national, ni de recensement consolidé de ces relations. Ainsi, que ce soit MENJ ou MESR, la gestion des données témoigne de l'existence de relations locales variées mais ce qui les diffère, c'est la facilité de consolidation nationale pour l'un et la difficulté rencontrée par l'autre. Au final, dans les deux cas, les acteurs s'arrangent au local sans que ces interactions ne soient véritablement connues et maîtrisées par le niveau central.

Parcoursup constitue une exception à cette observation dans la mesure où, par définition, c'est une collecte de données centralisée (recensement des vœux des futurs étudiants d'une part, recensement du processus de décision des établissements d'accueil d'autre part).

Ainsi, disposer d'une stratégie et d'une gouvernance partagées des conditions de réutilisation des données pour assurer une continuité des données en matière de suivi de cohortes par exemple, et en permettre un usage facilité à des acteurs tiers (dont les collectivités), y compris dans le parcours des jeunes dans l'enseignement supérieur, sont ainsi une nécessité au vu des ambitions numériques de la nation ; pour autant, les pratiques de collecte et d'usage de la donnée ainsi que les cultures professionnelles ne sont, malgré l'apparente évidence, pas toujours instantanément convergentes et posent des problèmes de coordination temporelle dans la disponibilité de la donnée tout particulièrement.

La base SISE par exemple, réalisée par le SIES à partir de certaines remontées de données nécessite la levée du secret statistique dès lors qu'une utilisation étendue au-delà du SIES est envisagée, ce qui rend plus compliquée la disponibilité de la donnée pour des bénéficiaires extérieurs.

L'exemple de la base SISE

Au niveau national, les bases compilées pour établir la base SISE donnent lieu à valorisations sous la forme d'indicateurs de pilotage (*data* ESR) et d'extractions à façon permettant de dégager des indicateurs territoriaux (régionaux) qui forment le commun dénominateur quantitatif des documents STRATER.

Les modalités de collecte des bases de données du sup se font sur la base de l'agrégation d'une douzaine d'enquêtes qui est compilée à partir de dénominateurs communs sous la forme d'un tableau de synthèse incorporant parfois des données individuelles. Cette table de synthèse qui présente des données protégées ne peut donc être transmise à l'extérieur mais elle fait l'objet d'une double utilisation : elle sert à la production des données statistiques nationales portées par le SIES (consolidation) et elle est diffusée auprès des services statistiques académiques (SSA) pour qu'ils travaillent à partir de ces données à l'élaboration d'indicateurs adaptés aux particularismes de leurs territoires. Pour les données « établissement », l'avis de ce dernier est requis pour les mettre à disposition, ce qui ne peut se faire au niveau du SIES vis-à-vis des collectivités territoriales.

Dans ce contexte plus compliqué que celui du scolaire pour ce qui relève de la collecte et du traitement de la donnée, la valorisation faite par le MESR de cette consolidation des données trouve aussi néanmoins ses limites : L'enquête SISE reste un outil important pour récupérer de la donnée mais tous les sujets n'y sont pas traités. On reste démuné quand on a besoin de données qui lient par exemple les résultats obtenus dans le cadre des projets investissement avenir (PIA) et leur impact sur la réussite des étudiants.

Le statut d'enquête statistique conduit aussi à ce que les besoins nouveaux soient relativement difficiles à couvrir. Par exemple, la relation entre les apports des PIA et l'existence de nouveaux cursus de formation est

difficilement explorée et couverte. La production de données est importante pour les statistiques mais aussi pour éclairer au mieux le pilotage des structures. L'AMUE avec PEGASE ne semble pas apporter plus d'informations sur les étudiants.

Ces exemples montrent que, comme la mission l'a déjà évoqué, la question des conditions d'une gouvernance partagée entre État et collectivités prime sur la stricte question du « transfert de données ». Dit autrement, l'enjeu technique reste subordonné à des formes de gouvernance qui sont l'expression d'un consensus collectif dont les contours restent à dessiner. C'est le sens que la mission entend donner à d'éventuelles recommandations. Il s'agit ainsi, pour elle, moins de délimiter ou d'architecturer les termes d'une convention qui viendrait relayer le seul repère existant en la matière (convention DEPP de 2015) que de promouvoir les conditions d'une gouvernance permettant le recours à un scénario qui vise à mieux organiser et fluidifier les relations entre les services académiques et les collectivités territoriales.

2.1.2. La feuille de route et la stratégie définie

En conséquence de cette circulaire du 27 juin 2021, le MENJ a produit une feuille de route ministérielle qui qualifie son ambition et ses objectifs en matière de transformation numérique. Améliorer l'efficacité, faciliter la transparence, proposer des ressources pour l'innovation sont au cœur de l'ambition affichée.

Deux axes majeurs structurent cette feuille de route ministérielle : promouvoir le développement et l'utilisation des interfaces de programmation d'application (API)⁷ d'une part, installer une plateforme de données scolaires pour alimenter l'ensemble des acteurs de l'écosystème scolaire et promouvoir en conséquence de multiples formes d'innovation dans le champ de l'éducation d'autre part.

S'agissant de l'essor des API du MENJ, deux catégories sont envisageables : des APIs spécifiques (*ad hoc*) à destination des acteurs publics et des collectivités territoriales en particulier, des APIs ouvertes en direction du grand public sous la forme d'un large catalogue venant répondre aux multiples attentes de ce dernier. Dans les deux cas, la DNE apparaît comme le bras opérateur d'un tel objectif.

S'agissant de l'installation d'une plateforme de données scolaires, deux projets distincts rentrent dans le champ d'une telle ambition :

- un premier projet (« Innovations, données et expérimentations en éducation » - IDEE) où la DEPP est partenaire de cet équipement d'excellence (Innovations, données et expérimentation - projet IDEE) porté par l'université de Paris sciences et lettres (PSL) ; cet Equipex, qui vise à promouvoir le lien entre chercheurs et acteurs de l'éducation par le développement d'expérimentations contrôlées, reste centré sur une importation et la promotion d'une recherche de nature translationnelle dans le champ de l'éducation. En ce sens, il n'épuise pas le sujet des interfaces de données entre le ministère et les acteurs publics, dont les collectivités territoriales ;
- le second projet (« Education Data Hub » - EDH) doit permettre à l'ensemble de la communauté éducative et ses partenaires, au rang desquels les chercheurs mais aussi les collectivités territoriales, de pouvoir disposer d'une structuration bien organisée des données en matière d'éducation afin de créer un écosystème organisé et cohérent, le tout guidé par une éthique commune et partagée de l'utilisation de ce patrimoine de données. En partenariat avec l'INRIA, ce projet poursuit un objectif de lisibilité et, par conséquent, cherche à pallier le constat d'un émiettement et de fragmentation de la donnée en éducation, ainsi qu'une volonté d'augmenter la qualité des données mises à disposition de l'ensemble de la communauté des acteurs du service public de l'Éducation actuellement dotées d'une faible interopérabilité. Après un travail exploratoire et de préfiguration, ce projet est actuellement en cours de finalisation dans sa définition.

La présentation de ces deux projets issus du MENJ montre de manière explicite qu'il n'est pas encore acquis pour ce ministère d'associer les collectivités dans une réflexion et démarche conjointes autour de la donnée

⁷ Une API (*application programming interface* ou « interface de programmation d'application ») est une interface logicielle qui permet de « connecter » un logiciel ou un service à un autre logiciel ou service afin d'échanger des données et des fonctionnalités. Les API offrent de nombreuses possibilités, comme la portabilité des données, la mise en place de campagnes de courriels publicitaires, des programmes d'affiliation, l'intégration de fonctionnalités d'un site sur un autre ou l'« open data ». Elles peuvent être gratuites ou payantes (glossaire de la CNIL, <https://www.cnil.fr/fr/definition/interface-de-programmation-dapplication-api>)

scolaire. Il n'y a pas de contribution ou de participation attendues de celles-ci, elles restent tout au plus des destinataires des APIs nouvelles disponibles.

2.2. Une administration centrale intervenant en fonction du champ de responsabilité de chacune de ses composantes

Les feuilles de route numérique de ces deux ministères expriment dès lors des contrastes marqués avec, au fond, l'existence de deux modèles distincts pour ne pas dire « alternatifs » d'échanges de données avec les collectivités territoriales : modèle centralisé et consolidé de type *Education Data Hub* vs modèle déconcentré mobilisant le potentiel régional des services académiques parmi lesquels les SSA jouent un rôle de premier rang. Dans les deux cas, le constat de situation montre que les données placées en « *open data* » sont peu opérantes au vu des attentes exprimées par les collectivités qui affirment leur intérêt de partager « en temps réel » des données de gestion avec les services académiques.

Pour des raisons de temps et de recherche d'efficacité, la mission a fait le choix de centrer l'analyse sur la situation du MENJ et sur les perspectives possibles de par l'essor constant du numérique et des données produites, source de multiples innovations portées par des acteurs directs et indirects du champ éducatif. Dans cette perspective, ci-après, la mission propose de recenser et d'analyser au niveau du MENJ chacune des directions concernées par la gestion de la donnée.

2.2.1. La DEPP, référent historique de la donnée scolaire

Le rôle de la direction de l'évaluation de la prospective et de la performance (DEPP) en matière de collecte, de production et de mise à disposition de la donnée en éducation est central au sein du MENJ. Par définition, cette direction assure le rôle de service statistique ministériel de l'éducation pour le compte de l'autorité de la statistique publique qui garantit qualité et indépendance dans l'élaboration et la diffusion de la donnée publique à des fins statistiques ; elle est aussi un appui à la décision pour le ministre, ce qui lui confère en l'état une réelle ambivalence : garantie d'indépendance dans la production statistique et appui d'aide à la décision pour les attentes du ministre. À ce titre, de multiples études sur les diverses facettes du fonctionnement ministériel (élèves, personnels, structures et territoires) sont livrées chaque année. De même, la DEPP joue un rôle important dans l'aide au pilotage du système éducatif au travers de la production d'outils tels que les bases d'aide au pilotage et à l'auto-évaluation des établissements (APAE premier et second degrés) et l'ensemble des dispositifs de mesure des acquis des élèves qui forment un outil de positionnement et de suivi multicritères des élèves et des structures scolaires.

La Base centrale de pilotage (extrait des différentes présentations en ligne de la BCP)

La Base centrale de pilotage est depuis 30 ans l'entrepôt de données de référence du système éducatif français. Elle est le fruit d'une collecte de statistiques exhaustives, stockées en vue de mettre à disposition un véritable outil d'aide à la décision.

La direction de l'évaluation, de la prospective et de la performance (DEPP) procède à la validation et à la classification thématique des statistiques qui entrent dans la base; elle en assure l'alimentation et l'enrichissement.

La BCP constitue un outil moderne de pilotage à différents échelons, répondant aux besoins d'information et d'aide à la décision. Il permet à tous les décideurs du système éducatif de partager une même information, d'utiliser un système de références commun ; d'avoir une meilleure visibilité dans l'élaboration des projets éducatifs de disposer de données traitées et vérifiées ainsi que d'éléments très précis de comparaison.

Business Objects® est l'outil de consultation de la BCP et celle-ci a la vocation de répondre à un large panel d'utilisateurs en leur offrant des références partagées. Son contenu doit donc satisfaire à des critères d'excellence et de qualité. Pour cela, elle met à disposition des données historicisées, exhaustives, contrôlées et homogénéisées, détaillées ou agrégées. Cette fiabilité des données en assure la comparabilité.

Enfin si la DEPP relève du périmètre « service statistique ministériel » (SSM), ce n'est pas le cas des services statistiques académiques qui sont pourtant les relais indispensables de cette direction. À cet égard, la DEPP en assure toujours l'animation, notamment au travers des journées nationales SSA. Cette remarque est d'importance au sens où la variété des situations académiques dans leurs relations aux collectivités territoriales sera soulignée infra.

Ces éléments posés, il est important de rappeler que, si l'intérêt de la production statistique est l'obligation faite au propriétaire de la donnée de la fournir, la limite réside dans l'interdiction de sa réutilisation. En d'autres termes, il s'agit là de poser le principe de la responsabilité qui reste le nœud du problème dans le partage des données avec les collectivités territoriales.

2.2.2. La DGESCO comme maître d'ouvrage de la politique du numérique éducatif au travers des remontées d'information nécessaires au pilotage du MENJ

La DGESCO constitue la « tour de contrôle » de la politique numérique du MENJ et les trois services qui forment l'ossature de cette direction sont directement concernées par sa mise en œuvre et son évolution.

La politique de l'« *open data* » permet de disposer d'un certain nombre de données ouvertes qui explicitent le fonctionnement général du système éducatif, elle satisfait le plus grand nombre, mais ne répond pas à la problématique des échanges de données avec des acteurs tiers en lien avec la transmission de données à caractère personnel visant à assurer le respect du code de l'éducation.

Les échanges de données avec les acteurs extérieurs au système éducatif concernent essentiellement l'utilisation des traitements réalisés par la DEPP pour le premier et le second degré en direction des communes (et des maires) pour ce qui concerne le contrôle du respect de l'obligation scolaire (ONDE pour le premier degré, SCOLARITÉ / SIÈCLE pour le second degré).

Au-delà des maires, ces échanges de données se font en direction des collectivités territoriales (départements et régions) pour ce qui relève des problématiques de la restauration scolaire, de l'identification des élèves boursiers, de l'organisation du transport scolaire, de l'organisation de la sectorisation et de la mixité sociale, ou encore de l'identification des élèves en situation de décrochage scolaire. Se rajoutent aussi les échanges avec les MDPH pour les élèves en situation de handicap ainsi qu'avec les PMI dans le premier degré pour la problématique de la santé des élèves (cf. l'explicitation proposée en 2.5.2).

En tant que responsable des traitements, la DGESCO est vigilante quant à la diffusion des données à caractère personnel, considérant que les conditions de réutilisation de ces données par les destinataires doivent être éclairées et bien encadrées. En ce sens, la DGESCO appelle de ses vœux une clarification de cette problématique de l'usage des données à caractère personnel par l'installation d'une sorte de cadre national de transmission de ces données, qui permettrait des pratiques plus harmonisées en la matière, entre les différentes directions du MENJ⁸.

Une telle avancée conforte l'idée d'une plateforme nationale des données scolaires dont le maître d'ouvrage serait le ministère qui poserait le cadre, identifierait les bénéficiaires, et délimiterait les contours des données concernées pour une mise en œuvre territoriale (régionale) en capacité de répondre à des attentes parfois différentes puisque dépendantes des contextes économiques, sociaux et structurels.

2.2.3. La DNE, à l'interface des directions et en appui technique des acteurs

La direction du numérique éducatif est le deuxième acteur majeur concerné par les échanges de données internes et externes au MENJ dans la mesure où la grande majorité de son activité est d'intervenir comme maître d'œuvre et en soutien des attentes de la DGESCO. La DNE est ainsi à l'interface de multiples acteurs, au cœur d'un écosystème complexe et foisonnant : la DGESCO mais aussi les autres directions métiers de l'administration centrale du MENJ, les académies elles-mêmes, les opérateurs du ministère, les collectivités territoriales et des acteurs privés relevant de ce que l'on appelle l'environnement des « *Ed-techs* ».

L'objectif premier de la DNE est de maîtriser au mieux les relations avec cet écosystème pour le rendre plus transparent et homogène pour l'utilisateur (mot de passe unifié, navigation simplifiée, évitement de saisies redondances de données, etc.). Les relations que la DNE construit avec les collectivités au niveau régional le sont par l'intermédiaire des DANE et des DSI.

Dans ce contexte, il faut observer que le paysage n'est pas encore suffisamment harmonieux pour l'utilisateur et, face à ce constat, les enjeux premiers sont la bonne maîtrise de la sécurité et de l'interopérabilité de la

⁸ Cf. préconisation 2 : élaborer et installer une doctrine nationale partagée de l'usage de la donnée scolaire permettant une plus grande homogénéité des comportements et des procédures en académies.

donnée, et la possibilité de développer une expérience fluide pour l'utilisateur, indépendamment de l'acteur qui offre le service pratiqué.

C'est en ce sens qu'il faut comprendre la feuille de route ministérielle et les différentes actions qui la composent. L'enjeu est en premier lieu de s'entendre sur une définition de l'urbanisation fonctionnelle du numérique éducatif. Cela requiert une cartographie fine par blocs fonctionnels, et la bonne identification des responsables de la donnée. Par exemple, le poids et la responsabilité des chefs d'établissements dans la production des données sont importants, en particulier en matière de gestion des données relatives à la vie scolaire. La DNE vise à fournir de bons repères, à proposer un cadre de référence technique approprié (en matière de sécurisation et des conditions de disponibilité des données, notamment). De façon plus générale, tout comme la DGESCO, la DNE appelle de ses vœux la fixation d'un cadre mieux établi des échanges de données qui puisse constituer un référentiel partagé par tous⁹.

De plus, l'État doit aussi avoir pour objectif d'apporter des « briques socles » à cet écosystème coordonné, qui viennent offrir des réponses partagées à des enjeux fonctionnels et techniques majeurs. Maîtriser la variété des ENTs par exemple, leur permettre un égal accès aux annuaires, produire des référentiels qui permettent une messagerie fluide et facilite le fonctionnement des activités de vie scolaire, est important. C'est le sens donné à l'accessibilité par *Educonnect*, qui simplifie l'accessibilité et le mot de passe. L'objectif affiché est notamment que tous les ENTs des lycées soient « éduconnectés » d'ici la fin de l'année. C'est aussi dans cette finalité qu'il faut comprendre l'installation du GAR (gestionnaire d'accès aux ressources) et des services qu'il véhicule, notamment par l'intermédiaire d'un catalogue des APIs disponibles.

L'administration centrale du MENJS se doit de fournir une base de services communs numériques, notamment pour que chaque établissement scolaire (école et EPLE) puisse accéder à une gamme de services et d'applications partagés et disponibles au national. Des progrès restent à mener en la matière. L'exemple de l'instauration de classes virtuelles, en conséquence de la crise sanitaire, en témoigne. Chaque ENT, donc chaque région, a déployé ses propres outils en parallèle là où le ministère s'est appuyé sur son opérateur, le CNED, pour proposer son propre outil (les classes virtuelles CNED), avec plus ou moins de soutenabilité pour une utilisation à grande échelle. Il en est de même pour des outils d'interaction numérique où la DNE promeut BBB (*Big Blue Bottom*) là où les acteurs mobilisent de multiples autres solutions techniques.

Pour la DNE, la cohérence globale du paysage numérique devra passer par la loi pour instituer un cadre d'interopérabilité partagé des données et reconnaître une plateforme interopérable dans le champ de l'éducation en voisinage de ce qui a été mis en place dans le champ de la santé (*Health Data Hub*). Cela devrait permettre de faciliter l'interopérabilité et le partage de formats et de protocoles de services. Il est important de rendre transparent l'accès à des briques aussi variées que les listes d'élèves, les listes des responsables d'établissements, les emplois du temps, les cahiers de textes des élèves, en soulignant que ces évolutions sont aussi parfois source d'inquiétude des enseignants tout autant que des parents, inquiétude renforcée par la diversité des situations observables d'une académie à l'autre.

En référence à la feuille de route ministérielle, la DNE considère la gouvernance de la donnée, interne au MENJ, comme claire et bien établie ; ce sont les interfaces avec les partenaires externes qu'il faut continuer de travailler, les conventions constituant en fait une phase transitoire en attendant la disponibilité d'une plateforme dédiée à l'éducation et répondant à un cadrage national mieux établi.

S'agissant de la réalité présente, il est clair que le comité des partenaires porté par la DNE reste la brique centrale de gouvernance par laquelle le point de vue et les attentes des collectivités territoriales en matière de gestion des données sont véhiculées. Pour autant, outre l'absence de périmètre défini explicitement pour cette instance, le niveau de représentation des interlocuteurs reste variable et n'apparaît pas nécessairement opérationnel car les informations qui y sont émises ont du mal à redescendre vers l'ensemble des territoires ; l'instabilité de la représentation ne crée pas non plus continuité de dialogue et ce comité des partenaires est davantage perçu comme une chambre d'enregistrement que comme un lieu de co-construction.

⁹ Cf. préconisation 2 : élaborer et installer une doctrine nationale partagée de l'usage de la donnée scolaire permettant une plus grande homogénéité des comportements et des procédures en académies.

2.2.4. La DJEPVA, nouvelle entité du MENJ, autre culture et autres enjeux dans la gestion de la donnée

Le rattachement de cette direction (jeunesse, éducation populaire et vie associative) au périmètre du MENJ et son évolution récente expliquent en partie qu'elle ne soit pas intégrée dans la feuille de route du ministère. La présentation de ses activités fait apparaître une double approche, métiers d'une part, partenariats d'autre part.

L'approche « métier » fait principalement référence au service national universel (SNU) qui a concerné 40 000 jeunes en 2022 et à l'accueil collectif des mineurs (ACM), qui concerne 1,8 millions d'enfants fréquentant des « centres aérés », en accueils périscolaires sans hébergement portés par les communes et dont une majorité s'appuie sur des conventions de services passées avec le milieu associatif, ainsi qu'un peu plus d'un million d'enfants bénéficiant de séjours en accueil collectifs avec hébergement (centres de vacances) durant les vacances scolaires. L'État exerce une fonction de contrôle de ces accueils, au plan départemental, et l'essor du numérique et de la gestion dématérialisée permettrait à l'administration de déployer une vision et une gestion plus systématique, et d'exercer, en conséquence, un contrôle plus efficace. Les données numériques actuelles restent en effet déclaratives et à ce stade, malgré la fonction statistique en cours de structuration au sein de l'INJEP, c'est bien la dimension « suivi de structures » qui est au cœur du fonctionnement métier de cette direction, ce qui explique le non recours à l'exploitation des données individuelles. En ce sens, contrairement à l'éducation nationale qui dispose de bases de données individuelles et de remontées d'informations bien établies avec le terrain, les publics de la DJEPVA ne sont pas des publics « bien identifiés » et demander des remontées de données individuelles pour l'ACM supposerait une réforme du cadre juridique des accueils collectifs de mineurs.

Malgré tout l'intérêt qu'il pourrait y avoir à interfacier des bases de données individuelles entre la DGESCO et la DJEPVA, l'identification (et la collecte) de données individuelles dans la logique métier de la DJEPVA devrait nécessairement passer par un changement législatif ce qui rend plus complexe, au moins dans la temporalité, la réalité d'une telle ambition.

Les relations entretenues par l'État avec les collectivités territoriales en matière de politiques de jeunesse relèvent d'une approche principalement partenariale. La DJEPVA s'appuie sur les réseaux déconcentrés académiques chargés de la jeunesse, de l'engagement et du sport (DRAJES en région académique, SDJES au plan départemental), créés depuis le 1^{er} janvier 2021. Ces services servent d'interfaces locales pour l'ensemble des politiques en faveur de la jeunesse et s'appuient sur des partenariats développés avec les collectivités locales et les partenaires associatifs, qui peuvent prendre la forme, pour ces dernières, de conventions. L'État doit cependant désormais composer, y compris dans ces domaines d'intervention non décentralisés avec les collectivités locales, et plus particulièrement avec les conseils régionaux qui sont les chefs de file sur la compétence des collectivités locales en matière de « politique de la jeunesse », notamment sur les sujets liés à l'information de la jeunesse, à la mobilité internationale et plus largement la promotion des activités de jeunesse et de l'engagement des jeunes¹⁰. En ce sens, le fonctionnement de la DJEPVA conforte au sein du ministère l'intérêt de considérer la région comme la bonne unité de compte pour avancer une réflexion partagée entre services académiques et collectivités.

En matière d'engagement des jeunes et de continuité éducative (entre les temps scolaire, périscolaire et extrascolaire) la DJEPVA fait cohabiter ses dispositifs « État » avec ceux relevant des collectivités. C'est par exemple le cas de la politique « vacances apprenantes » incluant le dispositif des « colo apprenantes », bâti sur un projet structuré, dont les opérateurs sont des associations et bénéficiant du concours des collectivités territoriales.

À ce stade, les politiques de jeunesse et d'engagement ne sont pas dans des logiques de guichet, de prestation et le recours à la donnée n'est pas de ce fait, un élément structurant.

Mais des avancées existent pour conforter l'impact de l'approche numérique dans la gestion des partenariats avec les associations. Un premier pas qui sert cette impulsion nouvelle et participe de la simplification

¹⁰ Alinéa 2 de l'article L. 1111-4 du code général des collectivités territoriales (modifié par l'article 104 de la loi n° 2015-991 du 7 août 2015 portant nouvelle organisation territoriale de la République) : « Les politiques publiques en faveur de la jeunesse menées par l'État, les régions, les départements, les communes et les collectivités à statut particulier peuvent faire l'objet d'un débat au sein de la conférence territoriale de l'action publique mentionnée à l'article L. 1111-9-1. Ce débat porte notamment sur l'articulation et la coordination de ces politiques entre les différents niveaux de collectivités et l'État ».

administrative attendue au travers du programme interministériel « Dites-le nous une fois », réside dans la valorisation de l'outil « Le Compte Asso » construit comme un portail unique pour identifier les bénéficiaires associatifs de subventions de l'État et en écho éviter les doublons d'aides vers les mêmes structures.

La DJEPVA a la volonté de structurer les deux approches, « métier » et partenariale, grâce aux outils digitaux. Face à ce contexte, la direction, faute de moyens humains et financiers, s'est tournée vers la DINUM pour créer, à l'instar du « *data sport - éducation* », une plateforme spécifique dénommée « *data asso* » avec l'aide d'une *start-up* d'État. Pour autant, à moyen terme, l'un des objectifs du rapprochement avec l'éducation nationale est de parvenir à mieux travailler dans ce domaine de la donnée au sein des services académiques. À ce titre, la mission constate que les DRAJES ont besoin de se rapprocher des SSA, qui eux-mêmes vivent ce rapprochement comme une source de travaux supplémentaires.

La mission considère que, du point de vue de la production et des interfaçages de la donnée, l'efficacité devrait conduire à faire fonctionner plus étroitement le triptyque INJEP - SIES - DEPP en créant davantage de connexions entre ces différents services sur des travaux conjoints pour disposer d'une meilleure capacité à renseigner la notion de « parcours de formation et d'engagement » des jeunes.

En conséquence à moyen terme, la mission pense qu'une évolution des services statistiques académiques (SSA) est souhaitable, pour qu'ils deviennent des têtes de réseaux de proximité incontournables, au carrefour de l'éducation nationale bien sûr, mais aussi de l'enseignement supérieur (devant l'autonomie des établissements du supérieur), des politiques de jeunesse et d'éducation populaire (outil d'analyse de la réalité de la vie associative de proximité), et des collectivités territoriales qui sont en attente de co-construction de stratégie éducative régionale¹¹.

Une telle évolution n'est toutefois possible qu'accompagnée d'un renforcement des services et donc des moyens, en particulier déconcentrés, ciblés sur la production et la gestion de la donnée.

2.2.5. La DAJ, une direction prudentielle dans la gouvernance de la donnée

Si la DAJ n'a pas de responsabilité directe dans la gestion et la gouvernance des données, il reste qu'au travers de son rôle de direction d'appui aux autres composantes du ministère, elle développe une vision relativement prudentielle de la problématique de la gestion et des échanges de données publiques, en particulier des données à caractère personnel. Le principe de protection des données personnelles dans le respect du RGPD (notamment le respect du bien-fondé et de l'exercice du droit d'opposition) reste la ligne de conduite absolue.

En conséquence, la gestion de la donnée se distingue selon son caractère interne ou externe. En interne, la mobilisation des données à caractère personnel à des fins de gestion ou de pilotage (information décisionnelle ou analyse de parcours des élèves par exemples) ne pose pas de difficulté. En ce sens, les interactions entre directions et l'animation de communautés au sein de l'administration peuvent être encouragées. À contrario, étendre les données à caractère personnel envers des acteurs tiers, dont les collectivités territoriales, requiert une analyse au cas par cas, une segmentation liée au RGPD qui oblige, aux yeux de la DAJ, un examen spécifique. Il faut ainsi raisonner traitement par traitement pour bien identifier les destinataires légitimes. Cette observation vaut également dans la perspective de l'instauration d'une plateforme pour en guider les conditions d'accès et de diffusion. C'est pourquoi la DAJ voit dans la possibilité de développer des APIs spécifiques une orientation opportune respectueuse des contraintes juridiques relatives à la gestion des données à caractère personnel. En perspective de la mise en œuvre de ces principes, la DAJ souligne l'importance d'une gouvernance de la donnée concertée et anticipée entre les directions de l'administration centrale du MENJ pour éviter le contexte actuel où la régularisation juridique des traitements de données du ministère est toujours en cours. À cet égard, la question de savoir quels destinataires considérer comme légitimes au regard du droit (et en particulier les collectivités territoriales régionales) fait l'objet d'échanges avec les autres directions du ministère. Il s'agit pour la DAJ de traduire en des termes juridiques (les arrêtés des traitements) les doctrines prônées et déterminées collectivement en la matière.

¹¹ Cf. préconisation 6 : renforcer la connivence INJEP-DEPP-SIES en académie auprès des SSA pour développer une compréhension adaptée des parcours scolaires et d'engagement des jeunes.

2.3. Le contexte de confinement a confirmé la nécessité du développement du numérique et a mis au premier plan de nouveaux besoins dans le champ de l'éducation

Cette prise de conscience institutionnelle et le renforcement de la gouvernance de la donnée publique ne sont évidemment pas indépendants de la conjoncture récente et de la crise sanitaire en particulier. Il est indéniable que les échanges numériques ont connu un essor considérable durant les différents épisodes de cette crise sanitaire qui, par définition, oblige à la mobilisation de données à caractère personnel. Cela est aussi marqué dans le champ sectoriel de l'éducation où la consolidation nationale des effets de la crise sanitaire (présence / absence des élèves ou des enseignants par exemples) ne s'est pas révélée chose aisée dans un suivi au quotidien.

2.3.1. Les interfaces écoles - ménages

Au démarrage de la crise en mars 2020, la situation numérique dans le champ de l'éducation était principalement marquée par les interfaces de vie scolaire et par le recours pédagogique de certains enseignants à divers supports numériques (tablettes, tableaux interactifs, etc.) en classe. La crise sanitaire a complètement changé la donne et placé les échanges numériques internes au cœur des préoccupations éducatives. Les échanges par voie de mails se sont révélés précieux pour les contacts et interactions entre le corps enseignant et les familles, y compris les élèves. Les différents confinements ont certes distendu les relations en proximité physique entre les familles et les enseignants mais l'essor des échanges numériques a permis de consolider la réalité éducative et de maintenir une réelle continuité dans les apprentissages de nos jeunes. De nouvelles pratiques de communication et d'échanges ont été générées, des supports numériques pour, par exemple, organiser la diffusion de supports pédagogiques physiques dans le premier degré notamment, ont été mobilisés.

Si ces échanges numériques se sont renforcés, il est significatif de noter que le premier degré et l'enseignement supérieur ont eu recours à des pratiques plus informelles et décentralisées (par les courriels, les SMS ou les réseaux sociaux) là où pour le second degré, le recours aux environnements numériques de travail (ENTs) qui étaient déjà plus développés a été d'une grande aide pour structurer les contacts.

2.3.2. Les ENTs comme outils de liaison et moyens d'enseignement

Les relations entre l'État et les collectivités territoriales sont régies selon le code de l'éducation par un principe de division du travail entre des collectivités qui assurent l'infrastructure éducative et un ministère qui assure l'infrastructure pédagogique, ce pour le premier comme pour le second degré. La ligne de partage est différente dans l'enseignement supérieur où l'autonomie des établissements d'enseignement confère une complexité supérieure dans l'expression de cette division¹².

Au périmètre du champ d'intervention du MENJ, les environnements numériques de travail (ENTs) constituent une interface tout à fait particulière. Tout d'abord, parce que ce sont les collectivités territoriales qui les financent dans la plupart des cas au titre des investissements éducatifs tout en observant que l'objet ENT est tel qu'il est extrêmement difficile de distinguer la composante infrastructure de celle concernant les contenus de ces ENTs. Ensuite, parce que les collectivités ont mobilisé, dans la plupart des cas, des prestataires privés, des éditeurs de logiciels qui viennent à l'appui de cette infrastructure.

Il faut insister ici sur le fait que la collectivité territoriale régionale est centrale pour asseoir l'importance et l'architecture globale de ces ENTs, en particulier pour le second degré. La disponibilité de ces outils s'étend vers les collèges et tend aussi à s'articuler voire à s'élargir vers le premier degré avec le souci d'une convergence ou plutôt d'une cohérence territoriale globale à assurer à l'échelle de chaque région avec, de facto, la collectivité territoriale régionale comme chef de file visant à disposer de la meilleure cohérence possible de cette architecture.

La crise sanitaire a provoqué une transformation de ce paysage. Même si les opérateurs de ces ENTs sont multiples et proposent aux collectivités territoriales des solutions différenciées, la place prise par le logiciel de vie scolaire PRONOTE est quasi-généralisée. En ce sens, ce logiciel privé, mis en place et opéré par la

¹² Cf. partie 1.3.

société Index Éducation, souvent plébiscité par les chefs d'établissement, a pris une place importante dans le paysage du numérique éducatif, et les collectivités ont vu l'intérêt de mieux développer les ENT en intégrant Pronote pour de fait bénéficier d'informations et de données scolaires non encore en leur possession.

Depuis 2020, la filiale numérique du groupe La Poste (Docaposte) a pris une participation majoritaire dans le capital d'Index Éducation. Menée avec l'appui de la Banque de territoires (CDC), l'acquisition répond à une volonté de consolider une présence stable et renforcée dans le numérique éducatif.

Les différents univers de Pronote permettent d'accumuler de l'information sur des sujets centraux pour la vie de l'élève : organisation de l'emploi du temps bien sûr mais aussi relevés de notes et absentéisme. En ce sens, ces données sont à la frontière entre les responsabilités éducatives de l'éducation nationale et les responsabilités logistiques des collectivités territoriales. S'il faut noter que ces données ne font pas à ce jour l'objet de traitements systématiques et consolidés, elles constituent une zone d'interrogation et de partage potentiel d'information qui pose explicitement la nécessité de s'entendre et de porter vision commune sur la réalité éducative.

2.3.3. Le renforcement du lien entre les acteurs via les échanges numériques

Les observations menées par la mission ont témoigné d'une grande hétérogénéité et d'une grande segmentation dans le jeu des acteurs. Cette affirmation vaut tout autant pour les acteurs de l'éducation nationale que pour ceux des collectivités territoriales.

À ce jour, pour l'ensemble de ces acteurs, les données, dans leur production et leur usage, restent encore très largement affaire de spécialistes. À cet égard, ce sont les services statistiques et les services informatiques qui sont au cœur de ce débat. Cette régulation des acteurs reste tout de même assez structurellement différenciée entre État et collectivités.

Du point de vue de l'État, dans le champ de l'éducation scolaire, il existe une longue tradition de production, de suivi et de traitement des données relatives à l'éducation dans le cadre de la statistique publique qui reste facilement consolidable, de l'unité de base que représente l'établissement à la situation agrégée au niveau académique ou national, ce qui permet d'en dériver de nombreuses et multiples études sur l'état du système éducatif. Si l'enseignement supérieur contribue à éclairer les parcours scolaires, l'action de l'administration centrale est plus complexe, et s'assimile à celle d'un assemblée qui ne dispose pas nécessairement de données immédiatement homogènes émanant des établissements d'enseignement supérieur, ce qui légitime un travail d'ajustement, de correction et d'harmonisation pour disposer d'une vision agrégée de la situation nationale pour l'ensemble du champ éducatif.

Du point de vue des collectivités territoriales, on observe là aussi une grande diversité de situations selon un axe vertical et horizontal. L'axe vertical consiste à rappeler que les moyens assignés à la production et au suivi des données sont proportionnés à l'amplitude territoriale de la collectivité : il n'est pas rare que les communes délèguent à l'Éducation nationale l'activité de recensement des élèves dans le premier degré là où il est plus rare de trouver des collectivités territoriales régionales qui ne disposeraient pas d'un service dédié à la production et au suivi des données éducatives. La variété horizontale réside dans la structuration très disparate d'une collectivité à l'autre des ambitions affichées et des moyens alloués à la fonction statistique. Si certaines collectivités territoriales régionales se sont dotées d'une vraie infrastructure pérenne, d'autres ne disposent pas véritablement de compétences internes stables et établies, et sous-traitent pour l'essentiel les études et apports pouvant être retirés des données scolaires.

2.4. Des services déconcentrés marqués par l'absence de stratégies coordonnées dans la gestion de la donnée scolaire

À l'issue de cette analyse descriptive et commentée des principaux acteurs de l'administration centrale dans leur positionnement en matière de production, de traitement et de diffusion des données scolaires, il est clair que si la feuille de route ministérielle a contraint les acteurs centraux à définir une orientation stratégique partagée et à l'assortir d'une gouvernance raisonnée, une compréhension assez disparate des enjeux de la donnée selon les caractéristiques fonctionnelles des différentes directions demeure.

Le numérique pénètre le monde éducatif avec force, particulièrement sous les effets secondaires de la crise sanitaire récente ; l'essor du numérique éducatif et les données qui en découlent créent de surcroît un environnement foisonnant qui bouleverse la donne en matière de production et d'usage de la donnée et ce sans que l'ensemble des conséquences soient encore complètement prises en compte et donnent lieu à des enseignements stabilisés. Cette appréciation, retirée des observations au niveau national est encore plus fortement perceptible au niveau des services déconcentrés de l'éducation nationale, que ce soit au niveau des académies ou des régions académiques.

Le travail d'observation mené en académie s'est fait selon deux voies complémentaires :

- appréhender, par l'intermédiaire d'un questionnaire épuré, les situations académiques dont l'ampleur des disparités observables dans les relations entretenues avec les collectivités territoriales s'est révélée importante. Les enseignements tirés de ces questionnaires peuvent constituer une base de réflexion pour les acteurs centraux du ministère et apprécier ainsi l'utilité à poursuivre ce travail de compilation des réalités académiques. Il faut ici préciser que les retours d'enquête ont été quasi exhaustifs (toutes les académies ont répondu à ce questionnaire hormis deux académies d'outre-mer) ;
- procéder *in situ* à un travail d'observation dans trois régions, l'Île-de-France, l'Auvergne-Rhône-Alpes et la Corse pour apprécier plus qualitativement cette variété des situations régionales avec les académies ou régions académiques correspondantes et satisfaire ainsi un fonctionnement compatible avec les compétences en responsabilité de chaque entité.

2.4.1. Un état des lieux éclairé par une enquête passée en académie

La convention passée entre la DEPP et l'ARF en 2015 envisageait la possibilité de déclinaisons académiques et c'est l'une des facettes que le questionnaire a pu éclairer¹³.

Si près de 90 % des académies répondantes connaissent l'existence de la convention de 2015, ce ne sont que 60 % d'entre elles qui ont pu la prendre en référence pour promouvoir une déclinaison locale, la plupart du temps au niveau académique. Pour autant, plus de 80 % des académies déclarent engager des collaborations avec d'autres collectivités territoriales (locales et départementales) par voie de conventions. Cette observation est d'importance dans la mesure où elle conduit à renforcer l'utilité d'un travail de recensement systématique de ces collaborations par l'administration centrale pour apprécier le périmètre de l'ensemble des données partagées à l'initiative académique.

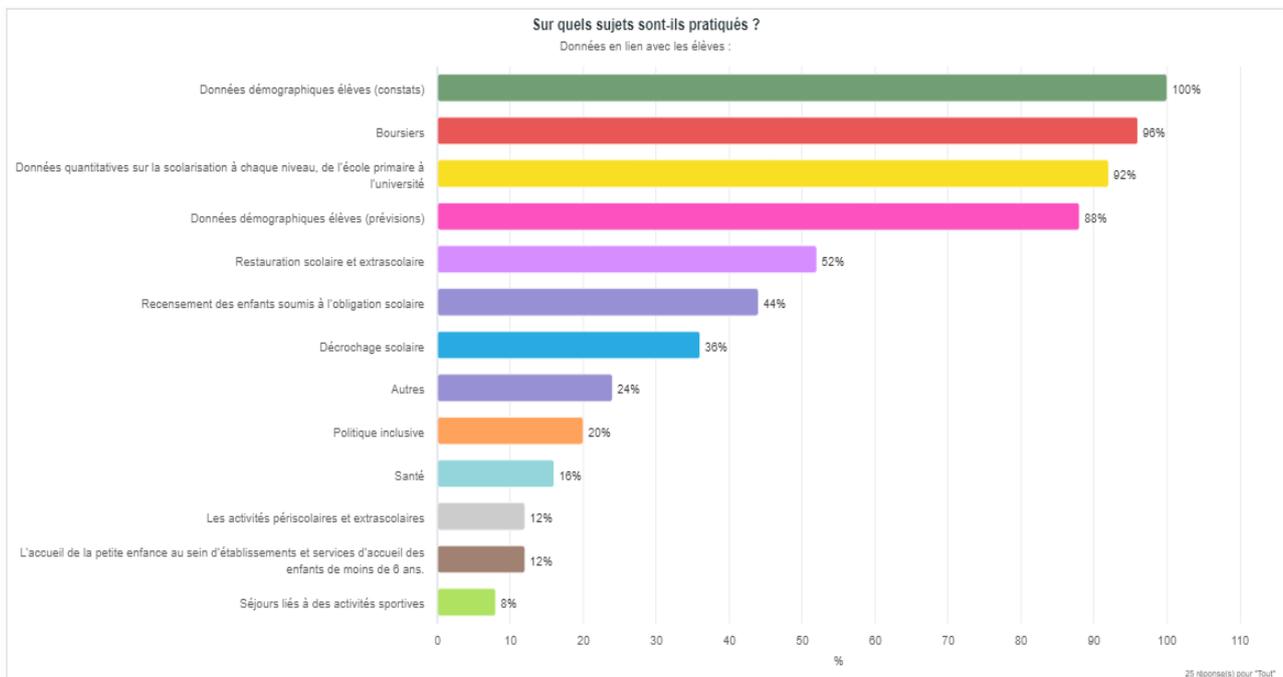
Les principaux acteurs académiques en charge des échanges de données avec les collectivités territoriales sont très majoritairement (60 %) les services statistiques académiques (SSA) qui sont au cœur de la réflexion portée en académie. Les DPD (10 %) ou les DSI (7 %) ne sont que marginalement en responsabilité sur ce sujet. La position dominante des SSA est vraisemblablement due à l'expertise requise pour à la fois collecter, traiter et analyser les données scolaires.

Quelle qu'en soit l'origine fonctionnelle, les acteurs en charge de la gestion des échanges de données le sont pour 60 % d'entre eux en disposant d'une délégation transmise par le recteur pour répondre aux sollicitations au cas par cas. De plus, de façon intéressante, 54 % de ces acteurs indiquent aussi répondre directement aux sollicitations émanant directement des IA-DASEN, confortant ainsi l'existence de coopérations infrarégionales et départementales plus ou moins encadrées et formalisées.

Le sujet des échanges de données avec les collectivités territoriales fait l'objet d'un point-bilan une fois l'an dans la majorité des cas mais il est intéressant de noter que 5 % des académies indiquent n'avoir jamais de point-bilan tout comme 20 % des académies témoignent de points réguliers instaurés une fois par trimestre. Ici encore, la variété des comportements relationnels est forte et on peut pressentir que certaines académies, qui ont déjà fortement structuré leurs relations, sont engagées dans des projets ou des dynamiques partagées qui nécessitent des points de rencontres plus fréquents.

Le tableau suivant explicite la fréquence des réponses en fonction des sujets thématiques les plus sollicités par les collectivités territoriales pour pratiquer des échanges de données :

¹³ Cf. copie du questionnaire en annexe 3 du rapport.



Source : Questionnaire adressé aux académies par l'intermédiaire des COAC - annexe 4 du présent rapport

Cette distribution, qui reflète d'après les services académiques les sujets d'intérêt exprimés par les collectivités territoriales, met en évidence un intérêt premier pour les données relatives aux élèves, et tout particulièrement les données de constat des élèves, les données de situation démographique par niveau et les données relatives aux prévisions. On peut aisément affirmer que cette forte connivence construite autour du suivi quantitatif des élèves (état et perspectives de moyen terme) s'explique par le fait que ces données sont nécessaires aux collectivités pour déterminer des choix d'investissement conséquents sur le bâti scolaire et prendre des décisions en lien avec la carte des formations (professionnelles tout particulièrement). Il est aussi important de préciser que ce sont ces mêmes enjeux qui ressortent également en premier rang dans les observations qualitatives menées avec les collectivités¹⁴.

Les aspects relatifs aux prérogatives couvertes par ces collectivités arrivent en seconde position (restauration scolaire, contrôle de l'obligation scolaire et décrochage scolaire) pour environ la moitié des académies. Assez loin derrière ces deux attentes, d'autres préoccupations thématiques apparaissent dont la politique inclusive, la santé des élèves, les activités périscolaires, mais pour un nombre d'académies beaucoup moins élevé.

Il est également intéressant de noter combien des préoccupations « de nature pédagogique » ne ressortent pas aussi systématiquement que ce l'on pourrait attendre. La mission pose l'hypothèse que la maturité très hétérogène des compétences techniques et statistiques des collectivités territoriales peut en partie expliquer ce constat. Certaines régions, bien structurées en la matière, représentent de vrais interlocuteurs pour les autorités académiques et peuvent, en conséquence, élargir fortement leur regard sur le suivi quantitatif des données scolaires dans leur région là où d'autres, moins avancées en la matière, n'ont pas réellement investi ces champs pédagogiques. Pour autant, il est possible d'envisager qu'à court ou moyen terme, cet élargissement thématique fera l'objet d'une progressive généralisation au vu de l'expérience acquise par les premières.

2.4.2. Des acteurs très variés et diversement coordonnés

Les résultats de l'enquête montrent une grande hétérogénéité des comportements des acteurs territoriaux en matière de gestion et d'échanges de données scolaires.

Tout d'abord, la plupart des académies ont intégré dans leurs stratégies l'importance des échanges de données scolaires avec les collectivités. 90 % des académies le font au niveau académique, 80 % le font également avec les collectivités départementales. Près de 50 % des académies considèrent que ces échanges sont essentiels à l'efficacité de la politique éducative et de façon intéressante, au niveau d'une coordination

¹⁴ Cf. partie suivante.

académique de ces échanges, la moitié des académies appellent de leurs vœux une amélioration de la régulation de ces échanges. Il n’y a par contre que 30 % des académies qui considèrent que ces échanges pourraient être améliorés au niveau infra-académique (départemental).

Pour autant, c’est pour partie paradoxal dans la mesure où ce sont aussi dans les relations infra-académiques (entre DASEN et collectivités) que des difficultés relationnelles peuvent apparaître pour 15 % des réponses académiques, contrairement au niveau académique (0 %). La proximité territoriale et la précision dans l’échelle de la donnée semblent être des facteurs qui tendent les relations entre autorités académiques et collectivités là où une amplitude plus grande dans l’échelle d’observation semble créer plus facilement un consensus dans un usage partagé de la donnée scolaire.

Enfin, l’enquête conforte le fait que les enjeux autour de la donnée scolaire restent concentrés sur le temps scolaire, les préoccupations relatives au temps hors scolaire continuant d’occuper une part très minoritaire des sujets traités entre collectivités et services académiques. Cela est vraisemblablement lié aux modalités de collecte et d’usage des données précédemment décrites pour ces activités hors scolaires.

2.4.3. La régionalisation comme facteur d’hétérogénéité dans les échanges avec les collectivités territoriales

Les modalités d’engagement du processus de régionalisation des académies ne sont pas étrangères à la grande variété observée des situations académiques. D’un point de vue institutionnel, l’installation de la RGPD a profondément marqué les pratiques professionnelles. Le respect des droits d’information et la possibilité d’exercer un droit d’opposition ont obligé les académies à revisiter les conventionnements passés avec les collectivités.

Mais ceci ne semble pas avoir été source d’harmonisation dans les pratiques académiques. Les lignes de clivage sont claires entre les régions académiques qui ont mutualisé la fonction de production et d’analyse de la donnée au niveau inter-académique et celles qui ont laissé aux académies une responsabilité plus forte en la matière. Comme nombre d’académies ont déployé des conventionnements au niveau départemental, la cohérence et la visibilité régionale des échanges de données ne sont pas évidentes à percevoir et à établir, du point de vue des autorités régionales. 60 % des académies indiquent de pas disposer d’instances régulières qui permettent l’échange de données avec les collectivités territoriales.

Entre un niveau national qui reste animé par la DNE par l’intermédiaire du comité des partenaires et un niveau académique (et intra-académique) qui reste l’unité de régulation historique des échanges de données, il n’y a que peu de modalités d’échanges intermédiaires et on ne peut qu’être frappé par l’absence de régulation intermédiaire au niveau de l’espace régional, de la part des autorités académiques tout comme des collectivités territoriales régionales¹⁵.

2.4.4. Un besoin de doctrine nationale élaborée préalablement pour assurer des échanges avec les collectivités territoriales sécurisés

L’enquête passée auprès des académies montre en dernier lieu combien les questions sont nombreuses et ne trouvent pas de réponse simple et partagée. Trois préoccupations ressortent avec force :

- la première est une attente de meilleure formalisation du transfert physique des échanges de données. Au vu des histoires académiques, les échanges avec les collectivités territoriales sont souvent empreints de relations interindividuelles entre personnels techniques qui se connaissent et fonctionnent en connivence. Ce schéma de fonctionnement que l’on pourrait qualifier d’artisanal ne peut satisfaire complètement aux exigences actuelles de sécurisation et encore moins futures si l’on tient compte de l’importance croissante du numérique. Ainsi, il est intéressant de noter que 80 % des académies indiquent procéder à des échanges de fichiers de données par internet. Même si la mission ne considère pas nécessairement que de tels transferts de fichiers puissent concerner des fichiers de données à caractère personnel, il reste que des améliorations procédurales peuvent être envisagées pour garantir un niveau de sécurisation de

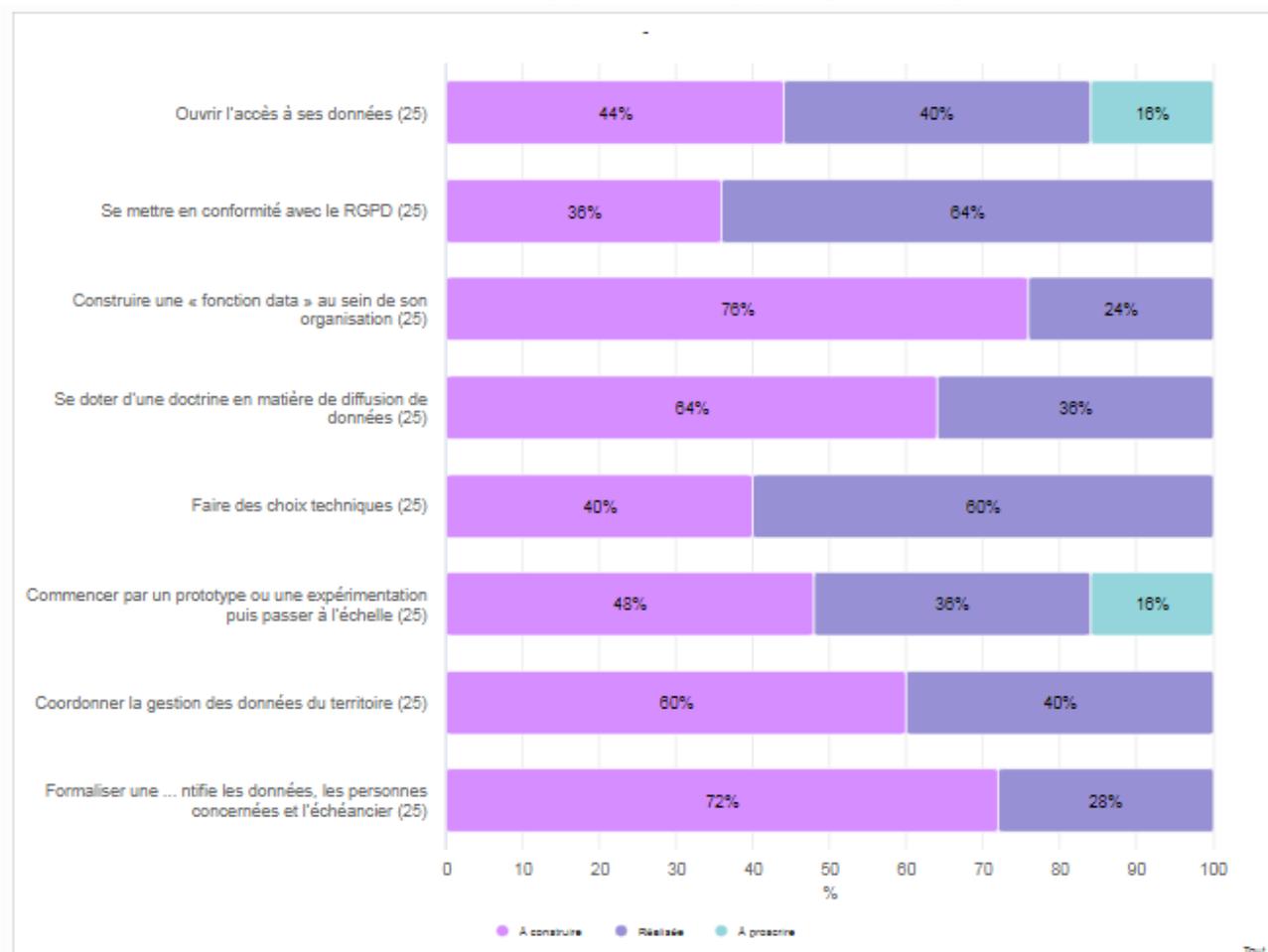
¹⁵ Cf. préconisation 3 : renforcer l’acculturation du MENJ aux enjeux de la donnée scolaire : du producteur (écoles et EPLE), des « consolidateurs » (SSA - DEPP ; DGESCO - DNE) et du régulateur (DAJ) pour mieux maîtriser les effets induits de la RGPD, en particulier l’exercice du droit d’information et du droit d’opposition des usagers.

ces échanges qui puissent être mieux encadré. À cet égard, il faut noter que certaines académies (36 % d'entre elles) se sont dotées d'une plateforme commune d'échanges de données avec les collectivités territoriales. Ici encore, cette observation plaide pour aller plus loin en matière de recensement des pratiques académiques, pour en tirer enseignement et apprécier les bases d'un partage de bonnes pratiques et généraliser les avancées observées ;

- les ENTs constituent un deuxième sujet de préoccupations. Si ces ENTs sont la plupart du temps financés par les collectivités territoriales, deux difficultés sont assez récurrentes. La première est d'établir une cohérence « verticale » de ces ENTs tout au long du parcours de l'élève, du premier degré jusqu'à l'enseignement supérieur. Les ENTs sont en effet le plus souvent ancrés dans le second degré scolaire et de plus en plus de collectivités cherchent à promouvoir une cohérence de ces dispositifs d'interaction numérique en amont (premier degré) et en aval (enseignement supérieur), ce qui pose des difficultés de coordination, soit entre collectivités territoriales, soit entre collectivités et établissements du supérieur. Mais la seconde difficulté est celle de la dépendance d'acteurs privés qui se sont imposés dans ce domaine. En effet, certains éditeurs privés de logiciels se sont de facto imposés sur ce marché des interactions numériques scolaires, et en particulier en matière de logiciels de vie scolaire¹⁶. Le type de relations pouvant être établies avec ces éditeurs privés est une question soulevée mais non traitée dans ce rapport. Il s'agit simplement de relever ici que les difficultés d'interaction sont de même nature pour les autorités académiques que pour les collectivités territoriales. Une multiplicité de données existe qui pourrait faire l'objet de traitements tout à fait utiles et éclairants pour un certain nombre de sujets importants (tels que les relevés de notes, la gestion des emplois du temps, ou encore la mesure de l'absentéisme des élèves) mais ne font pas encore l'objet de consolidation au-delà de l'unité de compte que représente l'EPL. Il y a là une marge de coordination tripartite entre les acteurs qui relève d'une marge de progrès à travailler et dont il sera difficile de faire l'économie de promouvoir la régulation des échanges de données entre services académiques et collectivités territoriales ;
- la troisième préoccupation relève de l'expression des académies quant aux marges de progrès attendues, transcrite dans l'enquête et synthétisée par le tableau suivant :

¹⁶ L'exemple le plus significatif reste la situation de Pronote ; voir le chapitre 2.3.2

Dans le cadre d'un processus mis en place pour de l'échange de données, quelles sont les étapes qui vous semblent : réalisées (R), à construire (C), ou à proscrire (P)



Source : Questionnaire adressé aux académies par l'intermédiaire des COAC - annexe 4 du présent rapport

Plusieurs enseignements sont aussi à retirer de ce tableau. Le premier réside dans le fait qu'une frange des académies (16 %) exprime une réelle réticence à partager des données avec les collectivités territoriales. Si cette quotité reste très minoritaire, c'est toutefois un sujet de préoccupation à prendre en considération dans l'objectif d'avancer vers une doctrine collectivement partagée. Cela peut exprimer aussi une forme de réelle responsabilité des autorités académiques en matière de gestion et de maîtrise de la donnée dont on peut aussi trouver indice dans le fait que 64 % d'entre elles disent s'être dotées d'une doctrine académique en matière de diffusion des données, que 72 % indiquent avoir formalisé une stratégie globale qui identifie les données, les personnes concernées et un échéancier assorti. Cela va de pair avec l'indication donnée que 60 % des académies assurent une coordination dans la gestion de la donnée au niveau de leurs territoires de responsabilité. Mais cette version optimiste est contrebalancée par le fait que 60 % d'entre elles considèrent qu'elles ont encore des progrès à fournir pour se mettre en conformité avec les attendus de la RGPD et qu'elles ont nécessité de faire des choix techniques pour conforter leurs stratégies d'échanges de données.

Il est vraisemblable que les progrès attendus passent par le renforcement d'une politique volontariste de formation autour de la donnée et de ses enjeux, en particulier en terme de la responsabilité engagée par l'autorité académique pour assurer un cadre de fonctionnement sécurisé et surtout pour s'assurer de la compatibilité entre la définition d'une doctrine nationale avec les pratiques en œuvre dans les territoires. Cette médiation, renforcée par la formation entre niveau central et niveau déconcentré, est aussi le moyen de s'assurer de l'existence ou de l'insuffisance, le cas échéant, des moyens et compétences en académie et en région pour maîtriser une politique de la donnée scolaire.

Au final, les enseignements de l'enquête, au-delà de la variété des comportements observés, constituent un plaidoyer pour avancer à la fois sur un recensement bien organisé des conventionnements passés avec l'ensemble des collectivités territoriales et sur l'établissement d'une doctrine nationale mieux établie et plus unifiée ainsi que par l'importance d'une coordination entre académies à l'échelle d'une région académique. Un schéma global de partenariat proposé par le national est attendu et pourrait se résumer dans une forme de proposition telle que formulée par l'une des académies en conclusion de l'enquête : « *il serait sans doute utile de relancer les conventions cadre [région] et d'en créer de nouvelles selon les interlocuteurs [départements, agences d'urbanisme, CARIF-OREF] en les actualisant avec les contraintes RGPD. Une partie des données devrait certainement être accessible au niveau national et le niveau local pourrait compléter par des données spécifiques. Les conventions cadre devraient aussi préciser les données attendues des collectivités locales* ».

Au-delà de « l'outillage conventionnel » que les académies appellent de leurs vœux, c'est aussi une attente en matière de gouvernance académique (et/ou régionale) qu'elles expriment : « *ces différentes réflexions nous amènent à penser qu'une instance de management (gouvernance) de la donnée au sein du rectorat serait quelque chose d'important à mettre en place, tant pour la visibilité des échanges (entrants et sortants, création d'une cartographie) que pour la rationalisation, conformité légale des informations, leur sécurisation, ainsi que le développement d'une culture et d'une stratégie commune sur cette gestion* ». L'utilité d'une forme de comité régional de la donnée scolaire, pendant régional de l'instance nationale que constitue le comité des partenaires, est ainsi affirmée par les acteurs académiques et régionaux¹⁷.

2.5. Une forte attente des collectivités territoriales

La mission a pu bénéficier d'échanges qualitatifs avec les collectivités territoriales à différents niveaux : national au travers de leurs principales représentations, mais aussi en collectant les points de vue de différentes collectivités, par voie de sondage, au niveau communal comme départemental. La mission a également pu confronter les regards entre collectivités territoriales régionales et services académiques dans trois régions : l'Île-de-France, l'Auvergne-Rhône-Alpes et la Corse. En introduction, les régions considèrent que la convention de 2015 entre la DEPP et l'ARF a principalement été un affichage politique et que, depuis lors, le contexte a changé et requiert d'aller plus loin en matière d'accès aux données. Le transfert de certaines compétences illustré par la loi pour la liberté de choisir son avenir professionnel et les responsabilités nouvelles en matière d'orientation des élèves illustrent parfaitement le propos.

2.5.1. Une insatisfaction systématiquement relevée sur la disponibilité de la donnée scolaire et sa fiabilité

Il est marquant de noter combien les collectivités territoriales formulent des griefs assez systématiques à l'encontre des services académiques en matière de gestion de la donnée scolaire. Certaines observations semblent fondées, d'autres beaucoup moins. En fait, les critiques sont le plus souvent inversement proportionnées à la taille de la collectivité. C'est au niveau communal (et intercommunal) que les relations avec les services académiques semblent les plus fluides. Les municipalités disposent d'une longue historicité partenariale avec l'éducation nationale pour le recensement des élèves et leur accès au service public de l'éducation. Cependant, c'est bien au niveau régional que les questionnements stratégiques en matière de formation initiale se dessinent. À cet égard, l'ARF, en tant que représentation coordonnée des régions, confirme un véritable intérêt à installer une plateforme de données dédiée et permettant de fluidifier ces échanges tout en rappelant l'existence antérieure de plateformes gérées en partenariat avec l'État : SYNERGIE, qui est un portail de dépôt de demande d'aides européennes pour les fonds structurels européens fourni par l'ANCT (Agence nationale pour la cohésion des territoires) ou encore OSIRIS, logiciel fourni par l'ASP (Agence de services et de paiement) pour la gestion des fonds européens agricoles pour le développement rural (FEADER). Ces différentes expériences ayant eu des résultats nuancés quant aux apports possibles pour prendre des décisions étayées, l'ARF attend d'une nouvelle convention d'échanges de données un véritable projet politique réel et partagé avec l'État en matière d'éducation.

¹⁷ Cf. préconisation 5 : élaborer un cadre de concertation mieux structuré entre MENJ et collectivités territoriales au national (comité des partenaires) tout comme en territoire par l'installation d'un comité régional de la donnée scolaire.

2.5.2. La complémentarité des préoccupations entre État et collectivités autour de la donnée scolaire impose un nouveau paradigme

L'instauration du RGPD en 2018 a complètement changé les modalités de la gestion des données dans la mesure où c'est la responsabilité des collectivités territoriales qui est engagée pour ce qui concerne les données dont elles disposent. Au lieu de s'en remettre à des déclarations préalables auprès de la CNIL, elles ont l'obligation de tenir à jour une documentation de leurs traitements qui permettent de témoigner de leur bonne foi dans la maîtrise de l'usage de ces données. Ce « renversement » de la preuve oblige à une gouvernance plus professionnelle de la donnée en perspective d'un contrôle devenu ex post par la CNIL, qui se caractérise par l'installation d'un délégué à la protection des données (DPO / DPD), par un recensement à jour des traitements et par la maîtrise de la sécurisation de données, qui inclut les éventuelles entreprises de sous-traitance pouvant être mobilisées (traitement ou hébergement des données).

Le RGPD a engendré, comme cela a été souligné, l'installation d'un DPO / DPD en académie, interface avec la CNIL pour ce qui concerne les traitements académiques mobilisant des données à caractère personnel.

L'exemple de la gouvernance de la donnée en AURA

Ce qui fait débat en matière de gouvernance de la donnée réside dans l'arbitrage de la notion d'anonymisation de la donnée pour respecter le RGPD et considérer que la doctrine en matière de diffusion de données à caractère personnel est correctement appliquée. En l'espèce, les appréciations sont disparates sur ce qu'il est possible de faire ou pas et il n'y a pas de « doctrine nationale » bien arrêtée, ce qui peut conduire à des comportements différenciés selon les académies. L'exemple Auvergne Rhône Alpes (AuRA) est ici instructif : l'académie de Lyon conventionne de longue date avec les collectivités pour la mise à disposition de fichiers de données individuelles anonymisées sur les élèves du second degré, notamment un fichier individuel anonymisé de localisation géographique des élèves, comprenant l'établissement de scolarisation et les formations suivies l'année en cours et l'année précédente, ainsi que l'adresse complète de l'élève. En effet, les éléments d'adressage complets sont nécessaires pour traiter les questions de sectorisation, questions qui relèvent aussi de la compétence des conseils départementaux et des conseils régionaux.

Le DPD de l'académie de Lyon considère que la mise à disposition d'éléments précis de localisation géographique des élèves pose problème au regard du RGPD, car elle rend les données individuelles fournies indirectement nominatives, surtout si elles sont croisées avec d'autres sources de données nominatives en possession des collectivités territoriales (données sur les transports notamment).

Pour autant, la question étant remontée à la DEPP par le SSA, la réponse rapportée à l'académie confirme (à juste raison, NDR) que le traitement des adresses postales des élèves est possible, si la finalité du traitement est justifiée (au sens de l'article 5 b du RGPD). En l'espèce, ici la finalité du traitement de ces données est la sectorisation des élèves, compétence de l'État mais qui relève bien aussi de la compétence des conseils départementaux et des conseils régionaux. Il importe donc que le DPD de l'académie de Lyon déclare ce traitement et en indique la finalité sur la fiche de registre.

Mais la difficulté réside surtout dans le fait que le respect du RGPD impose une information auprès des élèves et de leurs parents pour être informés de la finalité du traitement, des destinataires des informations, de l'existence et des modalités d'exercice de leurs droits.

Ici encore, l'application du RGPD et les voies et moyens mobilisés pour répondre à ce devoir d'information sont très variablement mobilisées selon les académies.

La mise en place de règlements disjoints fixant le cadre des responsabilités selon la nature des données produites ainsi que cela se passe avec la commission européenne au profit des agences nationales qui mettent en œuvre les programmes européens de mobilités et de partenariats pourrait servir d'exemple et être source d'inspiration¹⁸.

¹⁸ Cf. le règlement 2018/1725 qui s'applique aux institutions européennes et concerne tous les actes faisant partie de l'accord de contribution avec la Commission et le règlement 2016/679 (RGPD) pour l'ensemble des activités qui ne sont pas couvertes par la réglementation applicable aux institutions européennes. Un exemple : en France, l'agence nationale Erasmus+ peut donc soit être responsable du traitement des données (dans le cadre RGPD, typiquement l'ensemble des traitements qui s'appliquent aux sites internet, ou données RH), soit être utilisateur des données personnelles dans la mesure où la Commission est responsable du traitement des données sur l'ensemble du programme Erasmus+, et procède à un transfert de données vers les agences nationales.

L'exemple des interrogations posées au sein du MENJ quant aux conditions d'utilisation de la donnée scolaire témoigne aussi des interrogations portées par les collectivités quant aux modalités, opportunités et limites qu'elles peuvent identifier en la matière. La réversibilité s'entend pour les collectivités territoriales. Ces dernières font aussi face aux attendus de la RGPD, ce qui les engage dans une transition qui les conduira de facto à une professionnalisation de la fonction « production et gestion des données scolaires » qui nécessite de disposer de compétences renforcées dans la maîtrise de la donnée, en réponse à la construction d'instruments modernes, fondés sur la donnée scolaire pour exercer plus efficacement les responsabilités que leur confèrent la loi et le code de l'éducation en particulier.

Une première responsabilité qui incombe aux collectivités est le contrôle de l'obligation d'instruction et de formation. Dans le premier degré, les maires agissent en qualité d'agent de l'État pour établir la liste des enfants résidant dans la commune qui sont soumis à l'obligation d'instruction (article L. 131-6 sur le fondement duquel sont pris les articles R. 131-10-1 et suivants). Dans le second degré, ce sont les départements et les régions qui assurent ce suivi.

Au-delà de cette mission de contrôle, la prérogative majeure des collectivités relève de l'immobilier scolaire. Ce sont les collectivités territoriales, chacune à leurs niveaux d'intervention, qui doivent construire les bâtiments, effectuer la maintenance et le gros entretien, les faire fonctionner, voire les fermer après avis du représentant de l'État qui prononce la désaffectation. Cela suppose qu'elles puissent disposer des informations nécessaires et en particulier de prévisions d'effectifs les plus fiables possibles, partagées avec l'éducation nationale. La variété des situations académiques observées est en partie due au fait que le code de l'éducation ne dit rien quant à la mobilisation des données nécessaires pour répondre à cette mission.

Ensuite, les départements ont également pour responsabilité la localisation et l'identification des capacités d'accueil et d'hébergement en collèges¹⁹, ce en vue d'une possible amélioration de la mixité sociale (L. 213-1), et les régions délimitent les objectifs et les contours de la sectorisation des lycées même si la décision à cet égard relève de l'autorité académique. Ici, à nouveau, le recours aux bases élèves, et aux prévisions de court et de moyen termes semblent requis pour répondre à cette attente.

Les régions ont des responsabilités plus importantes encore en matière éducative dans la mesure où elles élaborent tout à la fois un schéma prévisionnel des formations (collèges, lycées, EREA, maritime et agricole) en associant le privé sous contrat à la réflexion (L. 214-1), ce après accord des départements. La région établit également le schéma régional de l'enseignement supérieur, de la recherche et de l'innovation (SRESRI) et arrête la carte de la formation professionnelle initiale après accord de l'autorité académique (L. 214-13-1). Enfin, la région définit et met en œuvre le SPRO et la politique régionale de la formation professionnelle (L. 214-12, L. 214-16-1). On peut ajouter le suivi de l'obligation de formation, avec les missions locales, qui sont des associations (L. 114-1 et L. 313-7 et R. 114-4, R. 114-5 en partie réglementaire).

Ces dernières responsabilités confèrent aux régions la responsabilité de la coordination de l'action publique en matière de formation professionnelle, étendue à l'orientation depuis la loi du 5 septembre 2018, via la maîtrise de l'information des jeunes sur les formations et les métiers ; cette responsabilité peut engendrer une attente de données sur les caractéristiques sociales des élèves et des apprentis, ainsi que des données relatives à l'insertion professionnelle des jeunes.

Au final, si les collectivités territoriales sont en charge des infrastructures scolaires et de l'organisation des activités périscolaires (dont restauration, hébergement, animations périscolaires), l'exercice de l'ensemble de leurs responsabilités justifie le fait de pouvoir disposer de données élèves individuelles et de leur évolution. À ce jour, au vu des arrêtés organisant le recensement des élèves scolarisés et l'usage de ces informations (cf. article 6 de l'arrêté BE1D devenu ONDE, et article 7 de l'arrêté SCONET devenu SIÈCLE), seuls les maires des communes concernées (et les agents de la collectivité territoriale en charge de la gestion de la restauration et des bourses des élèves) sont recensés comme destinataires de ces informations.

Au fond, la question que posent les collectivités territoriales dans l'exercice de ces responsabilités est celle des modalités par lesquelles un système de collecte de la donnée, d'analyse et de production statistique associée, historiquement très centralisé au sein de l'éducation nationale, peut se transformer pour être

¹⁹ Ce qui, dans le langage ordinaire, relève de « la sectorisation ou de la carte scolaire » pour faire la parallèle avec le premier degré et le rôle des communes.

également mobilisé par elles, afin de faire face à une logique croissante de décentralisation de la décision publique et à la nécessaire efficience allocative qu'elle suppose.

2.5.3. La volonté d'assortir les échanges de données à un projet éducatif partagé avec les services académiques au niveau régional

De la même manière que le regard est placé sur la donnée scolaire au sein du MENJ, le choix fait dans le rapport est de porter regard sur le positionnement stratégique des collectivités territoriales en centrant l'analyse sur la situation des régions en tant que chefs de file de l'ensemble des collectivités. Au-delà de l'expression de leurs représentants au niveau national, la mission a pu observer plus précisément la réalité des trois situations rencontrées : l'Île-de-France, l'Auvergne-Rhône-Alpes et la Corse.

Que ce soit en Île-de-France ou en Auvergne-Rhône-Alpes, le constat et l'attente exprimée sont les mêmes : améliorer et stabiliser la relation avec les services académiques afin d'accéder plus aisément aux données nécessaires au bon usage de la dépense publique engagée dans le cadre de leurs responsabilités. Toutes les collectivités rencontrées partagent la même volonté d'inscrire leur action en matière de gestion de la donnée scolaire dans une complémentarité plus active avec l'éducation nationale.

Elles considèrent que les échanges de données (et les conventions afférentes) ne sont pas une fin en soi et que ces dernières ne constituent qu'un outil pour exprimer d'abord la confiance établie entre les deux parties puis rendre compte des besoins identifiés et partagés avec les services académiques. La plupart des acteurs rencontrés se placent dans une phase d'apprentissage au regard du traitement de la donnée et considèrent disposer de capacités de traitement assez faibles et de ressources humaines en nombre limité à cet effet.

La première préoccupation concrète et partagée par les trois régions investiguées est une attente de meilleur partage des données élèves, que ce soit dans le constat ou dans les prévisions d'effectifs. C'est un enjeu d'abord et avant tout financier : celui de la bonne maîtrise de l'infrastructure que constituent collèges et lycées. La région Auvergne-Rhône-Alpes rationalise l'historique de ses relations autour de la donnée scolaire en trois étapes assez distinctes, dans sa capacité à mener une réflexion sur les données, à produire des études spécifiques sur les populations élèves (carte des formations, mixité sociale, etc.). Ces propos peuvent se transposer de manière assez similaire avec ce qui s'est fait dans la région IDF :

- **étape 1** : « on a subi une relation à sens unique conduisant à imposer les choix rectoraux aux collectivités » ;
- **étape 2** : « on a installé une réflexion commune de façon collaborative par l'intermédiaire de groupes de travail partagés mais où l'expertise des services académiques était dominante » ;
- **étape 3** : « la région dispose maintenant, elle aussi, d'une expertise et d'une capacité propre de réflexion : par exemple, il n'est pas rare que des maires interrogent la collectivité sur l'opportunité de création de structures (collèges ou lycées) en fonction de la pression démographique. Aujourd'hui, la région dispose de la capacité à croiser les bases éducation nationale avec d'autres sources (bases INSEE régionales par exemple) pour établir la base de sa propre réflexion et être plus solide dans l'étayage de ses propositions en direction de l'autorité académique (c'est le schéma qui a prévalu pour la décision de projection d'installation de six nouveaux lycées en région) ».

La situation de la Corse est à cet égard un peu différente. Le volontarisme politique pour investir dans la réalité éducative est clairement annoncée par la collectivité qui vient de se doter d'un schéma directeur du numérique dont elle démarre la mise en place. Dans cette phase de démarrage, la collectivité cherche encore à bien paramétrer le périmètre de sa réflexion :

- de quelles données a-t-elle besoin ?
- comment piloter au mieux une politique publique territoriale de l'éducation ?
- comment croiser des informations administratives de différentes origines ?
- faut-il créer un service / une entité transversale de traitement des données en général ou installer des approches thématiques / sectorielles de la donnée pour mieux les comprendre et être plus efficace ?

Ces différences de positionnement, au-delà des particularismes de ces trois régions, reflètent assez bien une situation générale, au national, où les régions n'avancent pas nécessairement à la même vitesse dans leurs capacités à se saisir de l'essor du numérique et à l'intégrer opérationnellement et en lien avec leurs ressources humaines dans leur stratégie de développement éducatif territorial. Certaines d'entre elles disposent aujourd'hui d'une expertise avérée dans le traitement des données scolaires, d'autres en sont encore aux balbutiements et en découvrent le sens, l'importance et les potentialités. Ici, on retrouve l'apport des ENTs dont les régions les plus avancées dans l'expertise technique de la donnée mesurent bien tout l'intérêt à observer le fonctionnement de l'infrastructure par le biais des emplois du temps et à en déduire des hypothèses quant à l'efficacité du taux d'occupation et d'usage des locaux en lycées.

Si produire de la donnée peut sembler aisé dès lors que l'on en a les moyens, il n'en est pas de même pour installer une connivence collaborative. L'accès aux données de l'éducation nationale n'est pas nécessairement simple et fluide. La région Auvergne-Rhône-Alpes par exemple a établi avec ses équipes, une carte d'identité de chaque EPLE dont les indicateurs lui sont propres (élèves, annuaire des chefs, offre de formations, caractéristiques d'hébergement et de restauration, patrimoine immobilier) ; ce travail est mené sans le recours à la base APAE (aide au pilotage et à l'auto-évaluation des établissements), base pour laquelle la région ne dispose pas d'un accès ; c'est une dépense de temps et d'énergie qui pourrait être évitée.

Si des sujets structurels de ce type gagneraient à être traités en collaboration, d'autres sujets moins majeurs mais qui rythment le quotidien de l'action publique des collectivités permettraient néanmoins des gains importants de temps, de coûts et d'énergie qui auraient aussi des impacts sur la reconnaissance faite aux agents qui travaillent ces sujets et leur dégagerait du temps pour des activités additives (cf. exemple ci-dessous).

**Un exemple d'amélioration possible grâce à une concertation mieux organisée
en région Auvergne-Rhône-Alpes**

S'agissant des données liées aux bourses au mérite délivrées par le conseil régional, la situation a évolué par l'instauration de la RGPD.

Jusqu'alors, la région recevait de l'académie les fichiers des élèves concernés (mentions au baccalauréat) et pouvait directement construire une convention spécifique avec chaque jeune pouvant prétendre à une bourse au mérite (entre 7 000 et 9 000 jeunes). C'était une instruction automatisée à partir des fichiers pour récupérer les informations et valider tous les dossiers en un clic.

Ce fonctionnement a été mis à mal avec la RDGP qui implique un consentement préalable et volontaire de la part du jeune pour établir les conventions ; aujourd'hui, une information préalable est requise et elle est redevenue manuelle très largement. Cela a un coût dans la mesure où cette situation requiert une charge de travail (établir le consentement individuel) qui se traduit par l'embauche de CDD et par des délais plus longs de délivrance de ces bourses. On pourrait envisager un relais par les autorités académiques pour « automatiser » ce consentement mais il n'existe pas de gouvernance locale adaptée pour échanger sur ces sujets et, l'année dernière, d'après les représentants de la région « *on a perdu 8 points dans l'efficacité du dispositif* ».

Cet exemple peut se démultiplier (exemples des boursiers, du Pass culture, du faible taux d'utilisation des salles informatiques, etc.). Dans l'idéal, il faudrait, par un travail de concertation préalable État-région, délimiter la liste des services pour lesquels l'obtention d'un consentement de l'élève (ou de ses parents) est requise pour obtenir en une seule fois par an (lors de l'inscription par exemple) sa participation à tous ces services. On retrouve ici l'utilité précédemment évoquée d'un comité régional de la donnée scolaire, équivalent académique du comité des partenaires, porté au national par la DNE.

De même, les choix d'investissement des collectivités restent marqués par des décisions politiques devant lesquelles les réalités techniques ne sont et ne restent que des appuis à la décision marqués par des aléas et beaucoup d'incertitude ; trois types d'expériences sont mentionnés qui témoignent de la complexité des décisions et de l'importance de croiser les regards avec l'éducation nationale :

- la construction d'un lycée qui, au départ, pouvait être perçue comme fragile pour l'éducation nationale devant l'estimation du besoin, et qui se révèle au final très utile parce que l'essor démographique s'est accentué ;

- la construction d'un lycée qui répond à un besoin territorial partagé mais dont la conjoncture économique (et les saignées industrielles locales) ont fragilisé la pertinence du choix, aux yeux des deux parties ;
- la construction volontariste d'un lycée (et accepté de la part des deux parties) pour conforter un projet territorial, qui reste encore un pari partagé mais à conforter.

Il est utile ici de rappeler les conclusions d'une récente note du CNIS : « *Cependant, si la qualité des données est importante, faut-il rappeler que les données ne sont pas des statistiques ? Produire des données certes, mais pour quoi faire ? Le passage aux statistiques suppose leur inscription dans une perspective d'analyse précise qui vise à poser les problèmes que l'on veut tout à la fois comprendre et expliquer.... Autrement dit, les politiques statistiques doivent être définies à travers un dialogue permanent entre les producteurs et les utilisateurs de statistiques. La production de données de qualité est une chose, la production de statistiques en est une autre qui s'inscrit nécessairement dans une problématisation à travers laquelle seront précisés les questions de recherche, les méthodes d'analyse, la nomenclature des données nécessaires, les référentiels de comparaison pertinents, etc.* » (note du CNIS, *Les données statistiques sur les collectivités locales*, 2019, page 37).

Cette observation à caractère général s'applique avec pertinence au champ de l'éducation et au partage de la donnée scolaire entre services académiques et collectivités. Les témoignages recueillis en région explicitent parfaitement cette situation et la qualité de l'investissement des collectivités en éducation est proportionnée à la qualité et à la stabilité des relations qui peuvent être établies, moins d'un point de vue technique car les relations interpersonnelles prévalent entre techniciens, que d'un point de vue « politique » au sens d'une médiation régulière et solidement construite entre collectivités et services académiques.

Ainsi, le constat mené en région montre que les relations d'échanges de données se construisent plutôt de façon informelle et sur la base des bonnes volontés interpersonnelles entre services statistiques respectifs quand ils existent au sein des collectivités. Cette situation est fragile, ne garantit pas de continuité et peut exposer les deux parties à des risques de sécurité. La plupart des acteurs rencontrés appellent de leurs vœux a minima la mise en place d'un comité de suivi de la donnée conjoint pour mieux se comprendre et rapprocher les cultures professionnelles. Un tel comité, existant au national, pourrait se décliner en région et donc en région académique sans pour autant nier le niveau académique pour lequel demeurent des spécificités intra-régionales qu'il faut continuer de prendre en compte. En son sein, et en fonction d'un accord politique partagé, il s'agirait d'identifier communément les données nécessaires pour étayer les choix, de construire l'analyse partagée et de prendre des décisions dans le respect des compétences de chacune des parties mais connues et reconnues de tous²⁰.

Certains de nos interlocuteurs ont mis en avant l'existence du comité régional de l'emploi et de la formation professionnelle (CREFOP) et la possibilité de le mobiliser pour, par exemple, installer une commission supplémentaire de type données de formation. La concertation élargie aux partenaires sociaux en matière d'évolution de l'offre de formation initiale sous l'égide du CREFOP est une des options envisageables mais la mission privilégiée en matière de formation initiale l'existence d'une interface bilatérale comme cela existe pour la carte des formations professionnelles. Il s'agit donc plutôt d'une extension de la médiation existante sous la forme d'un comité régional de la donnée scolaire qui permettrait une réelle acculturation aux différents acteurs pour atteindre une culture partagée.

Généraliser une telle évolution dans la gouvernance de la donnée scolaire est aussi un prérequis pour ne pas dire une condition nécessaire à l'établissement d'un lieu commun qui rejoint les propositions retirées de l'enquête passée en académie où le national pourrait être à l'initiative d'une sorte de plateforme-entrepôt de données, alimentée par le national et complétée par les données et/ou études statistiques académiques au bénéfice conjoint des services académiques et des collectivités. L'objectif partagé exprimé par tous est de concourir par ce travail préalable commun à une meilleure compréhension de ce qui fonde la continuité des parcours des élèves et des étudiants, selon les spécificités territoriales. Cet état de fait engendre un besoin de données multi-sources pour lesquelles le principe d'une complémentarité entre alimentation nationale et alimentation locale (académique et régionale) se justifie pleinement.

²⁰ Cf. préconisation 5 : élaborer un cadre de concertation mieux structuré entre MENJ et collectivités territoriales au national (comité des partenaires) tout comme en territoire (par l'installation d'un comité régional de la donnée scolaire).

En conclusion de l'expression des attentes des collectivités, il faut souligner l'élément facilitateur qui réside dans le fait que ces dernières sont très proches des attentes exprimées par les services académiques, chacun concourant à sa manière et selon les responsabilités qui lui sont dévolues à l'objectif de la réussite scolaire du plus grand nombre. Cette symétrie affichée renforce la demande appuyée des collectivités vis-à-vis de l'éducation nationale pour contribuer et non pas recevoir un projet éducatif partagé, garant de la politique et exigences de l'éducation nationale mais adapté aux réalités et spécificités territoriales locales.

Une telle ambition pourrait poser les bases d'une véritable innovation procédurale, possible du fait de l'essor du numérique et de son positionnement comme révélateur d'une nouvelle donnée éducative, inscrite dans le dispositif de la République ouverte. Elle pourrait aussi donner aux projets éducatifs régionaux une envergure nouvelle et valoriser des investissements récents mais encore peu efficaces comme c'est le cas pour les ENTs par exemple, encore trop souvent restés dans les limbes des perspectives de collaboration.

L'enjeu contemporain est ainsi d'instaurer des espaces politiques de partage de la donnée, au-delà du relationnel technique et interpersonnel entre les services correspondants. Cela doit se réfléchir et être impulsé au national (peut-être sous la forme d'instauration d'assises nationales de la donnée scolaire²¹) ; mais cela doit surtout se décliner en région pour absorber les spécificités des territoires. Les régions avancent à des rythmes différents et c'est une dimension à prendre en compte pour absorber la variété des écosystèmes que constituent les académies et les régions aujourd'hui.

Il est essentiel de dépasser les deux écueils parfois observés : d'une part, le fait de doubler les efforts dans la production de la donnée, le plus souvent par méconnaissance des possibilités et des réalités partenariales ; d'autre part, le face-à-face qui est rarement constructif car chacun arrive avec ses convictions fondées sur des données qui ne sont pas convergentes parce que, soit la temporalité de leur production, soit le champ couvert ne sont le plus souvent pas immédiatement comparables.

Arriver à concilier les apports du national et les spécificités locales (régionales) est un défi d'efficacité et d'égalité des chances dans le parcours scolaire des élèves ; maîtriser cette diversité régionale et la partager avec les autorités académiques est aussi un gage de meilleure utilisation de la dépense publique d'éducation.

2.5.4. Les ressources humaines pour produire et utiliser de la donnée

La mission, à partir de ses observations, ne peut que reprendre à son compte le constat dressé par le CNIS qui soulignait le « *déficit d'appropriation de la statistique publique lié au faible équipement de la plupart des collectivités territoriales en expertise statistique* » et qui rappelait aussi que « *cette faiblesse est maintenant plus grave compte tenu de la place de la statistique dans la gestion publique* » (CNIS, 2019, page 8, *op. cit.*).

Les observations menées en région montrent à la fois une disparité des structures de production et d'usage de la donnée, une diversité des formes de gouvernance retenues région par région, une difficulté partagée à recruter des personnels expérimentés, une faible connaissance des possibilités offertes par le traitement des données scolaires et un besoin de formation / acculturation des services et des élus aux réalités éducatives et aux conditions d'anticipation raisonnée que devraient déployer les collectivités.

C'est aussi à une telle acculturation partagée de la donnée scolaire que doivent servir des projets mobilisateurs tels que ceux précédemment indiqués. Dans cet esprit, disposer d'un lieu / équipement concentrant les jeux de données, d'une sorte de plateforme ou de plateau technique régional, s'avère nécessaire mais non suffisant dès lors que tous les acteurs impliqués ne peuvent participer d'une dynamique collective partagée. L'enjeu est moins « technique » que qualitatif à double titre :

- l'existence et la maîtrise des ressources humaines nécessaires pour faire fonctionner une telle plateforme ;
- l'existence d'une réflexion préalable aboutie pour identifier l'usage précis attendu de la donnée scolaire.

Pour permettre une acculturation collective plus forte autour des enjeux de la donnée, il est nécessaire de développer une sensibilisation des acteurs par l'essor des métiers autour de la donnée, les « *data scientists* »

²¹ Cf. préconisation 1 : tenir des « Assises nationales de la donnée scolaire » : ouvrir la réflexion MENJ-Collectivités territoriales-Edtechs pour établir les termes d'un projet partagé autour de la donnée complétant ressources nationales et ressources régionales.

certes, mais aussi tous les métiers d'accompagnement, de techniciens, de gestionnaires des équipements et des infrastructures plus largement. La prise en compte de la donnée comme variable stratégique de l'éducation requiert une qualité de la donnée, une intelligence des traitements à en retirer, une rigueur dans l'emploi des nomenclatures et dans le renseignement des métadonnées, de sorte que l'usage de la donnée soit intelligible et reproductible. Cette ambition n'est pas seulement affaire de spécialistes. Elle requiert une acculturation collective qui reste encore largement à installer.

Au-delà des structures qui gèrent la donnée scolaire, le premier effort consiste à sensibiliser les acteurs de premier niveau que sont les chefs d'établissement pour garantir la qualité de la donnée scolaire. Leur rôle est essentiel dans l'usage qui pourra être fait ultérieurement aux différents échelons territoriaux (du local au national). Dès lors, une meilleure acculturation des chefs d'établissements sur ce sujet est essentielle pour leur permettre de jouer leur rôle d'interface entre collectivités territoriales et services académiques et en écho d'installer la confiance entre les protagonistes. Leur rôle pivot est central à considérer pour l'ensemble de la chaîne ultérieure de valorisation de la donnée et pour que les applications nationales soient robustes et fiables. Il est par ailleurs indéniable que les chefs d'établissement ont, aujourd'hui plus qu'hier, un intérêt marqué pour les synthèses statistiques des situations des établissements scolaires (radars et nuages de points) et qu'ils mobilisent de façon plus forte les données APAE soit à des fins de pilotage personnel, soit pour répondre aux attentes de l'institution, lors des dialogues de gestion notamment. Cette évolution peut aussi nourrir positivement et promouvoir un fonctionnement similaire auprès des acteurs de la collectivité (et des élus en particulier).

3. Une stratégie qui doit évoluer au bénéfice de l'enfant et du jeune : favoriser une gouvernance territoriale partagée de la donnée scolaire

Cette troisième partie du rapport s'appuie sur le constat et l'analyse précédemment dressés pour mettre en avant des perspectives de développement de la situation qui soient le plus conciliables possibles avec les objectifs affichés et poursuivis par chaque acteur impliqué (administration centrale du MENJ, services académiques, collectivités territoriales, services techniques de ces collectivités ou encore, représentations collectives de ces collectivités. Ainsi, chaque sous-chapitre correspond à des enjeux partagés et à des actions plausibles permettant de les avancer, et l'écriture de cette partie vise à expliciter la façon dont cette conciliation des ambitions et des actions peut progresser collectivement. Un objectif prioritaire commun : la compréhension des parcours et des destins scolaires de chaque élève

3.1. Définir les conditions de mobilisation des données « élèves »

L'enquête réalisée auprès des académies²² ainsi que les observations qualitatives menées avec les collectivités territoriales confortent le fait que l'intérêt partagé entre les services académiques et les collectivités concerne en premier rang la disponibilité et l'usage des données « élèves ».

Il a aussi été établi que les données retirées des environnements numériques de travail (ENTs) pouvaient fournir un ensemble important d'informations, relatives à la fois à la vie des élèves (ex. mesures de l'assiduité) et à la mobilisation des infrastructures (ex. efficacité de l'usage des locaux)²³.

Cela amène à confirmer que l'objectif de conventionnement entre l'État et les collectivités dans le champ éducatif doit viser en premier lieu une meilleure compréhension des trajectoires scolaires des élèves ; cet objectif ne doit pas non plus s'affranchir des progrès à attendre de la structuration de la donnée pour ce qui concerne le temps hors scolaire. L'analyse des parcours et des conditions d'exercice de l'enseignement doit être la finalité partagée par tous les signataires, un commun dénominateur en vue d'une meilleure harmonisation nationale des conventionnements passés au niveau régional.

À cet égard, le modèle de conventionnement établi en région Normandie²⁴ relève d'une démarche que l'on peut qualifier de « bonne pratique ». Elle pourrait préfigurer d'une forme de généralisation en posant le principe de partage d'un annuaire fédérateur entre collectivités et services académiques au travers des ENTs.

²² Voir paragraphe 2.4.

²³ Voir paragraphe 2.3.2

²⁴ Cf. annexe 5 du rapport.

L'exemple normand responsabilise l'ensemble des acteurs dont il définit les rôles et les attributions dans une approche complémentaire :

- les services académiques en tant que garant de la qualité de la donnée (entretien et sécurisation) ;
- le chef d'EPL, en tant que responsable des traitements mobilisant des données à caractère personnel (protection des usagers) ;
- les collectivités en tant qu'utilisateurs et seuls bénéficiaires garant de la non circulation des données en dehors de leur propre usage.

Un tel conventionnement tripartite engageant chefs d'EPL, autorité académique et collectivité territoriale régionale en vue d'un usage partagé des bases de données scolaires, incluant le traitement des données individuelles des élèves, est une démarche qui mérite d'être suivie, cadrée, voir encadrée par le niveau national. Proposée dans le cadre des contrats d'objectifs tripartites, elle méritera à cet égard de servir d'enseignement et d'expérience pour en observer la bonne réalisation et la meilleure capitalisation possible.

3.1.1. Développer les apports réciproques entre État et collectivités territoriales

Dans la perspective d'installer une plateforme de données allant dans le sens d'un renforcement et d'une pérennisation des relations MENJ-collectivités territoriales, il est nécessaire que chaque acteur soit un véritable « contributeur » au regard de ses responsabilités. Ainsi, pour les collectivités, les données relatives au bâti scolaire ou aux ressources numériques devraient être intégrées à la banque de données partagées. De même, les données relatives à la programmation et à l'usage des plateaux techniques des EPL participent de cette capacité à caractériser un « écosystème de l'éducation » où État et collectivités ont nécessité de coordonner leurs actions et leurs investissements pour une meilleure efficacité.

Pour autant, réaliser une telle ambition est compliquée. Les DSI des régions académiques, qui font face à une charge de travail importante en interne, ne sont pas systématiquement acquises à l'idée d'une consolidation des périmètres de données ou des pratiques à moyens constants.

Pour avancer, y compris sur les enjeux de codification législative qui pourraient venir à la fin sédimenter le projet, il est nécessaire d'installer un groupe de pilotage dédié, ayant une mission de préfiguration d'un tel projet dont les contours devront être définis sur la base d'une concertation élargie de l'ensemble des partenaires (en référence aux attendus des assises nationales proposées, par exemple).

3.1.2. Déterminer un équilibre entre une disponibilité de la donnée consolidée au niveau national et la possibilité de développer des usages spécifiques de la donnée au niveau régional

L'importance des usages des données scolaires sera grandissante dans les années à venir. La mission souhaite mettre l'accent sur la nécessaire complémentarité de travail autour de la donnée entre le MENJ et les collectivités et, au regard des enjeux de continuité des formations post bac, entre le MESR et les collectivités.

Si le MENJ dispose d'une mission et d'un champ de compétences bien délimité en matière scolaire, les collectivités sont concernées par d'autres champs de compétences et leurs préoccupations dépassent largement la dimension éducative. À ce titre, elles doivent faire face à une problématique numérique dont la cohérence englobe d'autres champs sectoriels publics (formation - emploi - développement économique, notamment). Les collectivités sont en ce sens en recherche d'arbitrage sur la manière dont elles gèrent leurs investissements numériques entre spécialisation thématique et/ou concentration d'une infrastructure technique répondant à tous leurs besoins.

L'exemple de la collectivité territoriale de Corse est intéressant à évoquer ici. La collectivité s'est dotée d'un schéma directeur du numérique dont elle démarre la mise en place après avoir mené une réflexion sur les besoins pour un pilotage efficient de sa politique territoriale. Cela a supposé une phase d'inventaire des sources et des croisements utiles en matière d'interfaces des données administratives. Si la collectivité est aujourd'hui en réflexion sur la gouvernance la plus opportune, elle est consciente de la nécessité de conventions sectorielles de mobilisation des données avec les différents producteurs pour assurer, quel qu'il soit, le fonctionnement du modèle qui sera choisi et l'atteinte des objectifs auxquels il répond.

Ce type d'interrogations reste assez emblématique des évolutions qu'engagent ou que doivent engager les collectivités territoriales, partagées entre approches techniques et approches sectorielles dans la gestion de

la donnée. Ce débat n'est pas non plus indépendant de la réflexion qu'elles mènent en matière de potentiel de ressources à dédier à la gestion des données. Le caractère centralisé ou décentralisé dans la gestion des données n'a pas la même incidence en termes de moyens humains à dédier à cette thématique.

Dès lors, instruire une réflexion en matière de plateforme de données partageable entre le MENJ et les collectivités implique une double réflexion :

- **interne à l'éducation nationale** pour articuler au mieux une coordination horizontale entre les directions concernées et une coordination verticale entre administration centrale et services académiques²⁵ ;
- **externe, entre partenaires**, pour assurer une convergence d'intentions et de positionnement stratégique entre les régions, ce qui légitime l'idée d'un temps de concertation partagé qui pourrait prendre la forme « d'assises nationales de la donnée scolaire » : moment d'échange entre administration centrale et collectivités régionales - mais aussi entre collectivités territoriales elles-mêmes - afin de prendre conscience et la mesure de la grande variété des modèles en cours, de leur maturation et des investissements réalisés et à venir. Cette concertation devrait s'inscrire dans la suite des états généraux du numérique éducatif pour englober à la fois les enjeux d'infrastructures et de contenus de cet ensemble plus large d'acteurs qui concourent à la mise en œuvre du service public de l'éducation et délimiter le contour de pratiques mieux coordonnées. De telles assises devraient se poursuivre par un groupe de pilotage dédié, ayant une mission de préfiguration d'un projet de plateforme et d'une gouvernance régionale assortie.

3.2. Une gouvernance à repenser au niveau national comme au niveau académique

3.2.1. D'une gouvernance formelle à une gouvernance opérationnelle au sein de l'administration centrale

Comme indiqué en partie 2 du rapport, la feuille de route ministérielle issue de la circulaire interministérielle du 27 avril 2021 a été l'occasion de formaliser une stratégie ministérielle et d'asseoir, dans une architecture explicitée, les différentes modalités de la gouvernance ministérielle de la donnée.

Pour autant, il reste à cette gouvernance à faire ses preuves opérationnelles en matière de schéma organisé et de protocoles clarifiés dans les modalités d'usage de la donnée scolaire.

3.2.1.1 Le rôle des directions

La DEPP et la DNE jouent historiquement un rôle central dans la capacité à mobiliser la donnée scolaire, en interne et en externe. De même, la DGESCO assure un rôle essentiel de producteur de la donnée. Mais l'importance stratégique aujourd'hui prise par la donnée ne permet pas de conférer à ces directions l'exclusivité de la relation à la donnée (production et utilisation). L'importance croissante de la DJEPVA pour ce qui concerne la compréhension des temps hors scolaires dans les parcours des enfants et des jeunes est un paramètre à prendre en bonne considération. De même, La DAJ semble également d'importance, non seulement pour amplifier la couverture réglementaire du ministère mais aussi pour mettre toute son expertise juridique au service d'une stratégie ministérielle coordonnée qui puisse se décliner en éléments d'information et de doctrine partagée par l'ensemble des académies tout en facilitant des ouvertures possibles sécurisées au niveau local. Une doctrine, pas seulement un cadre.

Rassembler l'ensemble des directions du MENJS dans le cadre du « comité stratégique de la donnée » autour de l'administrateur ministériel de la donnée est une avancée significative dans la prise de conscience de l'importance des données scolaires. Tout en favorisant l'acculturation du MENJ aux enjeux de la donnée scolaire (du producteur – écoles et établissements, des consolidateurs – SSA - DEPP, DGESCO - DNE et INJEP - DJEPVA, au régulateur – DAJ), il doit devenir un élément central de la stratégie du ministère ainsi qu'une instance de conseil et d'accompagnement de l'ensemble des acteurs en région qui devront bénéficier de parcours de formation par l'intermédiaire des plans nationaux de formation (PNFs).

²⁵ Cf. partie 2.3.

3.2.1.2 Le rôle de l'AMDAC

L'AMDAC est aujourd'hui rattaché à la DNE. Toutefois, au vu des attentes de coordination entre les directions, un rattachement au secrétariat général serait plus opérationnel pour accompagner au mieux la montée en puissance de l'usage de la donnée scolaire dans les années à venir²⁶.

L'AMDAC doit être en mesure de jouer un rôle de coordonnateur et d'animateur ministériel selon les termes de la mission qui lui est confiée. À l'exemple des conventions, il n'existe pas aujourd'hui de consolidation au national de l'ensemble des collaborations passées entre les académies et les collectivités alors que l'objectif pourrait être rapidement atteint à partir des éléments dont dispose l'AMDAC. En effet, dans l'attente d'un recensement explicite et exhaustif, l'AMDAC dispose d'une première esquisse des situations régionales et des types de conventionnement existants.

Il n'existe pas non plus une culture de collaboration dans laquelle le comité des partenaires, animé par l'AMDAC, serait le réceptacle des demandes de données remontant des collectivités territoriales. La gestion des données scolaires, essentiellement traitée au cas par cas en région, nécessite un travail de consolidation au national qui incombe assez naturellement à l'animation du réseau des SSA, animation pour laquelle une participation régulière de l'AMDAC pourrait être profitable à tous.

3.2.1.3 Le rôle du comité des partenaires

Bien que la volonté de médiation assurée par le comité des partenaires – porté par la DNE – soit appréciée par les représentants des collectivités (ARF), ceux-ci soulignent aussi ses limites, l'instance étant perçue comme unilatérale et « *trop descendante* » de la part du ministère.

Une réelle co-présidence serait un progrès ainsi qu'une véritable déclinaison en région. Il ne s'agit pas de créer une instance supplémentaire dédiée aux données – les associations représentatives des collectivités reconnaissent que les instances actuelles (conseil départemental de l'éducation nationale - CDEN et conseil académique de l'éducation nationale - CAEN) ne favorisent pas l'appréhension de ce sujet – mais de décliner en région les principes portés par le comité des partenaires de la DNE sous la forme d'un comité régional de la donnée scolaire pour contribuer à améliorer l'apprentissage nécessaire en matière d'échanges de la donnée.

3.2.2. Le rôle déterminant des SSA et leurs liens avec les services statistiques des collectivités territoriales

Les observations menées confortent la place stratégique que jouent les SSA dans le dispositif de collaboration à engager avec les collectivités territoriales. « Bras armés » de la DEPP sous l'autorité des recteurs mais sans appartenance effective au « service statistique ministériel éducation », leur positionnement est délicat et explique la variété des situations régionales.

En matière de production et de gestion de la donnée scolaire, les SSA disposent de l'expertise technique de la donnée scolaire et sont les garants de la qualité des traitements statistiques pouvant être menés en région. Cette expertise est appelée à se renforcer au vu de l'extension des périmètres du ministère aux problématiques « jeunesse ».

Face à ce contexte, comme cela a déjà été souligné, ce sont les pratiques locales, fondées sur les relations interpersonnelles entre SSA et services techniques des collectivités, souvent construites de longue date qui prédominent et qui structurent les caractéristiques des échanges de données entre académies et collectivités. Cette situation explique aussi les variations observables d'une région à l'autre quant aux conditions de mobilisation des données à caractère individuel, aux modalités de conventionnement existantes et aux arrangements parfois informels qui peuvent exister d'une région à l'autre.

Tout en s'appuyant sur cette expertise « historique » avérée, il est nécessaire, d'une part, de démultiplier l'effort de coordination entre SSA au travers de journées animées par la DEPP pour renforcer l'usage des données, la maîtrise et la mutualisation des « bons traitements », d'autre part, de stabiliser une visibilité

²⁶ Cf. préconisation 7 : renforcer le rôle de l'AMDAC et, à travers lui, mieux opérationnaliser la gouvernance nationale et le suivi en territoire de l'usage de la donnée scolaire.

d'interaction et de procédures correspondantes dans les échanges avec les acteurs externes dont les collectivités.

3.3. Des actions à conduire pour garantir la réalisation d'évolutions souhaitables

Au regard des échanges et des constats réalisés par la mission, la situation actuelle n'est pas complètement satisfaisante notamment si l'on considère la multiplicité des sollicitations des collectivités et les réponses des SSA qui peuvent être amenés à leur transférer tout un ensemble de données qui font souvent l'objet d'ajustements, de corrections successives, etc. conférant à ces échanges une forme d'insécurité à la fois technique et juridique ainsi qu'une charge de travail additive non négligeable. À cette insécurité s'ajoute la disparité des modalités d'accès, certaines collectivités territoriales disposant d'une possibilité d'accès privilégié aux données (via un réseau privé virtuel - VPN) tandis que d'autres voient leur accès organisé à partir d'échanges directs de fichiers.

Malgré ces obstacles, l'appétence pour les données continue à se développer. Elle a été renforcée par l'article 61 de la loi du 24 août 2021 qui, après amendement parlementaire, met l'État via les services statistiques du ministère en demeure de fournir chaque année aux collectivités territoriales « *des données sociales anonymisées des élèves relevant des établissements publics et privés de la circonscription dont ils ont la charge* », ce en vue de travaux pour améliorer la mixité sociale. Cette évolution renforce la légitimité de trouver un mode de partage des données qui soit fiable, robuste et sécurisé et invite à faire évoluer les modalités du conventionnement actuel toujours lié à la convention de 2015 entre deux acteurs dont la pertinence et aussi à rediscuter compte tenu de ce nouveau contexte.

Alors qu'elle relève a priori plus du ministère dans son ensemble, c'est la DEPP qui porte actuellement cette responsabilité conventionnelle. Au regard des éléments exposés dans le présent rapport, il serait souhaitable de proposer une convention globale voire interministérielle – MENJ et MESR – avec les trois niveaux concernés des collectivités territoriales (ARF, ADF, AMF et AMRF). Une telle proposition portée au plus haut niveau de responsabilité des ministères (ministres, secrétaire générale) permettrait de conjuguer l'ambition politique et la nécessaire sécurisation des échanges²⁷.

Pour atteindre cet objectif et rendre les évolutions opérationnelles, il est nécessaire de conduire simultanément plusieurs actions.

3.3.1. Concilier l'exigence nationale relative à la qualité de la donnée avec un usage facilité au niveau académique, y compris avec des acteurs tiers (collectivités territoriales)

L'« *open data* » et le développement des APIs disponibles sont des pas en avant pour une meilleure transparence et une démocratisation des données scolaires auprès du plus grand nombre. Mais l'essor de l'« *open data* » ne résout pas les interrogations relatives aux collaborations entre services académiques et collectivités qui appellent le maniement concerté de données de gestion et, ce faisant, l'usage plus ou moins important de données à caractère personnel, selon les sujets thématiques traités qui relèvent de leurs responsabilités.

À cet égard, la mission confirme qu'il est indispensable de revenir sur le contenu de la convention de 2015 passée entre la DEPP et l'ARF, qui reste encore aujourd'hui le principal repère encadrant ces collaborations. Cette convention doit être revisitée au sens où elle dépasse les strictes prérogatives de la DEPP (voir paragraphe précédent) pour expliciter une stratégie ministérielle, tout particulièrement depuis le transfert de nouvelles compétences aux collectivités et la production des feuilles de route ministérielles.

Le nouveau cadre conventionnel devra permettre de référer aux choix ministériels et fournir un support de collaboration en matière de données scolaires avec l'ensemble de ses partenaires tels qu'évoqués dans ce rapport²⁸. Il est nécessaire de délimiter non seulement le périmètre des données partagées mais aussi les conditions d'accès et de sécurisation de ces données. Quel que soit le scénario finalisé qui sera arrêté, la gestion de cette plateforme collaborative devra être le reflet de la gouvernance ministérielle installée,

²⁷ Cf. préconisation 4 : revenir sur la convention DEPP-ARF de 2015 et la transformer en convention cadre MENJ - représentants de chacune des différentes collectivités territoriales en région.

²⁸ Développement des APIs vs projet de plateforme collaborative.

bénéficier d'une véritable coordination inter-directions et tenir compte des attentes territoriales différenciées qui se construisent en académies ou en inter-académiques.

À ce titre, comme mentionné au paragraphe 3.2.1.2, la déconcentration du « comité des partenaires » en « comité régional de la donnée scolaire » contribuerait à améliorer l'apprentissage nécessaire en matière d'échanges de la donnée. Il permettra de conduire la réflexion sur l'ouverture d'un accès aux bases de données élèves (ONDE pour le premier degré et SIÈCLE pour le second degré) pour les collectivités territoriales et sera envisagé comme un apprentissage fournissant à l'administration centrale les orientations de production spécifique de données scolaires qui auront pu être dégagées et réalisées.

3.3.2. Co-construire avec les collectivités territoriales un entrepôt partagé des données en éducation

La réalisation d'une telle ambition est compliquée et soulève de nombreuses interrogations, notamment au sein des DSI en région. La prise en compte des préoccupations politiques autour de la donnée scolaire ne peut être menée à bien que si des efforts sont réalisés en matière de moyens dédiés à celle-ci. Cela repose sur la capacité des académies à déployer de la ressource au sein des SSA pour conduire une politique coordonnée avec les collectivités territoriales qui soit fondée sur une meilleure imbrication opérationnelle de leurs responsabilités respectives en matière éducative.

Toutefois, il est possible de décrire les grandes lignes d'une infrastructure partagée de la donnée scolaire qui reposera sur une complémentarité entre une alimentation nationale et des apports régionaux, réunions des données du MENJ et celles des ENTs. Plusieurs options ou scénarii sont envisageables et peuvent constituer les étapes successives de maturation du projet :

1. **un entrepôt de données « éducation »**, géré au niveau central et partagé avec des acteurs déconcentrés et/ou externes par droit d'accès ;
2. **un entrepôt de données « éducation » co-construit** entre administration centrale et services académiques incluant un potentiel d'études répliquables par chacune d'entre elles ;
3. **un entrepôt de données « éducation » partageables** par un accès sécurisé géré au niveau académique ;
4. **un entrepôt de données « éducation » co-construit et partagé** entre les services de l'éducation nationale et les collectivités.

La réalisation du dernier scénario (n° 4) permettrait d'engager une nouvelle évolution dans la construction de l'entrepôt en intégrant à terme la mise à disposition pour l'ensemble des acteurs des données relatives aux infrastructures éducatives et celles relatives aux personnels afférents. Mais elle suppose une co-construction préalable qui devrait se vérifier dans les assises nationales de la donnée scolaire, et se poursuivre par un groupe de pilotage dédié, ayant une mission de préfiguration d'un tel projet de plateforme partagée²⁹.

3.3.3. Maîtriser les droits d'accès et les conditions d'utilisation de la donnée en éducation pour sécuriser les échanges informels : le choix d'un modèle

En conclusion de ce rapport, il apparaît que l'enjeu des collaborations entre l'État et les collectivités territoriales autour de la donnée scolaire peut se penser selon deux modèles de plateformes : le modèle du centre d'accès sécurisé aux données (CASD) versus le modèle de droits d'accès à la main des académies et des régions via une gouvernance assurée par un comité régional de la donnée scolaire, débat entre un projet de type « *Education Data Hub* » et des projets portés de façon déconcentrée qui impliqueraient plus encore les SSA, induisant ainsi la réflexion à mener sur le rôle qui peut être attendu d'eux et les conditions pour le remplir.

²⁹ Cf. préconisation 8 : scénariser une infrastructure partagée de la donnée scolaire dans une complémentarité entre une alimentation nationale et des apports régionaux. Dans cette perspective, les données relatives au bâti scolaire ou aux ressources numériques sont possiblement à intégrer, ce qui rejoint l'idée d'un entrepôt de données, bien commun au service d'un projet politique partagé ; les collectivités seraient contributrices à la fois dans la finalité (notamment en soulignant les enjeux d'insertion professionnelle des jeunes, de par leur responsabilité régionale dans la relation formation-emploi) et dans les apports de données (données relatives au bâti scolaire, données relatives à la programmation et à l'usage des plateaux techniques des EPLE, notamment).

Si la mise en œuvre du premier modèle est lourde, cette solution présente l'avantage de pallier le manque de compétences techniques parfois noté dans les services des collectivités. Elle permet de renforcer les possibilités de travailler de façon collaborative sur les données dans l'environnement physique des services académiques (SSA) pour garantir la pertinence, la sécurité et la confidentialité des données dans les traitements réalisés. La logique du CASD a la préférence de certains pour encourager au développement des collaborations tout en sécurisant fortement l'usage de la donnée. C'est aussi le modèle mis en place par l'INSEE avec ses délégations régionales pour promouvoir des relations en territoire. Ce modèle pourrait faire l'objet de conventions entre rectorats et collectivités.

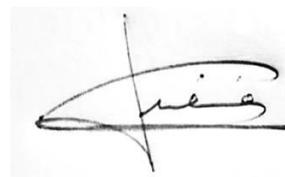
Le second modèle est plus exigeant au sens des enseignements retirés des observations faites en région. Toutefois, il vient répondre plus sagement aux attentes de collaborations en régions et aux volontés affirmer de co-construire. Sa valeur ajoutée tient aussi au fait qu'il peut permettre d'ouvrir sur des données relevant de l'économie, du travail / emploi et du monde de la recherche. Ce modèle permet aussi un enrichissement par mutualisation au sens où il peut permettre, à moindre coût, une consolidation de toutes les études locales existantes (parcours et trajectoires des élèves) qui n'est pas réalisée à ce jour, faute de moyens. Une telle démarche pourrait faire l'objet de plus fortes collaborations entre SSA et de mutualisation d'études thématiques en fonction des spécificités des compétences développées à l'échelle régionale. Cela permet enfin de pallier les différences entre régions et de donner à chacun un rythme d'avancée en phase avec ses aspirations et ses possibilités.



Nathalie COSTANTINI



Yves DELÉCLUSE



Michel QUÉRÉ

Annexe

Annexe 1 :	Lettre de saisine	50
Annexe 2 :	Feuille de route ministérielle MENJS	51
Annexe 3 :	Convention DEPP - ARF de 2015	68
Annexe 4 :	Questionnaire passé en académie	82
Annexe 5 :	Convention Normandie	87
Annexe 6 :	Liste des personnes rencontrées.....	89



RÉPUBLIQUE
FRANÇAISE

*Liberté
Égalité
Fraternité*

IGÉSR INSPECTION GÉNÉRALE
DE L'ÉDUCATION, DU SPORT
ET DE LA RECHERCHE

Section des rapports

N°20-21 059

Affaire suivie par :
Christine Chartier

Tél : 01 55 55 30 88

Mél : section.rapports@igesr.gouv.fr

Site Descartes
110 rue de Grenelle
75357 Paris SP 07

Paris, le 26/10/2021

La cheffe de l'inspection générale de l'éducation,
du sport et de la recherche

à

Monsieur le directeur de cabinet
du ministre de l'éducation nationale,
de la jeunesse et des sports

Objet : Désignation - Les conventions d'échange de données entre les académies et les collectivités territoriales : état des lieux, pratiques et préconisations.

La mission citée en objet, inscrite au programme de travail annuel de l'IGÉSR en 2020-2021, est reconduite pour la présente année scolaire.

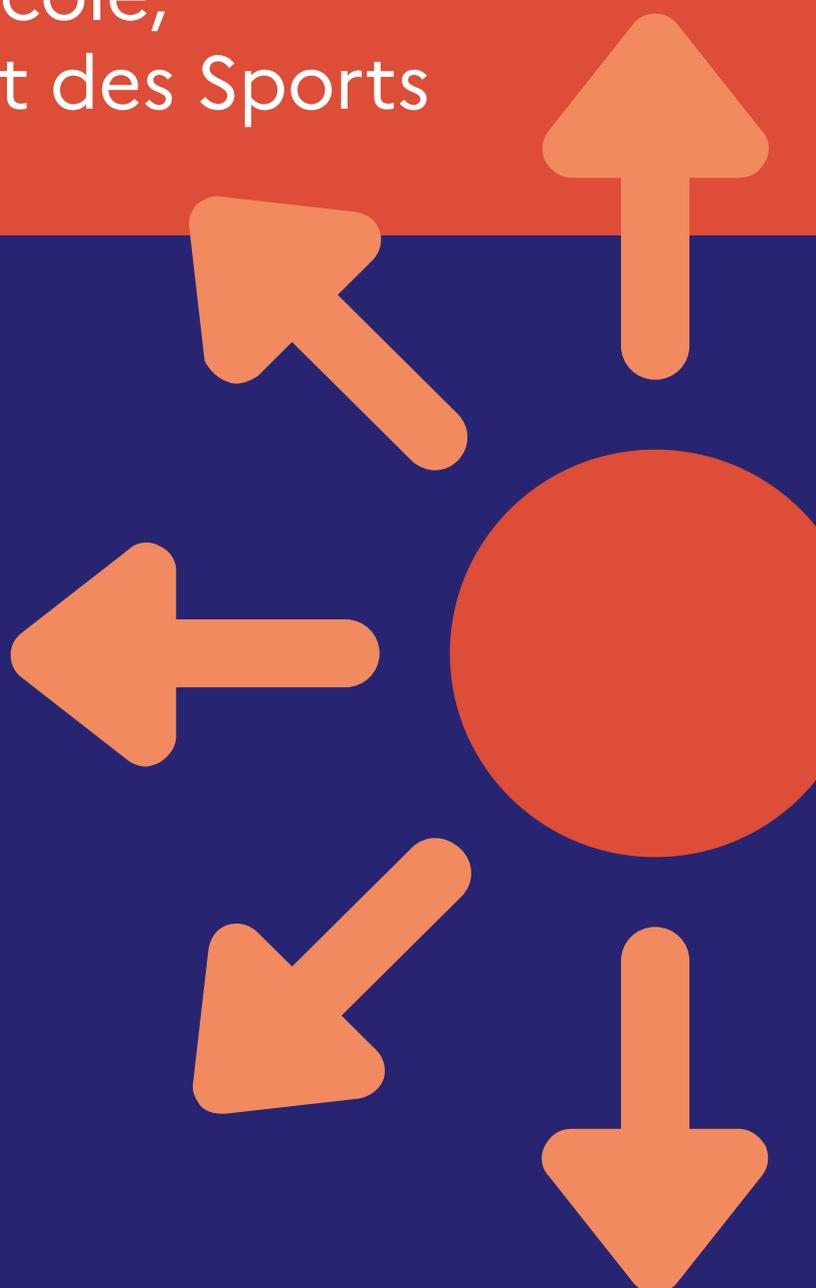
Je vous informe qu'elle est pilotée par Michel Quéré et composée de Alain Abécassis, Philippe Christmann, Yves Delécluse et Patrice Karam.

Caroline PASCAL



Pour une politique de la donnée ambitieuse au service de l'École, de la Jeunesse et des Sports

Feuille de route
Politique de la donnée
du MENJS



Synthèse

Le ministère de l'Éducation, de la Jeunesse et des Sports s'inscrit pleinement dans les orientations fixées par la circulaire du 27 avril 2021 sur la politique publique de la donnée avec l'ouverture d'un nombre croissant de jeux de données et la diffusion des codes sources des grandes applications nationales. Fondées sur les besoins exprimés par les utilisateurs, les mises à disposition de nouveaux jeux de données concernent prioritairement l'information des parents et des élèves (carte des formations, internats, diplômes professionnels, etc.). Après la diffusion des codes sources du processus national d'affectation des élèves en lycée Affelnet, ce sont les codes sources des applications sur le mouvement des enseignants du 1^{er} et du 2^d degrés qui seront bientôt diffusés. Cette politique d'ouverture mais aussi de partage des données s'illustre également avec le développement des API à destination des autres acteurs publics, et notamment les collectivités territoriales, et l'enrichissement du catalogue des API ouvertes.

Le développement du pilotage par la donnée correspond à une ambition forte pour le ministère. Celle-ci s'appuiera d'abord sur la diffusion d'une culture de la donnée à tous les niveaux, les sensibilisations et surtout les formations permettant d'en montrer tout le potentiel pour mesurer, évaluer et guider l'action publique mais également pour innover dans les processus et les services aux usagers. Ce pilotage par la donnée repose également sur une meilleure exploitation des données internes du ministère, en décloisonnant les données, encore trop liées aux applications, et en facilitant le partage avec les directions métiers et les académies. À cet égard, les données scolaires sont bien au cœur de la transformation numérique du ministère. Enfin, un grand nombre de données produites lors de la mise en œuvre du service public de l'éducation sont collectées par des acteurs privés alors qu'elles seraient fort utiles pour le pilotage opérationnel. Ces données liées aux différents outils de vie scolaire s'apparentent à de véritables données d'intérêt général dont devraient bénéficier les acteurs du système éducatif.

Si les activités d'enseignement produisent aujourd'hui un très grand nombre de données liées à la gestion des élèves et des personnels, à la vie scolaire, aux évaluations et aux résultats des élèves, elles sont aussi à l'origine d'une grande variété de données numériques, collectées, stockées et traitées par une multitude d'acteurs. Or ces données, ces traces numériques, ne sont pas encore suffisamment exploitées alors qu'elles offrent des potentialités majeures pour tous les acteurs de l'écosystème éducatif. L'exploitation et la valorisation des données scolaires, notamment numériques, qui doivent s'effectuer dans le cadre réglementaire de la protection des droits des élèves et des personnels de l'éducation nationale, sont porteuses de bénéfices pour les apprentissages mais également pour le travail des chercheurs et pour le développement de nouveaux services innovants. Dans cette perspective, la création d'une plateforme des données scolaires, permettant la rencontre entre les responsables de données et les utilisateurs de ces données, est un projet phare pour le ministère. Développée en partenariat avec l'INRIA, cette plateforme proposera une grande variété de modalités d'accès, d'ouverture et de partage des données qui seront proposées par les fournisseurs de données publics ou privés qui le souhaitent. Ce projet de plateforme (Data Hub) s'inscrit également dans la stratégie nationale en intelligence artificielle qui comporte désormais un volet consacré à l'éducation.

Ces grandes orientations de la feuille de route sur la politique ministérielle de la donnée proposent ainsi des actions ambitieuses pour les deux prochaines années. Ces actions, dont certaines concourent également aux objectifs du Partenariat pour un gouvernement ouvert, doivent bénéficier d'un soutien important de la part des acteurs du ministère, qui déjà se matérialise par l'engagement du Comité stratégique de la donnée, et d'une série de mesures d'accompagnement. Au dialogue constant avec les communautés de réutilisateurs et à l'animation des écosystèmes, notamment avec les collectivités territoriales, doit s'ajouter le développement d'une politique dynamique de renforcement des compétences numériques sur les données, condition nécessaire à la mise en œuvre de la politique des données du ministère de l'Éducation nationale, de la Jeunesse et des Sports (MENJS).

Préface

La collecte, l'ouverture et l'exploitation systématique des données sont au cœur de la transformation des administrations. Leur circulation contribue à renforcer l'efficacité de l'action publique : plus grande transparence pour le citoyen, simplification des démarches et adaptation de l'offre de service public aux usagers. Elle est aussi un facteur d'innovation pour l'ensemble de la société.

Le ministère de l'Éducation nationale, de la Jeunesse et des Sports a pleinement pris la mesure de ces enjeux, en instaurant, dès 2019, un Comité d'éthique pour les données d'éducation, en créant la fonction d'Administrateur ministériel des données en septembre 2019, en mettant en place en 2020 une organisation et une gouvernance de la politique de la donnée couronnées par le Comité stratégique de la donnée réunissant tous les directeurs d'administration centrale et en accélérant l'ouverture des données du ministère.

Aujourd'hui, la nouvelle impulsion donnée par la circulaire du Premier ministre du 27 avril 2021 sur la Politique publique de la donnée, des algorithmes et des codes sources, conduit le MENJS à poursuivre ses efforts pour ouvrir, partager, diffuser et valoriser les données produites ou collectées dans le cadre du service public de l'éducation ainsi que pour lever les freins qui demeurent, notamment en matière de compétences relatives aux données.

Les données sont également au cœur de la transformation numérique du MENJS. La révolution numérique que connaissent les organisations comme les activités économiques et sociales est fondée sur l'utilisation et l'exploitation des données massives qu'elles produisent en permanence. Dans cette perspective, et dans le respect permanent des règles de protection, l'exploitation des données numériques, issues des interactions des élèves et personnels de l'Éducation nationale avec les outils, ressources et services numériques mis à leur disposition, s'avère très prometteuse. Les chercheurs, les innovateurs, les acteurs du système éducatif travaillent déjà à l'identification des potentialités offertes par le recueil et l'analyse des données scolaires.

Cette feuille de route fixe ainsi les grands axes stratégiques de la politique des données pour le MENJS ainsi que l'ensemble des actions à mener dans les deux prochaines années, en articulation avec le plan ministériel défini pour le Partenariat pour un gouvernement ouvert. Issue d'une large consultation des directions, des régions académiques et des opérateurs, elle a été discutée et validée par le Comité stratégique de la donnée.

Grâce à l'engagement de tous ces acteurs, l'ouverture et la diffusion d'un plus grand nombre de données d'éducation seront un levier pour renforcer la confiance dans l'action publique de notre ministère et créer les usages de demain.

La Secrétaire Générale

Enjeux : améliorer l'action publique grâce à la donnée

I. Améliorer l'efficacité de l'action publique

Le MENJS a, dès la décision gouvernementale en 2011 d'aller vers **l'ouverture des données publiques**, contribué largement à la publication de données scolaires sur la plateforme gouvernementale.

Fin 2014, en cohérence avec l'engagement de la France dans le PGO (Partenariat Pour un Gouvernement Ouvert) et la publication le 7 octobre 2016 de la Loi Pour une République numérique qui instaure le principe d'une ouverture par défaut des données publiques, le ministère franchit une nouvelle étape avec la mise en œuvre d'une démarche d'ensemble très volontariste étendue à toutes les directions de l'administration centrale. Fin 2016, une plateforme ministérielle dédiée Open data synchronisée avec la plateforme gouvernementale et en lien avec la plateforme dédiée à l'Enseignement supérieur et à la Recherche était créée pour faciliter la mise à disposition des données et mettre en valeur les jeux de données les plus pertinents.



<https://data.Éducation.gouv.fr>

Cette stratégie d'ouverture et de partage des données se situe dans une tradition établie depuis longtemps au MENJS avec la mise en ligne des données statistiques de la Direction de l'évaluation, de la prospective et de la performance (DEPP).

En effet la DEPP, précurseur de l'open data sur les données d'éducation, a mis en place très tôt une politique d'ouverture des données qui accompagne la diffusion directe de l'information auprès de tous les utilisateurs par le canal de ses pages web.

Pour la DEPP, la mise en place de l'open data avec les bénéfices attendus sur la valorisation des données et l'innovation que leur exploitation peut engendrer, rejoint les missions de la statistique publique, au service des citoyens pour éclairer les débats publics.

Une des priorités de la DEPP est ainsi d'élaborer et de mettre à disposition des informations publiques pertinentes, objectives et documentées, diffusées en ligne gratuitement pour que chacun puisse s'en saisir :

- sous forme de tableaux, de cartes, de graphiques, de séries chronologiques;
- sous forme d'études et d'analyses avec les données associées.

Dans un contexte de prolifération des données, la statistique publique, et donc la DEPP, se doit de produire et diffuser des informations de qualité et référencées pour éclairer le débat public et les choix collectifs.

La loi pour une République numérique a donc rencontré une des missions de la statistique publique que la DEPP porte : publier le plus de données possible à des niveaux les plus détaillés possible, les plus actualisées possible, avec le plus de métadonnées possibles, et un accent sur la forme en veillant à utiliser des moyens de communication modernes, les rendant accessibles aux citoyens.

Ainsi, toutes les publications de la DEPP sont, d'une part téléchargeables en ligne gratuitement et d'autre part, accompagnées par la mise en ligne des tableaux, graphiques et cartes avec les données associées afin que chacun puisse s'en saisir librement.

Par ailleurs des séries longues d'indicateurs sur l'éducation ainsi que des données territorialisées sont de plus en plus nombreuses à être publiées et régulièrement mises à jour sur le site. Enfin des données par établissement sont également mises en open data pour pleinement répondre à la politique publique de la donnée.

La DEPP gère en effet, comme d'autres services statistiques comme l'Insee, un répertoire d'établissements. Il s'agit du fichier des établissements de formation. Ce fichier est accessible sur le site du ministère. En complément de la mise en ligne gratuite des fichiers de données et des publications, la DEPP a mis fin depuis janvier 2014 à la vente de fichiers comportant les données des établissements, issues du répertoire Ramsese. Ces données ont été mises en ligne sur le site du ministère au travers de l'application de consultation et cartographie des établissements (ACCÉ) en avril 2015. L'annuaire des établissements des premier et second degrés, publics et privés sous contrat, est disponible en open data et mis à jour quotidiennement.

Enfin la DEPP met en ligne le fichier des indicateurs de réussite des lycées (IVAL) depuis 1994, qui présente non seulement des résultats bruts, mais aussi des indicateurs de valeur ajoutée des établissements pour rendre compte des actions des lycées pour accompagner les élèves qu'ils accueillent jusqu'au baccalauréat. Ces données illustrent l'apport de l'expertise de la DEPP et font l'objet de nombreuses réutilisations. La DEPP a également mis en ligne début 2021 des données d'insertion par établissement de formation professionnelle (lycées et CFA) de niveau CAP à BTS, en gardant le même principe de calcul de valeur ajoutée.

Pour veiller à la satisfaction des besoins de l'ensemble des utilisateurs de statistiques sur l'éducation, la DEPP assure une veille sur les demandes extérieures de données qui lui sont adressées, afin de repérer les thématiques sur lesquelles il faudrait investir pour mettre en ligne plus de données.

Action 01-21S2 : Développer la mise à disposition de jeux de données issus des services statistiques (DEPP) en Open Data, notamment pour les données qui font l'objet de demandes récurrentes du public.

Enfin, la politique publique de la donnée comporte d'autres volets, au-delà de la nécessaire ouverture des données et des codes sources. Ceux-ci sont présentés et rassemblés dans le document stratégique de la Direction numérique pour l'éducation dont le titre (« **Les données au cœur de la stratégie numérique du ministère** ») fait référence à la première partie du discours du ministre de l'Éducation à l'université d'été du numérique éducatif en août 2018 dont l'idée centrale était de mieux protéger les données pour mieux les valoriser. Ainsi, depuis presque deux ans, la feuille de route du ministère en matière de numérique est centrée sur les données, couvrant les trois dimensions que sont la protection, l'urbanisation des systèmes d'information autour de la donnée et la valorisation des données scolaires qui offrent de grandes potentialités pour les élèves, les enseignants, les cadres, les entreprises de la EdTech et les chercheurs.

La protection des données vise à créer un cadre de confiance pour tous les acteurs de l'école, rassurant pour la communauté éducative et facilitant le développement des activités des entreprises de la EdTech (Gestionnaire d'accès aux ressources pour simplifier la diffusion des ressources et des services numériques, Code de conduite précisant les modalités de mise en œuvre du RGPD, création d'un Comité d'éthique pour les données d'éducation). L'urbanisation des SI autour de la donnée doit faciliter la mise à disposition des données et ainsi garantir les meilleurs liens avec l'ensemble de l'écosystème de l'éducation nationale (interministériel, collectivités locales, acteurs de la EdTech) et avec la société civile. Ainsi, la constitution de grands référentiels de données, découplés des applications, comme la mise en qualité des données faciliteront leur partage et leur circulation. Enfin, la valorisation des données vise à créer les conditions et les structures permettant la mise à disposition des données scolaires pour tous les acteurs intéressés. La collecte des données sur les usages numériques est, comme l'a montré la mise en place de la continuité pédagogique durant la pandémie, un volet essentiel du pilotage par la donnée.

Action 02- 22S2 : Développer un tableau de bord de la transformation numérique à destination des principaux acteurs du ministère et du grand public permettant d'évaluer les différentes composantes des politiques autour de l'équipement des établissements, la formation des personnels, l'utilisation du numérique et ses impacts dans les établissements.

Action 03-21S2 : Faire évoluer le plan de marquage (mesure d'audience) des Espaces Numériques de Travail (à date, 60% des établissements disposent d'un ENT avec 5 millions d'élèves, 400 000 enseignants et 6 millions de parents sont dans le dispositif national de mesure d'audience) afin de mieux connaître et d'évaluer les usages des services et des ressources numériques, de suivre les parcours utilisateur et d'étendre le dispositif au 1^{er} degré (2 millions d'élèves, en forte croissance).

Action 04-23S1 : Publier les données résultant des mesures d'audience des ENT (nouveau dispositif) en Open Data.

II. Favoriser la transparence et stimuler la démocratie

Le ministère de l'Éducation nationale, de la Jeunesse et des Sports, avec plus de 12 millions d'élèves, 20 millions de parents et 850 000 enseignants, dispose d'une immense caisse de résonance des attentes de la société en matière de transparence de l'action éducative. L'un des enjeux des années à venir pour favoriser la transparence est de poursuivre le développement d'une politique de la donnée dans l'éducation : il s'agit de capter les principales attentes des citoyens et de la communauté éducative au sens large puis de les traduire en actions concrètes.

Cet engagement s'inscrit par ailleurs dans le plan d'action ministériel 2021-2023 du Partenariat pour un gouvernement ouvert, dans lequel le ministère s'investit particulièrement en cherchant à impliquer systématiquement la société civile au plus près des acteurs de l'éducation, dans les établissements.

Action 05-21S2 : Mettre en Open Data l'ensemble des données publiées sur les sites institutionnels du MENJS après information de la direction responsable des données (Éducation.gouv.fr et eduscol.fr).

Action 06-21S2 : Ouvrir de nouveaux jeux de données en s'appuyant sur un dispositif innovant permettant la remontée des besoins du terrain et leur analyse.

Action 07-21S2 : Confier aux référents Open Data dans les directions et les académies un rôle de facilitateur d'accès à la donnée.

Action 08-22S1 : Intégrer les données Jeunesse et Sports dans l'Open Data du ministère.

Permettre un accès plus facile à la donnée pour les chercheurs contribue également à la transparence de l'action publique et à la démocratie.

La DEPP s'est engagée dans le projet EQUIPEX « Innovations, Données et Expérimentations en Éducation » (IDEE) du Programme d'investissement d'avenir (PIA 4, Stratégie d'accélération « Numérique et enseignement »). IDÉE est un projet multipartenarial visant à développer la recherche expérimentale en France dans le domaine de l'éducation. Il rassemble six établissements de recherche (Université Paris Sciences et Lettres, Ecole d'Économie de Paris, Sciences Po Paris, Université Grenoble Alpes, Commissariat à l'énergie atomique et CNRS) et bénéficie du soutien de l'Agence nationale de la Recherche (ANR) pour une durée totale de huit ans. Le projet IDÉE comportera trois objectifs principaux:

- Faciliter aux équipes de recherche l'accès aux données administratives en éducation ;
- Constituer et partager un ensemble d'équipements de mesure utiles aux équipes de recherche ;
- Mobiliser et structurer des réseaux de partenaires, chercheurs et professionnels de l'éducation nationale, intéressés par la recherche expérimentale en éducation.

La DEPP s'est ainsi mobilisée sur le premier objectif et apportera son soutien, ses compétences et ses données, à l'occasion de la réalisation par l'équipe projet de la construction d'un centre sécurisé qui permettra un accès des chercheurs plus souple et très sécurisé aux données du système éducatif. Ce centre sécurisé permettra de développer plus de collaborations entre les chargés d'étude et de recherche de la DEPP et les chercheurs en offrant un lieu sécurisé où tous pourront travailler ensemble dans un cadre adapté.

Action 09-22S1 : Faciliter l'accès aux données du système éducatif pour les chercheurs dans un cadre sécurisé.

III. Favoriser la création de valeur économique et proposer des ressources pour l'innovation

Une plateforme des données d'éducation (Éducation Data Hub)

Un des enjeux spécifiques au secteur de l'éducation porte sur la valorisation des données scolaires numériques.

L'ensemble des recherches et des expérimentations menées dans le secteur du numérique éducatif met en évidence les potentialités majeures offertes par le recueil, l'analyse et le traitement des données scolaires et des traces d'apprentissage, ouvrant ainsi la voie à une profonde transformation du système éducatif :

- au bénéfice de l'élève tout d'abord, qui pourra à terme disposer d'un environnement personnalisé d'apprentissage lui permettant de faire le point sur ses forces et ses besoins, d'accéder à des parcours pédagogiques spécifiques, à des aides adaptées ou encore à des suggestions d'activités ou de ressources correspondant à ses marges de progression ;
- pour le professeur ensuite, qui, libéré de certaines activités chronophages, pourra réinvestir ce temps dans la différenciation pédagogique grâce à une connaissance renforcée des besoins spécifiques de chacun de ses élèves ;
- pour le pilotage du système éducatif par la donnée avec les potentialités de l'analyse des données massives et de l'IA pour mieux évaluer les pratiques et en modéliser les évolutions futures (pilotage par la donnée) ;
- pour les entreprises de la filière EdTech, qui pourront mieux cerner les besoins des utilisateurs et proposeront des produits et des services innovants et performants ;
- pour les chercheurs en éducation enfin qui appréhenderont et comprendront mieux les mécanismes des apprentissages, permettant une évolution éclairée des pratiques.

Les données numériques issues des usages numériques des élèves, des enseignants et de l'administration (avec les Environnements numériques de travail ENT, les solutions de vie scolaire, les ressources éditoriales, les plateformes d'apprentissage, etc.) pourraient être collectées et rassemblées afin de les analyser et de les exploiter pour mieux comprendre les mécanismes d'apprentissage, aider les enseignants à proposer des ressources et des activités adaptées à chaque élève ou encore contribuer au pilotage des établissements et du système éducatif. Ces données, qui ne sont ni les données de gestion ni les données statistiques, sont produites massivement et en flux continu, principalement par des entreprises privées.

Des orientations politiques récentes mettent en avant la nécessité de collecter et de rassembler ces données massives pour en faciliter l'utilisation et ainsi conduire au développement de nouveaux outils comme l'IA (Stratégie nationale pour l'IA, Rapport Bothorel et stratégie européenne des données).

Le MENJS, en partenariat avec l'INRIA et avec un financement du PIA 4 (dans le cadre de la stratégie d'accélération « Numérique et enseignement »), a décidé de créer une plateforme des données scolaires qui mettra notamment les données numériques à la disposition de la recherche, mais aussi des entreprises de la EdTech et favorisera leur circulation. Cette plateforme permettra ainsi de collecter, d'analyser et d'exploiter les données numériques massives (anonymisées voire pseudonymisées) issues des pratiques et des usages numériques des élèves, des enseignants et des établissements scolaires (traces numériques). Ce projet concerne les chercheurs, qui pourront développer une meilleure compréhension des mécanismes d'apprentissage et du rôle des environnements scolaires, et les entreprises qui proposent des solutions innovantes et qui pourront ainsi enrichir leurs données par d'autres données, développer et tester leurs algorithmes d'intelligence artificielle, mutualiser leurs données ou encore évaluer leurs innovations.

C'est donc un guichet unique qui constituera un tiers de confiance pour les acteurs producteurs de données, notamment les EdTech, et une plateforme d'échanges entre producteurs et utilisateurs de données.

La réalisation d'une phase de préfiguration, permettant de consulter les utilisateurs et producteurs potentiels et d'en mesurer les attentes, sera entreprise au début de l'année scolaire 2021-2022 avant de lancer à la fin de l'année les travaux de création de la plateforme technique. Parallèlement, les nombreux chantiers de ce projet seront ouverts (gouvernance, juridique, éthique, modèle économique, interopérabilité, etc.). Particulièrement attentive à ce projet, la CNIL sera pleinement associée à sa conception comme à sa mise en œuvre.

La création de cette plateforme des données scolaires est un élément phare de la stratégie IA pour l'éducation, qui sera inscrite dans la future stratégie nationale pour l'intelligence artificielle, tout en favorisant l'ouverture et le partage des données avec les acteurs de l'écosystème éducatif dans le respect de la réglementation sur les données personnelles.

Action 10-22S2 : Créer une plateforme des données scolaires pour partager ces données avec les chercheurs et les entreprises de la EdTech (création de la plateforme dans 12 à 18 mois)

Mettre la donnée au service du sport de haut niveau français

L'apport des sciences dans la performance sportive évolue de manière exponentielle. La data joue un rôle de premier plan pour tenter d'objectiver et de comprendre ce qui se passe au niveau d'un sportif, de son environnement ou encore de sa concurrence, au cours de sa préparation ou lors d'une échéance sportive. Les différentes données collectées sont essentielles pour nos chercheurs et nos « sports scientists » qui travaillent à éclairer le fait sportif et essaient d'anticiper sur l'avenir.

Leurs analyses sont essentielles aux adaptations fréquentes réalisées par les entraîneurs et les encadrements techniques qui cherchent quotidiennement à construire un avantage concurrentiel en vue des grandes échéances mondiales, olympiques et paralympiques.

Au-delà de la construction d'une performance, l'analyse de différentes sources de données permet d'objectiver et d'évaluer des situations complexes, rendant la prise de décision des managers plus aisée, les politiques conduites plus efficaces et leurs évaluations plus objectives.

Face à l'obligation de performance du sport français dans les années à venir, le Ministère et son opérateur public l'Agence Nationale du Sport (ANS) cherchent à améliorer et objectiver la performance individuelle et collective du sport français à travers l'aide au pilotage de la performance par la data, l'accompagnement scientifique et l'analyse comparative agrégée des données aux niveaux national et international.

Pour cela, La Direction des sports (DS), l'ANS et l'Institut National de l'Expertise et de la Performance (INSEP) ont décidé d'unir leur force pour créer le Sport Data Hub (SDH), un nouvel élément clé dans la transformation du sport de haut niveau français.

Le SDH a donc pour ambition d'interconnecter différentes bases de données publiques ou privées afin d'accélérer et coordonner le développement de la data dans l'écosystème du sport français au bénéfice des fédérations et de leurs sportifs de haut niveau (SHN), des laboratoires de recherche dédiés au sport et des différentes organisations qui encadrent le développement du sport aux plans national et territorial, à des fins de pilotage, de recherche ou d'accompagnement des projets de performances de nos SHN.

Concrètement, le SDH est une offre globale de services composée :

- d'une plateforme technologique de collecte des données, sécurisée et conforme au RGPD;
- d'un catalogue de données dans les domaines de la gestion opérationnelle, de l'optimisation de la performance et du pilotage stratégique, provenant de différentes sources;
- d'un catalogue d'outils métier et d'outils d'analyse, adapté aux différents besoins et maturités des fédérations;
- d'une communauté d'experts et de référents data;
- d'un incubateur de projets innovants autour de la donnée.

La structure du SDH étant posée, il convient d'accélérer son déploiement :

- en renforçant l'interopérabilité des sources de données des différents acteurs publics et privés du sport;
- en augmentant la réalisation de cas d'usage dans les domaines du pilotage stratégique, de l'analyse de la concurrence, de la détection des potentiels, de l'évaluation du risque de blessure, de la quantification de la charge d'entraînement, de l'analyse de l'environnement, de l'interface homme/machine, etc...;
- en accélérant la collecte de données de performance à travers le déploiement d'outils de collecte adaptés aux besoins des différentes disciplines;
- en ouvrant les données de pilotage du sport aux instances territoriales intervenant dans la performance;
- en accompagnant l'acculturation des acteurs à l'utilisation de la donnée.

Action 42-23S2 : Renforcer le déploiement de cas d'usage au sein du Sport Data Hub dans les domaines du pilotage stratégique, de l'optimisation de la performance et de la gestion opérationnelle.

IV. Bénéficiaire des « données d'intérêt général » du système éducatif

Le pilotage des politiques publiques ainsi que la plus grande individualisation des parcours scolaires nécessiteraient une plus large mise à disposition des données détenues par des acteurs privés et intéressant l'Éducation nationale. Actuellement, peu de données détenues par les acteurs privés peuvent être utilisées par le ministère.

Un certain nombre de plateformes, détenues par des acteurs privés, produisent des données numériques en flux continu, dans deux grands domaines de la scolarité que sont la vie scolaire (notes, cahier de texte numérique, temps de travail des élèves, absences) et le numérique éducatif (ressources numériques comme les manuels scolaires, messageries, espaces de stockage, de discussion, etc.). Ce sont donc des acteurs privés qui détiennent des données scolaires importantes, telles que les notes et les absences d'une très grande majorité des élèves de l'enseignement secondaire, ou encore les nombreuses sociétés éditrices des ENT (Environnement numérique de travail) qui, avec des commandes des collectivités, fournissent des services numériques aux écoles et aux établissements scolaires, collèges et lycées.

Si les établissements scolaires disposent d'une grande partie de ces données, puisqu'ils ont chacun contractualisé avec la société éditrice des solutions numériques, ce n'est pas le cas de l'académie ou du ministère lui-même. Pourtant, il serait utile de pouvoir disposer d'une agrégation nationale, voire académique ou départementale, d'un certain nombre de ces données. La crise sanitaire actuelle en a d'ailleurs montré l'importance (repérage et mesure du décrochage par exemple avec le nombre d'élèves qui ne se connectent plus à leur ENT à leur cahier de texte numérique). Certaines actions seraient facilitées par l'intégration de ces données aux autres éléments de pilotage (exemple des remplacements de courte durée avec l'élaboration d'un indicateur avancé). Par ailleurs, la DEPP aurait également besoin, à des fins de statistiques et pour alléger la charge des établissements (nombreuses enquêtes pour récupérer l'information), des données issues des logiciels privés de vie scolaire. Un travail mené sous l'égide du Secrétariat Général du ministère vise actuellement à bien identifier les besoins propres à chaque direction du ministère.

Des problématiques juridiques (contrats par établissement; propriété des données; protection des données personnelles) et techniques (interopérabilité insuffisante des outils privés avec les systèmes d'information de l'éducation nationale) contribuent à expliquer cette situation.

Un certain nombre de ces données, sans pouvoir être qualifiées de « données d'intérêt général » au sens de la loi pour une République numérique, s'en rapprochent néanmoins.

Le ministère continue à travailler à la mise en place d'une stratégie permettant de dépasser ces freins et d'assurer une véritable interopérabilité entre l'ensemble des applications utilisées par les établissements, condition indispensable à la consolidation de données de pilotage, en particulier de vie scolaire, et cela quelles que soient les solutions informatiques retenues localement.

Action 11-22S1 : Être en mesure de bénéficier des données scolaires collectées par les acteurs privés de la EdTech (solutions de vie scolaire notamment) après analyse des conditions juridiques de ces transferts vers les académies et l'administration centrale.

Faciliter et valoriser l'ouverture, le partage et l'exploitation des données

I. Ouvrir les données publiques

L'enrichissement des jeux de données en Open Data

Lors du CITP du 5 février 2021, un certain nombre d'ouvertures de jeux de données a été annoncé pour 2021, les jeux de données concernés figurant dans le tableau en annexe indiquant ceux qui sont prévus pour être diffusés sur data.Éducation.gouv avant l'été (Tableau de suivi de l'ouverture des données, codes sources et API publics).

D'autres ouvertures de données sont envisagées pour les mois suivants, avec notamment :

- Les référentiels des diplômes professionnels. Comme les programmes de l'enseignement général, qui figurent déjà sur le site Open Data du ministère, ces référentiels, qui demandent encore un travail de consolidation (nombreux arrêtés modificatifs, nombreuses évolutions récentes), seront publiés en Open Data.
- Les données sur les internats, qui sont actuellement recueillies sous trois formes différentes, font l'objet d'une harmonisation et une collecte unique auprès des établissements est envisagée, en associant l'ONISEP.
- La liste des formations professionnelles proposées dans les GRETA. Celles-ci concernent les formations correspondant aux diplômes nationaux professionnels et les formations à des certifications qu'ils créent eux-mêmes et proposent au RNCP. La rénovation du système d'information (SI2G) avec des nomenclatures communes nationales est en cours et un site proposera à l'automne 2021 l'intégralité de l'offre, avec des présentations communes pour un même diplôme. Il sera alors possible de mettre les données en Open Data.
- La carte des formations scolaires. Après la carte des formations prévisionnelles, réalisée par l'ONISEP et mise en Open Data en avril 2021 (<http://opendata.onisep.fr/actualite/41/8-l-onisep-diffuse-des-jeux-nationaux-sur-l-offre-de-formation-initiale.htm>), sera mise à disposition la carte des formations constatées, dont les données figurent dans les bases de gestion du ministère. La mise en Open Data de ces données permettra des valorisations comme les cartes interactives (exemple de la carte Parcoursup qui propose une recherche dans l'offre de formation).

À plus long terme, le MENJS étudie la possibilité de procéder à l'ouverture des données sur d'autres nouveaux domaines (décrochage scolaire par exemple).

Plus largement, la déclinaison au niveau territorial de cette politique de la donnée complète le dispositif ministériel d'ouverture des données. Elle doit se traduire notamment par la mise en place de plates-formes régionales permettant aux régions académiques et académies de disposer de moyens propres pour ouvrir leurs jeux de données, en articulation avec la politique nationale, et développer ainsi la culture de la donnée au plus près des territoires.

Action 12-22S1 : Poursuivre l'ouverture des jeux de données du MENJS et étendre le champ des domaines couverts (offre de formation, internats, offre de formation des GRETA, référentiels des diplômes professionnels, etc.) (6 à 12 mois).

Action 13-22S1 : Ouvrir une plateforme Open Data à l'échelon de chaque Région académique en articulation avec la plateforme Open Data nationale.

Développer les API

La démarche d'APIsation au ministère a été lancée par la Direction du numérique pour l'éducation, en 2018. Elle a pour objectifs stratégiques de valoriser les données du MENJS et d'identifier de nouveaux usages. D'un point de vue opérationnel, elle doit permettre de réduire le coût des échanges (réutilisation des services de mise à disposition des données via les APIs), de standardiser et de sécuriser les échanges entre SI.

Quatre canaux de besoins sont identifiés : besoins internes, pour les échanges entre les SI du MENJS; besoins des autres administrations, dont ceux de l'observatoire des démarches, que le MENJS souhaite intégrer dans le catalogue interministériel api.gouv.fr. (Ex : en cours la construction de l'API certificat de scolarité); besoins des acteurs privés contribuant à la mise en œuvre des politiques publiques de l'éducation; diffusion des données ouvertes (Ex : API Annuaire de l'éducation pour les établissements publics et privés qui figure désormais sur api.gouv.fr).

Différentes actions ont été mises en œuvre :

- Mise en place d'une infrastructure qui permet de mettre les APIs à disposition, le gestionnaire d'APIs OMOGEN API, et d'une équipe dédiée à son administration;
- Définition d'un cadre de gouvernance des APIs;
- Formation à la supervision du gestionnaire d'APIs et aux développements d'APIs;
- En cours : Mise en place du dispositif de sécurisation des échanges API via le protocole OAuth 2 (échéance mars 2022).

Plusieurs actions sont envisagées dans les domaines suivants :

- Construction du catalogue des APIs du MENJS;
- Définition du cadre d'usages des APIs par les consommateurs;
- Conception d'une fiche de demande, que le ministère souhaite mettre à disposition, afin de permettre aux autres administrations d'exprimer leurs besoins en termes de données du MENJS;
- Adhésion des directions métiers, propriétaires des données, dans la démarche.

Le MENJS pourrait mettre à disposition les API suivantes :

- API école (collectivités)
- API mairie (collectivités)
- API [diplome.gouv](http://diplome.gouv.fr) (Pôle emploi)
- API résultats examens (collectivités)
- API Inserjeunes (DGER MAA) – en place
- API scolarité (éditeurs privés) – en place
- API Adage (PASS CULTURE) – en cours
- API Association – version publique en cours
- API Engagement – en place
- API « Élève scolarisé » (nom provisoire)

Action 14-22S1 : Construire un catalogue des APIs du MENJS et développer des APIs ouvertes.

Action 15-22S1 : Définir un cadre d'usage des APIs par les consommateurs.

Action 16-22S1 : Associer les directions métiers à la gouvernance et à la mise en œuvre de la stratégie d'APIsation du ministère.

II. Partager les données entre acteurs habilités

Le ministère offre une gamme de services en ligne permettant de simplifier significativement les démarches des usagers autant que les tâches des agents. Cette gamme de services va s'enrichir, afin de renforcer sa cohérence, son accessibilité et son attractivité. Un nouveau portail national des démarches en ligne liées à la scolarité a ouvert, à cet effet, le 7 juin 2021. L'accès à ce portail national sera facilité grâce au dispositif d'accès EduConnect qui, en lien avec FranceConnect, permettra un parcours usager de connexion fluide dans les principaux outils de la sphère éducative (en particulier les ENT et les logiciels de vie scolaire).

Plusieurs nouvelles démarches en ligne font l'objet de travaux avec les directions métiers, notamment pour les certificats de scolarité. Les premiers travaux montrent un potentiel à deux niveaux. Le premier réside dans la mise en place d'un service pour les parents leur permettant de télécharger, au format PDF, le certificat de scolarité à tout moment de l'année. Le deuxième axe de dématérialisation réside dans la capacité de présenter les données de scolarité dans une API à disposition des autres administrations et organismes pour soutenir l'objectif « Dites-le nous une fois ». Les développements actuellement prévus pour mai 2022, pourraient être anticipés en fonction des plans de charge et aboutir avant la fin de l'année 2021. Les services viendront compléter l'API statut boursier en cours de réalisation en collaboration avec le CNOUS.

Action 17-22S1 : Développer une API Données de scolarité au profit des autres administrations dans le cadre de l'objectif « Dites-le nous une fois ».

Dans le domaine de la vie associative, l'API Association permet d'obtenir les informations administratives (données et documents) d'une association, à partir de son n°RNA, n°Siren/Siret ou de son nom et code postal du siège. Elle remonte les informations :

- Issues du Répertoire national des associations (associations loi 1901, géré par le ministère de l'intérieur) : gestion d'un flux de données quotidien issu du RNA + appel à l'API RNA
- Issues du répertoire Sirene pour les catégories juridiques de type « association » : appels à l'API Sirene (Insee)
- Saisies par les associations dans leur compte asso, en ce qui concerne les informations qui ne figurent pas dans les répertoires nationaux mais qui sont nécessaires à la réalisation des démarches administratives : relations d'affiliation à des réseaux, fédérations, agréments administratifs, moyens humains, RIB, etc.

- L'API dans sa version actuelle <https://www.associations.gouv.fr/l-api-association.html> facilite considérablement le « Dites-le nous une fois » car, intégrée dans des systèmes d'information de demande de subvention ou de services aux associations, elle évite à ces dernières de devoir saisir une nouvelle fois les informations que détient déjà l'administration. Pour les administrations, l'enjeu est aussi considérable car elles n'ont plus besoin d'effectuer les contrôles au sein des répertoires nationaux. La ville de Paris notamment consomme l'API depuis 2019 pour ses SI de la vie associative.

Elle a pour vocation à s'enrichir prochainement :

- Intégration des associations relevant du régime local (Alsace-Moselle, bases gérées par le ministère de la justice) et de l'ISEE (Nouvelle-Calédonie)
- Ouverture aux données qui permettent de valoriser les activités des associations : leur logo, une description générale, leur lieu d'activité, etc.

Dans sa version enrichie, elle sera semi-ouverte et permettra à des acteurs privés fournissant des services numériques aux associations (type HelloAsso pour les dons ou les adhésions en ligne) d'y intégrer des données d'activités, qui pourront alors être exploitées dans d'autres services numériques aux associations, leur évitant alors, là aussi, de devoir les saisir à nouveau.

Dans le domaine de l'engagement civique, l'API Engagement <https://api.gouv.fr/les-api/api-engagement> permet de récupérer et de diffuser les missions des différents réseaux d'engagement : service civique, réserve civique de l'Etat, mais aussi de collectivités territoriales et acteurs privés partenaires. Elle permet aux plateformes d'engagement publiques et privées de mettre en commun leurs missions, augmentant ainsi considérablement la visibilité des annonces et le taux d'occupation des places offertes aux potentiels bénévoles et volontaires.

Concrètement, chaque plateforme peut intégrer l'API pour diffuser ses missions vers les autres plateformes partenaires d'une part, et récupérer, d'autre part, les missions disponibles sur les autres plateformes selon des critères spécifiques comme la localisation, le domaine d'action, etc. Le tableau de bord de l'API permet à chaque partenaire de configurer les flux de missions entrant/sortant et de mesurer la contribution de chaque plateforme en missions et en trafic.

III. Exploiter les données

Plusieurs projets du ministère concourent à favoriser l'exploitation des données, qu'il s'agisse de ses propres données ou plus largement des données de l'écosystème éducatif (cf. le projet de plateforme des données d'éducation, supra). Le développement d'une stratégie Intelligence artificielle dans l'éducation contribue également à cet objectif.

Une meilleure exploitation des données internes

En matière d'exploitation des données, et dans le respect de la réglementation sur le droit des données à caractère personnel, le ministère pourrait valoriser davantage ses propres données.

À court terme, 2 projets ont été identifiés :

- La mise à disposition pour les académies et les régions académiques, sur une plateforme centralisée, des données issues de leur SI de gestion afin que celles-ci puissent développer des outils de pilotage opérationnels, actuellement d'une grande complexité liée à la démultiplication des sources de données.

- Le décloisonnement des données liées aux parcours scolaires afin d'améliorer le pilotage opérationnel de chacun des acteurs académiques, régions académiques et administration centrale, pour les procédures d'orientation, d'affectation et les campagnes de lutte contre le décrochage scolaire (Cockpit de pilotage des parcours).

Action 18-22S2 : Mettre en place une plateforme d'ouverture, de partage et de valorisation des données de gestion afin de permettre aux directions métiers et aux académies d'améliorer leur pilotage opérationnel.

Le projet GAIA-X de Data Space sur les données « éducation et compétences »

Le ministère a par ailleurs intégré le **projet GAIA-X**, projet visant à faciliter la circulation des données dans un cloud souverain européen. Lancé en juin 2020 par la France et l'Allemagne et piloté par un consortium qui regroupe plus de 200 institutions publiques et privées, avec les plus grandes entreprises européennes, ce projet maîtrise européenne des données dans un cloud souverain a pour ambition d'offrir une alternative aux géants anglo-saxons. L'objectif est de proposer des services de circulation de la donnée accessibles à partir de plusieurs solutions de cloud public, permettant de mettre en œuvre des solutions qui répondront aux grands principes européens fondateurs du RGPD et du futur Data Governance Act comme la transparence, l'interopérabilité et la portabilité des données.

Dans le cadre de cette stratégie, la création d'espaces de données sectoriels européens est devenu un axe essentiel (Data Space).

A l'initiative du MENJS, un projet de Data Space « Éducation and Skills » a été proposé puis validé, son objectif étant de fédérer les acteurs de la EdTech et les acteurs institutionnels (ministère, ONISEP notamment) pour fournir, consolider, partager les données produites (dans le cadre de la réglementation en matière de droit des données à caractère personnel), au bénéfice des chercheurs et des entreprises pour développer de nouvelles applications à base d'IA. Le projet doit permettre de définir les conditions d'échange et de partage des données d'éducation et d'orientation dans un cadre européen et éthique. A terme, l'objectif est de voir émerger un écosystème européen de la donnée d'éducation et des compétences.

Action 19-21S2 : Préparer la mise en place du « Data Space Éducation et compétences » au niveau français et européen en partenariat avec les entreprises de la EdTech (gouvernance, premiers cas d'usages, groupes de travail thématiques, etc.) dans le cadre du projet européen GAIA-X.

La feuille de route IA pour l'éducation

La stratégie IA pour l'éducation, qui doit être intégrée dans la future Stratégie nationale IA, trois ans après le rapport Villani, se caractérise par trois grands volets : des principes, des objectifs et des actions.

L'adoption et la mise en valeur de grands principes permettra de rendre acceptable une IA qui doit être au service des acteurs de la communauté éducative.

- L'IA doit aider l'enseignant et non le remplacer. Celui-ci doit avoir le choix d'activer/désactiver les utilisations de l'IA. Les enseignants doivent être associés à la conception des outils intégrant des solutions d'IA.
- L'IA ne doit pas conduire à des décisions individuelles automatiques pour les élèves (orientation, affectation, notes, etc.) sans intervention humaine.
- L'explicabilité (ou l'interprétabilité) des résultats de l'IA est nécessaire afin que les enseignants concernés soient en mesure de comprendre l'analyse réalisée par la machine et d'en expliquer les raisons éventuelles à un élève
- L'IA ne doit être utilisée avec les élèves qu'à des fins pédagogiques ou éducatives.

Les travaux du Comité d'éthique pour les données d'éducation, créé en 2019, accompagneront la mise en œuvre de cette stratégie IA pour l'éducation.

L'IA doit accompagner la transformation numérique dans l'éducation. Sensibiliser, démystifier, expérimenter, former, convaincre : c'est ainsi que l'IA pourra assister utilement les acteurs et améliorer les politiques publiques. Avec l'IA, plusieurs objectifs sont attendus :

- Améliorer les connaissances des enseignants au service de meilleurs apprentissages ;
- Fournir des outils pédagogiques ou éducatifs pour accompagner la transformation des pratiques d'enseignement et d'apprentissage ;
- Améliorer la gestion des parcours des élèves (orientation, formation, décrochage scolaire, etc.) par une utilisation massive des données de scolarité ou des données produites par le ministère ;
- Aider au pilotage des établissements et du système éducatif par une meilleure utilisation des données dans un cadre éthique et sécurisé ;
- Contribuer dans la durée à la formation des citoyens et à l'employabilité ainsi qu'au développement d'Edtech compétitives et responsables.

Les nombreuses actions du ministère portent ainsi sur l'innovation et la recherche, l'accompagnement des acteurs et la formations des utilisateurs de l'IA pour l'éducation.

Les premières actions concernent le travail avec l'écosystème du numérique éducatif, des partenariats s'établissant avec la recherche et les entreprises (Groupes de travail thématiques avec les laboratoires de recherche, partenariat avec le LORIA de l'Université de Lorraine pour la collecte et l'analyse des traces d'apprentissage, participation des laboratoires IA de l'INRIA, de CNRS et des universités, projet de Plateforme des données d'éducation).

Des solutions intégrant l'IA sont ensuite développées, pour la pédagogie ou pour la classe, selon des modalités qui s'apparentent à une recherche-développement (dont les assistants IA pour l'apprentissage en français et en mathématiques au cycle 2). Ainsi, le PIA 2 a ainsi permis le financement du Partenariat d'innovation pour l'IA, permettant à 6 start-up ou consortium, incluant des jeunes entreprises innovantes, de pouvoir développer et expérimenter des solutions d'IA pour les apprentissages fondamentaux (lecture, calcul). Le ministère a également lancé un projet pour créer et diffuser un assistant vocal pour l'apprentissage de l'anglais en primaire, permettant de travailler sur les données d'enregistrement (voix). De la même façon, un projet d'aide automatisée à la correction de texte, très prometteur pour faciliter le travail d'évaluation, permettra de constituer et de mettre à disposition un corpus important de textes (dans le Éducation Data Hub, afin de développer des solutions de traitement automatique des langues). Enfin, le dernier domaine concerne la sensibilisation et la formation à l'IA, qu'il s'agisse des enseignants et des cadres ou bien des élèves.

Action 20-23S1 : Mettre en œuvre la stratégie IA pour l'éducation.

Action 21-22S1 : Développer de nouveaux partenariats d'innovation pour l'IA (P2IA) en les étendant aux principaux enseignements du collège et du lycée.

IV. Codes sources et algorithmes

L'ouverture des codes sources et des algorithmes

Les codes sources et les algorithmes constituent des documents administratifs communicables.

Le ministère a dans le cadre d'un groupe de travail ad hoc (correspondants des directions et représentants d'académies) :

- Mené l'étude des textes juridiques liés aux algorithmes/codes sources
- Rédigé une mention légale type adaptable par direction
- Commencé un recensement des algorithmes entraînant une prise de décision.

Les codes sources et/ou les documents liés au développement de ces outils (spécifications fonctionnelles) ont fait l'objet de demandes récurrentes de communication ou de publication en ligne. Les réserves initiales opposées à la communication de ces éléments ont, tout d'abord, tenu, aux risques importants d'atteintes à la sécurité des systèmes d'information des administrations que présentait cette communication et large diffusion, et à tout le moins, aux difficultés pour l'administration d'évaluer ces risques. Par ailleurs, un des freins à la communication des codes source peut tenir aux difficultés techniques rencontrées par l'administration pour publier ces codes qui se sont constitués, au cours des années, par sédimentation progressive et qui n'ont pas été initialement conçus pour être ouverts et partagés.

Le ministère est désormais résolument engagé dans une démarche d'ouverture systématique de ses codes sources, dans la mesure des moyens dont il dispose et des contraintes techniques, organisationnelles et de sécurité qui s'imposent à lui.

Action 22-22S2 : Publier les codes sources des applications sur le mouvement des enseignants (1^{er} et 2^d degrés).

Action 23-22S1 : Réaliser un état des lieux des algorithmes et codes sources à publier en priorité (en fonction des attentes des usagers).

Action 24-22S2 : Adopter des démarches de développement permettant la publication du code source dès la conception, soit par réorganisation des équipes de développement internes au ministère, soit en intégrant des clauses ad-hoc dans les appels d'offre.

La stratégie en faveur des logiciels libres et des communs numériques

Depuis plusieurs années le ministère de l'Éducation nationale, ses opérateurs, les académies travaillent à la production de services et de ressources basées sur des logiciels libres et l'open éducation (par exemple, plateforme nationale M@gistère de formation continue des enseignants utilisant le logiciel libre Moodle; Apps.Éducation.fr qui regroupe un ensemble de logiciels libres pour le domaine de l'éducation, s'appuyant sur des logiciels libres existants). La Direction du numérique pour l'éducation contribue également activement au socle interministériel des logiciels libres, depuis 2010. Elle s'appuie notamment sur un pôle national dédié au logiciel libre, le pôle EOLE/logiciel libre basé à Dijon, qui développe depuis presque 20 ans des solutions libres de gestion des infrastructures (pare feu, gestion de comptes, partage de fichiers...) pour les EPLE et écoles.

LE MENJS souhaite encourager davantage l'utilisation de logiciels libres. Cette action établit des liens entre le renforcement de la stratégie numérique du ministère et les propositions issues des États généraux du numérique pour l'éducation de novembre 2020. Elle s'appuie notamment sur le travail réalisé par la direction interministérielle du numérique (DINUM) et le programme de la commission européenne « pour une Europe numérique ». Elle complète la stratégie du ministère pour les données et nécessite :

- la mise en œuvre d'une stratégie sectorielle du logiciel libre et de l'Open éducation (allant des infrastructures aux ressources à caractère pédagogique) de façon à favoriser les échanges de pratiques et de compétences entre enseignement scolaire et enseignement supérieur;
- l'animation de l'écosystème avec des dispositifs de soutien;
- une incarnation au sein de la direction du numérique pour l'éducation.

La mise en œuvre de cette action, ses principes directeurs et son déroulement nous aideront à élaborer et à fournir de meilleures solutions et de meilleurs services à l'ensemble de la communauté éducative, pour exploiter le pouvoir d'innovation et de collaboration de l'open éducation et du code source ouvert. Cela permettra d'intégrer le ministère dans la communauté de l'open éducation et du logiciel libre en tant que contributeur agile et participant actif. Cette stratégie renforcera par ailleurs la confiance et la transparence dans les outils numériques proposés par le ministère.

La création de services à l'aide de codes sources ouverts, avec une vision sectorielle, accessible à tous, contribuera à créer des communautés et des procédures interopérables, transparentes et non discriminatoires pour l'accès aux données, pour l'IA et pour des modèles et des méthodes de formation en matière d'apprentissage automatique. Cette action aura également pour ambition de favoriser la création de ressources pédagogiques libres et ouvertes, accessibles à tous, de façon à créer des processus collaboratifs et des communautés pour faciliter la diffusion de la production volontaire des enseignants (valorisation dans les carrières des enseignants, des agents et des cadres; soutien aux différentes associations et/ou communautés du libre; formation des enseignants, des agents et des élèves aux vertus et aux principes du libre en éducation).

Action 25-22S2 : Concevoir un « Open Source Program Office » sectorisé pour l'éducation, doté d'indicateurs, afin de fournir et de mutualiser des solutions, des ressources et des services pour l'ensemble de la communauté éducative (6 à 12 mois).

Développer le capital humain : acculturation et gestion des compétences

I. Acculturer à tous les niveaux de l'organisation

La consultation publique conduite par la mission Bothorel souligne un fort besoin d'acculturation des agents publics aux enjeux des données. Le rapport indique que « l'utilité et le potentiel des données et des codes sont mal compris, et restent identifiés dans l'imaginaire collectif comme des sujets purement techniques ». Les agents, en particulier des directions métiers, et quel que soit le niveau hiérarchique, doivent donc mieux percevoir la valeur de la donnée et son importance dans l'élaboration et le pilotage des politiques publiques.

Le rôle des données dépasse le cadre des métiers et des compétences professionnelles individuelles. Tous concernés par les données, que ce soit en tant que propriétaire ou responsable des données, propriétaire d'un processus de traitement des données, prescripteur d'applications embarquant des données ou en tant que simples utilisateurs, les agents ont besoin d'une culture commune. En particulier du fait de l'ouverture des codes et des données, l'administration doit présenter un visage unique et doit organiser une collaboration entre ses services pour « dé-siloter » les données et assurer ainsi leur cohérence voire leur unicité. Cette culture commune doit donc reposer sur les notions de collaboration et de transversalité. Développer un esprit « collaboratif » et reconnaître la dépendance vis-à-vis d'autres services, qu'ils soient en charge de la collecte, du traitement, de l'exploitation ou de l'utilisation, ou anticiper les coopérations interservices sont les objectifs à poursuivre.

Les actions de sensibilisation à l'exploitation et à la valorisation des données, avec un vaste programme d'acculturation aux données, codes sources et algorithmes, doivent concerner tous les niveaux, de l'administration centrale aux services déconcentrés, en commençant par l'encadrement supérieur et intermédiaire, lequel peut plus facilement entrer dans ces problématiques si elles sont liées à l'évaluation des politiques menées. Ce programme recensera, dans un premier temps, les initiatives de valorisation et de collaboration relatives aux codes et aux données puis élaborera un ensemble d'actions pouvant être déclinées dans les directions, les académies et les opérateurs du ministère : campagne d'affichage, conférences, sessions de mise en situation/jeux de rôle, journée de la donnée, page Pleiade « Politique de la donnée », etc.

Action 26-22S1 : Concevoir un « Programme d'acculturation aux données, algorithmes et codes » des agents en administration centrale et en services déconcentrés sur la base d'un bilan préalable (enquête qualitative et quantitative auprès des agents).

Action 27-22S1 : Acculturer les cadres à la démarche d'évaluation d'une politique publique pour l'éducation et à l'utilisation des données/indicateurs nécessaires au pilotage des politiques publiques.

II. Former les agents à la culture de la donnée

Les actions de sensibilisation et le travail d'acculturation doivent être poursuivis et pérennisés avec le développement d'actions de formation. Le rapport de la mission Bothorel recommande de prévoir un plan de formation aux enjeux du numérique non seulement pour les cadres dirigeants, mais aussi pour l'ensemble des niveaux hiérarchiques, avec un volet portant sur la politique d'ouverture des données et des codes sources. Le plan interministériel de formation proposé par la DGAFP et la DINUM doit être complété au niveau ministériel afin de prendre en compte plus spécifiquement les besoins du MENJS.

Action 28-22S1 : Inscrire des modules consacrés à l'Open Data et à la valorisation des données (pilotage par la donnée, data analyst, dataviz par exemple) dans le Plan national de formation de l'Éducation nationale (PNF), notamment en direction des cadres et relais de la formation dans ce domaine, et inciter les académies à reprendre ces formations dans leurs plans académiques de formation.

Action 29-22S1 : Assurer au niveau central et académique la publicité des formations interministérielles relatives aux enjeux du numérique, proposées par la DGAFP (notamment avec MENTOR) et la DINUM.

Fort de son expérience en matière de formation des enseignants, le MENJS peut contribuer à la production de contenus de formation avec le soutien à la réalisation d'un MOOC « Culture de la donnée » qui pourrait être largement diffusé sur la plateforme FUN. Les modalités de formation de type MOOC ont prouvé leur intérêt et leur efficacité. Ainsi, le MOOC développé par le Lab INRIA avec le soutien du ministère de l'Éducation nationale de la Jeunesse et des Sports comptait 27 000 inscrits en juin 2021 (ouverture en avril 2020) et près de 2 000 attestations d'achèvement. Un nouvel objectif est de concevoir, réaliser et diffuser un parcours dont le contenu viserait à faire comprendre ce qu'est une donnée, les types de données, leurs finalités, la mobilisation de l'Open Data et la réflexion sur le partage de communs et ses conditions. La cible serait les inspecteurs, les chefs d'établissement, les formateurs, les enseignants et leurs élèves, mais ce parcours permettra également l'acculturation et la formation de tous les acteurs, parents et citoyens en proposant une pédagogie de résolution de problèmes.

Action 30-22S1 : Proposer un parcours de formation et d'autoformation de type MOOC avec le Lab INRIA autour de trois modules principaux (Comprendre, Proposer-Exploiter l'Open Data, Relever des défis et éviter des biais, Construire et partager des cas d'usage) (12-24 mois).

III. Recruter et proposer des trajectoires attractives

En complément des questions d'acculturation et de formation des agents publics à la politique de la donnée et des codes sources, se pose la question des compétences disponibles pour mettre en œuvre cette politique, dans un contexte où tous les métiers du numérique au sein de l'État, et donc par capillarité au sein du MENJS, sont actuellement en tension.

Plusieurs actions doivent être menées pour renforcer l'attractivité du ministère et, en interne, l'intérêt pour les métiers de la donnée. Au-delà des actions habituelles sur la promotion des activités réalisées au sein du MENJS, l'attractivité du ministère peut ainsi reposer sur une politique dynamique de rapprochement avec de futurs ingénieurs formés à la data science. Ainsi, le recrutement d'entrepreneurs d'intérêt général (EIG) ou d'apprentis issus de grandes écoles peut ainsi contribuer à familiariser ces personnes avec les projets innovants du ministère, à constituer un réseau de personnes ressources pouvant ultérieurement rejoindre les équipes du MENJS tout en apportant de nouvelles compétences à ces équipes.

Action 31-21S2 : Recruter plusieurs EIG à la rentrée 2021 afin de participer à la construction du projet de Data Space européen sur la formation et les compétences dans le cadre de GAIA-X.

Action 32-22S1 : Favoriser le recrutement d'apprentis sur les compétences numériques autour de la donnée.

Les activités et les métiers autour de la donnée bénéficient d'un intérêt croissant de la part des collaborateurs des services informatiques et même au-delà (ainsi, un membre d'une délégation académique au numérique a pu développer le site <https://annuaire-Education.fr/> qui valorise et rassemble toutes les données ouvertes sur l'éducation). Certaines personnes ayant ainsi pu développer des compétences ou un fort intérêt pour les données, il serait pertinent de les recenser, en interrogeant tous les services académiques, de façon à créer un vivier de talents ou de futurs talents qui pourraient être mobilisés sur les nouveaux projets du ministère autour de la donnée.

Action 33-22S1 : Réaliser une cartographie des compétences en matière de données dans les académies.

IV. Animer des communautés au sein des administrations

Au-delà encore de l'acculturation et de la formation, le ministère doit s'appuyer sur ses personnels en valorisant et animant des communautés qui participent activement à la politique de la donnée et des codes. Les services doivent aussi pouvoir compter sur une personne ressource, qui participera à la conduite de la politique de la donnée déclinée au sein de leur service. Ces agents, réunis au sein de communautés (territoriales, fonctionnelles), bénéficieront des benchmark, kits, conseils, de leur communauté. Leur objectif sera également d'alerter et de conseiller le comité stratégique de la donnée.

Le PGO propose dans ce cadre une nouvelle action portant sur la création et l'animation du réseau des référents « données et codes sources » intégrant notamment l'ensemble des services déconcentrés (Régions académiques) et des opérateurs du MENJS.

L'animation nationale autour des JPIL (Journées de Pilotage), organisées régulièrement par la DNE, permettra de rassembler les référents et les personnes ressources sur l'ouverture et la circulation des données, l'objectif étant de présenter des sujets innovants autour de la donnée (acteurs privés ou interministériels), de faire des retours d'expérience sur des projets locaux et d'apporter une cohérence nationale autour des outils décisionnels liés à la donnée.

Action 34-22S1 : Désigner des référents « données » dans chaque direction et opérateur du ministère.

Action 35-22S1 : Organiser des journées trimestrielles sur les grandes problématiques du partage des données (dispositifs d'anonymisation, mise en qualité des données, normes et standard d'interopérabilité, etc.).

Mobiliser et évaluer : gouvernance et pilotage

I. S'appuyer sur l'Administrateur ministériel des données

Placé auprès du Directeur du numérique pour l'éducation, l'Administrateur ministériel des données (AMD) a été nommé en septembre 2019 et est chargé d'animer et de coordonner la politique ministérielle d'ouverture des données, qu'il s'agisse des données produites ou collectées par le ministère ou des données, de nature privée, produites lors de la mise en œuvre du service public de l'éducation.

Son action s'inscrit dans la stratégie de transformation numérique du MENJS portée par la DNE (protéger et valoriser les données scolaires).

- L'AMD a été chargé d'organiser et d'assurer le suivi d'une véritable gouvernance de l'ouverture des données du MENJS, qui s'articule autour des éléments suivants : association de toutes les directions aux choix et aux décisions d'ouverture des données dont elles ont la responsabilité ; mise en place d'une procédure d'instruction, systématique et collective, pour chaque proposition ou demande d'ouverture et de diffusion de données ; mise en place d'un Comité stratégique de la donnée, présidé par la Secrétaire Générale et comprenant l'ensemble des directeurs.
- L'AMD est chargé de mener des actions destinées à valoriser les données scolaires numériques. La valorisation des données scolaires issues de la généralisation des services et des outils numériques, qui sont utilisés de façon croissante pour les apprentissages, les évaluations, l'orientation, la vie scolaire, est un volet novateur et essentiel de la politique ministérielle des données. Dans le respect du cadre protecteur des données personnelles, la production, la collecte et l'analyse des données et notamment des données d'apprentissage (Learning analytics) doivent être mises en œuvre à travers une série d'actions structurantes (pilotage du projet de Data Hub pour les données scolaires par exemple).
- L'AMD assure un rôle de coordination pour le développement de l'intelligence artificielle dans l'éducation. L'IA représente un outil d'analyse des données massives contribuant ainsi au développement des Learning analytics mais, plus largement, c'est un domaine qui doit irriguer l'ensemble du système éducatif, qu'il s'agisse de la formation des élèves et des enseignants, de la production de nouveaux outils pédagogiques innovants utilisant ses apports ou encore des questions juridiques et éthiques soulevées par son utilisation.

Action 36-21S2 : Mettre en œuvre la feuille de route de la Politique ministérielle des données pour le ministère de l'Éducation, de la Jeunesse et des Sports.

II. Assurer la mobilisation au sein du ministère et développer les synergies

La gouvernance de l'ouverture des données au sein du MENJS

Le MENJ a mis en place initialement une **gouvernance dédiée à l'Open Data** :

- **Création d'un Comité de Pilotage ministériel Open Data** à partir de 2015, réunissant sous la présidence de la Secrétaire Générale les directeurs et des représentants des opérateurs et d'académies. Ce comité est une instance transversale d'échanges, de décision et d'impulsion sur la stratégie Open Data en matière d'organisation, de structuration et de publication de données. Il est réuni 2 fois par an.
- Désignation d'un référent correspondant ministériel Open Data sis auprès de la Secrétaire Générale
- **Mise en place d'un Réseau de correspondants Open Data dans chaque direction de l'administration centrale, opérateur EN et chaque académie.** Ce réseau est réuni régulièrement dans le cadre de réunions dédiées : atelier données et réunion « ouverture algorithmes/codes sources.

Cette gouvernance s'articule étroitement avec la gouvernance générale liée à la politique générale de la donnée.

La politique des données du MENJS repose, en effet, sur plusieurs entités chargées de :

- La protection des données (DPD, Direction des affaires juridiques, FFSI, Direction du numérique pour l'éducation) ;
- L'ouverture, la diffusion et la valorisation des données (Référént Open Data Secrétariat général, AMD, DEPP, DAJ)
- Les échanges de données dans le cadre de la dématérialisation des démarches administratives (SG, Direction générale de l'enseignement scolaire (DGESCO), DNE).

À la faveur de la création de la fonction d'AMD en septembre 2019, un réexamen de la gouvernance des données a été entrepris et un certain nombre de décisions ont été actées en février 2020.

Plusieurs principes sont désormais mis en œuvre :

- Association des directions aux choix et aux décisions d'ouverture des données dont elles ont la responsabilité ;
- Mise en place d'un comité stratégique de la donnée présidé par la Secrétaire Générale et réunissant les directeurs d'administration centrale pour les orientations et les arbitrages ;
- Maintien du comité de pilotage Open Data réunissant également les académies et les opérateurs afin d'assurer une information et une cohérence des ouvertures des données.

Enfin, le MENJS a instauré en 2019 un Comité d'éthique pour les données d'éducation chargé de conduire des réflexions sur l'utilisation de ces données, afin notamment de garantir un juste équilibre entre valorisation et protection des données, et d'émettre des avis et des recommandations. Il est composé de parlementaires, de chercheurs, d'universitaires et des responsables associatifs.

III. Assurer la bonne gouvernance et le pilotage de la feuille de route

Le Comité stratégique de la donnée

Le Comité stratégique de la donnée, présidé par la Secrétaire Générale et comprenant tous les directeurs, s'est réuni pour la première fois en janvier 2021. Il a vocation à se réunir deux à trois fois par an. C'est une instance de décisions et d'arbitrage qui doit permettre de façon très opérationnelle l'articulation entre les différents acteurs, les directions, la DNE avec l'administrateur des données, le DPD, la DAJ, et le Secrétariat général au titre de l'Open Data.

Le Comité veille également à l'alignement des différentes feuilles de route et définit, de façon collégiale, les priorités stratégiques du ministère en matière de politique de la donnée, notamment toute la problématique de l'exploitation des données que produit le ministère dans ses différentes composantes.

Le périmètre de son action est large, avec l'accès, le partage, la valorisation des données, sans y intégrer la protection des données, même si cette dimension reste bien identifiée. Il intègre également les thématiques jeunesse et sports dont les directeurs sont invités à ce comité. Sur le type de données, sont concernées les données du ministère mais aussi les données externes dans le cadre du service public de l'éducation, celles des éditeurs, des collectivités locales avec lesquelles des échanges sont importants.

Le Comité stratégique de la donnée s'articule avec les autres comités du MENJS : Comité d'éthique des données d'éducation, comité de pilotage Open Data qui intègre des représentants des académies et les opérateurs, et comité de suivi de la protection des données.

Cette gouvernance de la politique d'ouverture des données du MENJS contribue ainsi à créer un cadre de confiance avec l'instauration de règles claires, explicites et partagées permettant de répondre aux objectifs de la politique des données d'éducation.

Action 37- 22S1 : Organiser l'articulation entre les différents comités traitant des données dans l'objectif de la mise en œuvre des engagements de la feuille de route.

IV. Entretenir le dialogue avec les communautés de réutilisateurs et écosystèmes

Le MENJS dispose d'un grand nombre de dispositifs permettant le dialogue avec les communautés de réutilisateurs et les écosystèmes, au premier rang desquels figurent les différentes associations représentatives comme les associations de collectivités locales, réunies par la DNE au sein d'un Comité des partenaires (AMRF, ADF, Régions de France, etc.), ou les associations représentant les entreprises de la EdTech (AFINEF, EdTech France notamment). La Délégation à la communication comme la DEPP permettent également la remontée des demandes et des besoins des usagers et des utilisateurs des données.

Afin de donner une nouvelle impulsion à ces dialogues autour des réutilisations des données, plusieurs actions peuvent être menées parallèlement. Il s'agit d'abord de mieux associer les territoires aux problématiques de circulation des données, qu'il s'agisse des services déconcentrés des Régions académiques ou des collectivités elles-mêmes, qui recueillent nombre de données et souhaitent également approfondir les partages de données avec l'Éducation nationale.

La création de la Plateforme des données d'éducation comme le projet GAIA-X vont s'accompagner de la mise en place d'instances de gouvernance, faisant une large place aux représentants de l'écosystème dont les entreprises intéressées par l'utilisation des données du système éducatif.

Enfin, en relation avec les actions prévues dans le cadre du PGO, la société civile, avec les associations, les initiatives d'intérêt général et les citoyens doit pouvoir exprimer ses attentes et des besoins à travers des actions d'animation autour de l'utilisation et de la réutilisation des données.

Action 38-22S2 : Mettre en place une gouvernance régionale de la donnée dans les régions académiques sous l'égide des secrétaires généraux de région académique.

Action 39-22S1 : Initier une mission d'évaluation de l'Inspection générale de l'Éducation, du Sport et de la Recherche (IGESR) portant sur les conventions d'échange de données entre les académies et les collectivités territoriales (état des lieux, pratiques et préconisations).

Action 40-22S1 : Créer un groupe technique de la donnée dans le cadre du Comité des partenaires afin d'échanger sur les questions de circulation des données entre les acteurs publics.

Action 41-21S2 : Mettre en place des événements réguliers, associant la société civile, permettant de faciliter la manipulation et l'exploitation des données ouvertes brutes (ex : hackatons dataviz).

Ressources et moyens

Le MENJS consacre actuellement environ 4 ETP directement à l'Open Data. Pour la DNE et les directions métiers, on peut estimer que 2 ETP prennent en charge les volets fonctionnels et techniques d'ouverture des données (dont un ½ ETP pour la gestion et le déploiement des données en Open Data utilisés). À la DEPP, en dehors de leur activité d'élaboration des données qui n'est pas comptée ici, tous les chargés d'études de la DEPP doivent, au moment de la mise en ligne de leurs données en Open Data, s'assurer de la qualité de celles-ci. Une personne est chargée de la coordination de cette mission. Enfin, chaque année, les chargés d'étude doivent réfléchir et construire de nouvelles bases. Compte tenu du nombre important de fichiers mis en ligne par la DEPP, on peut estimer, en fourchette basse, le nombre d'ETP mobilisés à 2.

La plateforme utilisée, dédiée à l'Open Data, présente un coût annuel d'environ 70 K€ (environnement des données pour le national et les régions académiques); ce service est à la fois utilisé pour la mise à disposition des données (Open Data strict) mais également pour le développement de service, en s'appuyant sur les API produites nativement par la plateforme.

Le pilotage de la politique de la donnée, correspondant aux missions de l'AMD, peut être associé à 1,5 ETP, l'AMD adjointe ayant une part importante de son activité consacrée à l'Open Data (et comptabilisée supra).

Enfin, un grand nombre de collaborateurs de la DNE, estimé à une quarantaine (dont la moitié en tant que prestataires), travaillent actuellement sur le décisionnel.

Pour atteindre les objectifs et les actions fixés dans la feuille de route, le MENJS s'appuiera sur les compétences et les montées en compétences de ces équipes du décisionnel.

Annexe

N°	ACTIONS	RESPONSABLE	ÉCHÉANCE
Action 01	Développer la mise à disposition de jeux de données issus des services statistiques (DEPP) en Open Data, notamment pour les données qui font l'objet de demandes récurrentes du public.	DEPP	2021 semestre 2
Action 02	Développer un tableau de bord de la transformation numérique à destination des principaux acteurs du ministère et du grand public permettant d'évaluer les différentes composantes des politiques, autour de l'équipement des établissements, la formation des personnels, l'utilisation du numérique et ses impacts dans les établissements.	DNE	2022 semestre 2
Action 03	Faire évoluer le plan de marquage (mesure d'audience) des Espaces Numériques de Travail (à date, 60% des établissements disposent d'un ENT avec 5 millions d'élèves, 400 000 enseignants et 6 millions de parents sont dans le dispositif national de mesure d'audience) afin de mieux connaître et d'évaluer les usages des services et des ressources numériques, de suivre les parcours utilisateur et d'étendre le dispositif au 1 ^{er} degré (2 millions d'élève, en forte croissance).	DNE	2021 semestre 2
Action 04	Publier les données résultant des mesures d'audience des ENT (nouveau dispositif) en Open Data.	DNE	2023 semestre 1
Action 05	Mettre en Open Data l'ensemble des données publiées sur les sites institutionnels du MENJS après information de la direction responsable des données (Éducation.gouv.fr et eduscol.fr).	AMD	2021 semestre 2
Action 06	Ouvrir de nouveaux jeux de données en s'appuyant sur un dispositif innovant permettant la remontée des besoins du terrain et leur analyse.	AMD	2021 semestre 2
Action 07	Confier aux référents Open Data dans les directions et les académies un rôle de facilitateur d'accès à la donnée.	AMD	2021 semestre 2
Action 08	Intégrer les données Jeunesse et Sport dans l'Open Data du ministère.	AMD	2022 semestre 1
Action 09	Faciliter l'accès aux données du système éducatif pour les chercheurs dans un cadre sécurisé.	DEPP	2022 semestre 1
Action 10	Créer une plateforme des données scolaires pour partager ces données avec les chercheurs et les entreprises de la EdTech (création de la plateforme dans 12 à 18 mois)	AMD	2022 semestre 2
Action 11	Être en mesure de bénéficier des données scolaires collectées par les acteurs privés de la EdTech (solutions de vie scolaire notamment) après analyse des conditions juridiques de ces transferts vers les académies et l'administration centrale.	SG	2022 semestre 1
Action 12	Poursuivre l'ouverture des jeux de données du MENJS et étendre le champ des domaines couverts (offre de formation, internats, offre de formation des GRETA, référentiels des diplômes professionnels, etc.) (6 à 12 mois).	AMD	2022 semestre 1
Action 13	Ouvrir une plateforme Open Data à l'échelon de chaque Région académique en articulation avec la plateforme Open Data nationale.	AMD	2022 semestre 1
Action 14	Construire un catalogue des APIs du MENJS et développer des APIs ouvertes.	DNE	2022 semestre 1

N°	ACTIONS	RESPONSABLE	ÉCHÉANCE
Action 15	Définir un cadre d'usage des APIs par les consommateurs.	DNE	2022 semestre 1
Action 16	Associer les directions métiers à la gouvernance et à la mise en œuvre de la stratégie d'APIsation du ministère.	DNE	2022 semestre 1
Action 17	Développer une API Données de scolarité au profit des autres administrations dans le cadre de l'objectif « Dites-le nous une fois ».	DNE	2022 semestre 1
Action 18	Mettre en place une plate-forme d'ouverture, de partage et de valorisation des données de gestion afin de permettre aux directions métiers et aux académies d'améliorer leur pilotage opérationnel.	DNE	2022 semestre 2
Action 19	Préparer la mise en place du « Data Space Éducation et compétences » au niveau français et européen en partenariat avec les entreprises de la EdTech (gouvernance, premiers cas d'usages, groupes de travail thématiques, etc.) dans le cadre du projet européen GAIA-X.	DNE	2021 semestre 2
Action 20	Mettre en œuvre la stratégie IA pour l'éducation.	DNE	2023 semestre 1
Action 21	Développer de nouveaux partenariats d'innovation pour l'IA (P2IA) en les étendant aux principaux enseignements du collège et du lycée.	DNE	2022 semestre 1
Action 22	Publier les codes sources des applications sur le mouvement des enseignants (1 ^{er} et 2 ^d degrés).	SEMSIRH	2022 semestre 2
Action 23	Réaliser un état des lieux des algorithmes et codes sources à publier en priorité (en fonction des attentes des usagers).	AMD	2022 semestre 1
Action 24	Adopter des démarches de développement permettant la publication du code source dès la conception, soit par réorganisation des équipes de développement internes au ministère, soit en intégrant des clauses ad-hoc dans les appels d'offre.	DNE	2022 semestre 2
Action 25	Concevoir un « Open Source Program Office » sectorisé pour l'éducation, doté d'indicateurs, afin de fournir et de mutualiser des solutions, des ressources et des services pour l'ensemble de la communauté éducative (6 à 12 mois).	DNE	2022 semestre 2
Action 26	Concevoir un « Programme d'acculturation aux données, algorithmes et codes » des agents en administration centrale et en services déconcentrés sur la base d'un bilan préalable (enquête qualitative et quantitative auprès des agents).	AMD	2022 semestre 1
Action 27	Acculturer les cadres à la démarche d'évaluation d'une politique publique pour l'éducation et à l'utilisation des données/indicateurs nécessaires au pilotage des politiques publiques.	SG /AMD	2022 semestre 1
Action 28	Inscrire des modules consacrés à l'Open Data et à la valorisation des données (pilotage par la donnée, data analyst, dataviz par exemple) dans le Plan national de formation de l'Éducation nationale (PNF), notamment en direction des cadres et relais de la formation dans ce domaine, et inciter les académies à reprendre ces formations dans leurs plans académiques de formation.	DGRH	2022 semestre 1
Action 29	Assurer au niveau central et académique la publicité des formations interministérielles relatives aux enjeux du numérique, proposées par la DGAFP (notamment avec MENTOR) et la DINUM.	DELCOM	2022 semestre 1
Action 30	Proposer un parcours de formation et d'autoformation de type MOOC avec le Lab INRIA autour de trois modules principaux (Comprendre, Proposer-Exploiter l'Open Data, Relever des défis et éviter des biais, Construire et partager des cas d'usage) (12-24 mois).	DNE	2022 semestre 1

N°	ACTIONS	RESPONSABLE	ÉCHÉANCE
Action 31	Recruter plusieurs EIG à la rentrée 2021 afin de participer à la construction du projet de Data Space européen sur la formation et les compétences dans le cadre de GAIA-X.	DNE	2021 semestre 2
Action 32	Favoriser le recrutement d'apprentis sur les compétences numériques autour de la donnée.	DNE	2022 semestre 1
Action 33	Réaliser une cartographie des compétences en matière de données dans les académies.	DNE	2022 semestre 1
Action 34	Désigner des référents « données » dans chaque direction et opérateur du ministère.	AMD	2022 semestre 1
Action 35	Organiser des journées trimestrielles sur les grandes problématiques du partage des données (dispositifs d'anonymisation, mise en qualité des données, normes et standard d'interopérabilité, etc.).	AMD	2022 semestre 1
Action 36	Mettre en œuvre la feuille de route de la Politique ministérielle des données pour le ministère de l'Éducation, de la Jeunesse et des Sports.	SG/AMD	2021 semestre 2
Action 37	Organiser l'articulation entre les différents comités traitant des données.	SG	2022 semestre 1
Action 38	Mettre en place une gouvernance régionale de la donnée dans les régions académiques sous l'égide des secrétaires généraux de région académique.	SG	2022 semestre 2
Action 39	Initier une mission d'évaluation de l'Inspection générale de l'Éducation, du Sport et de la Recherche (IGESR) portant sur les conventions d'échange de données entre les académies et les collectivités territoriales (état des lieux, pratiques et préconisations).	SG	2022 semestre 1
Action 40	Créer un groupe technique de la donnée dans le cadre du Comité des partenaires (collectivités territoriales) afin d'échanger sur les questions de circulation des données entre les acteurs publics.	DNE	2022 semestre 1
Action 41	Mettre en place des événements réguliers, associant la société civile, permettant de faciliter la manipulation et l'exploitation des données ouvertes brutes (ex : hackatons dataviz).	AMD	2021 semestre 2
Action 42	Renforcer le déploiement de cas d'usage au sein du Sport data Hub dans les domaines du pilotage stratégique, de l'optimisation de la performance et de la gestion opérationnelle	DS	2023 semestre 2



**MINISTÈRE
DE L'ÉDUCATION
NATIONALE,
DE LA JEUNESSE
ET DES SPORTS**

*Liberté
Égalité
Fraternité*





**Convention cadre passée entre le ministre de l'éducation nationale,
de l'enseignement supérieur et de la recherche
et
l'Association des Régions de France
portant sur les échanges d'informations statistiques nécessaires au pilotage
du système éducatif**

Date de notification : 27.08.2015

Convention n° 2015-DEPP-023

LA PRÉSENTE CONVENTION EST ÉTABLIE ENTRE :

LE MINISTÈRE DE L'ÉDUCATION NATIONALE, DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE

Représenté par Catherine MOISAN, Directrice de l'évaluation, de la prospective et de la performance (DEPP),

Ci-après désigné par le terme « MENESR »

D'une part,

ET L'ASSOCIATION DES RÉGIONS DE FRANCE

Représentée par Alain ROUSSET, son Président,

Ci-après désignée par le terme « ARF »

D'autre part,

Vu le Code de l'éducation,
Vu le Code général des collectivités locales,
Vu le Code du travail,

IL EST CONVENU CE QUI SUIV

Préambule

La présente convention s'inscrit dans une volonté commune de renforcer le partenariat entre l'Etat et les Régions en respectant le cadre des compétences partagées en matière éducative. Elle doit être le vecteur d'un rapprochement entre les Présidents de Régions et les Recteurs d'académie pour pouvoir disposer conjointement d'une information de qualité sur le système éducatif.

Elle se situe dans un contexte d'affirmation des compétences des Régions dans le domaine de la formation et de l'orientation. Elle fait le constat des besoins accrus des collectivités en informations opérationnelles et décisionnelles pour construire des politiques territoriales et des conditions d'un dialogue positif avec les autorités académiques autour de compétences partagées. Elle insiste sur l'idée d'échange de données et de services entre partenaires institutionnels, Etat et Région, de la formation.

En ce sens, cette convention a donc également pour objectif de promouvoir et de faciliter la collaboration de l'Etat et des Régions en matière de mise en cohérence de leurs systèmes d'information.

La coordination des actions engagées de part et d'autre sur les territoires nécessite, en effet, de renforcer les échanges d'informations, en volume et en qualité, de façon à permettre aux différents décideurs de mieux dialoguer afin de construire des politiques, cohérentes et efficaces, à partir de données sélectionnées, documentées et communément reconnues.

En outre, une mise en commun des ressources pour la réalisation d'études thématiques intéressant les Régions et les Académies doit aussi être encouragée et si possible être développée avec d'autres partenaires présents sur le territoire.

Article 1 : Objet de la convention

Les éléments répertoriés dans le cadre de cette convention doivent servir de référence à la contractualisation des échanges de données entre les Présidents de Région et les Recteurs d'académie, agissant au nom du MENESR.

Cette convention cadre concourt à l'élaboration des conventions locales, sur un volume d'échange minimal nécessaire au bon fonctionnement du système éducatif. Elle est conçue, notamment dans son annexe technique, comme étant non limitative ou non exclusive, adaptable en fonction de la capacité à mobiliser certaines données ou de construire certains partenariats répondant aux souhaits des Recteurs et Présidents de Régions d'aller au-delà des échanges explicitement cités dans cette convention cadre.

Pour ce faire, il est important que, comme l'Etat reste le garant de la cohérence des systèmes d'information, les Régions puissent s'engager à garantir la cohérence, voire à terme l'interopérabilité, de leurs bases de données avec les référentiels (nomenclatures et répertoire établissements) administrés par le MENESR. En ce sens, le MENESR met à la disposition des Régions, directement et par l'intermédiaire de l'ARF, ses référentiels et la documentation dont il dispose. De même, il assiste, conseille et contribue à la formation des agents des Régions dans l'usage de ces référentiels.

Le MENESR informe les Régions, au plus tôt, des modifications des référentiels (contenus et structures), afin de permettre en amont une bonne adaptation des systèmes d'information des collectivités.

En outre, les restrictions aux échanges sont principalement liées aux règles du secret statistique et à l'engagement exprès de non cession des informations à des tiers sans accord préalable du président de Région ou du recteur d'académie.

La description technique des données fait l'objet d'une annexe à la convention.

Cette annexe technique distingue quatre parties : les données mises à disposition par la DEPP, les données mises à disposition par toutes les académies, les données susceptibles d'être mise à dispositions par des académies, les données mises à disposition par les Régions.

Article 2 : Accès aux Informations produites par le MENESR – Besoins des Régions

Les Régions, dans le cadre de leurs champs de compétences obligatoires en matière d'éducation et de formation tout au long de la vie ou dans le cadre de leurs politiques éducatives propres, ou en accompagnement de politiques éducatives académiques, ont besoin d'accéder à différents niveaux de sources d'informations détenues par l'éducation nationale : informations synthétiques, accès aux bases de données nationales et académiques, communication de fichiers standards de nature statistique ou opérationnel.

Pour rappel, les principales compétences Régionales concernées sont :

A-) Pour la planification des formations

La responsabilité des Régions a été réaffirmée sur l'ensemble des besoins de formation à travers notamment, les articles L214-1 à L214-4, et L214-13-1 du code de l'éducation.

Les Régions ont en charge l'adoption des schémas prévisionnels des formations des collèges, des lycées, des établissements d'éducation spéciale et des établissements d'enseignement agricole.

Elles arrêtent la carte des formations professionnelles initiales.

Elles définissent des schémas Régionaux de l'enseignement supérieur, de la recherche et de l'innovation qui s'articulent avec les contrats de plans Régionaux de développement des formations professionnelles.

B-) Pour les établissements d'enseignement, principalement les lycées

Les Régions définissent les programmes prévisionnels des investissements relatifs aux établissements scolaires relevant de leur compétence.

Elles assurent en matière patrimoniale, la construction ou reconstruction, les extensions, les grosses réparations et la maintenance du bâti, l'achat et la maintenance des équipements pédagogiques, ainsi que la maintenance informatique des établissements.

Elles ont en charge les dotations financières nécessaires au fonctionnement des établissements, ainsi que les fonctions d'accueil, de restauration scolaire et d'hébergement, et l'entretien général et technique des établissements.

En outre, la loi de refondation de l'école prévoit (article L421-4 du code de l'éducation) la possibilité de mise en œuvre de conventions tripartites Etablissement/Etat/Région dans le cadre du contrat d'objectif, et celle d'autoriser l'utilisation des locaux scolaires au bénéfice d'activités culturelles et artistiques, d'éducation populaire ou de formation.

C-) Pour la formation professionnelle et l'apprentissage

Les Régions définissent et mettent en œuvre les politiques d'apprentissage et de formation professionnelle, en particulier au moyen du contrat de plan Régional de développement des formations professionnelles.

D-) Les domaines mentionnés au présent article pourront être complétés en tant que de besoin au gré des évolutions que pourrait introduire la loi portant Nouvelle Organisation des Territoires de la République (NOTR), en particulier la dévolution aux Régions des transports scolaires d'une part, et la responsabilité conjointe des Régions avec les Recteurs sur la définition des districts des lycées d'autre part.

L'annexe technique décrit en partie 1 les données mises à disposition par la DEPP, en partie 2 les données mises à disposition par toutes les académies, en partie 3 les données susceptibles d'être mise à dispositions par des académies.

Article 3 : Accès aux informations produites par la Région – Besoins du MENESR

Il convient tout d'abord de rappeler que les Régions sont tenues de respecter les dispositions de l'article 130 de la loi du 13 août 2004 complétées par le décret n° 2007-572 du 18 avril 2007 relatif à la transmission à l'Etat par les collectivités territoriales d'informations statistiques relatives à l'accueil, la restauration, l'hébergement et l'entretien dans les collèges et lycées.

Celles-ci, doivent, en particulier, veiller à renseigner de façon exhaustive et particulièrement qualitative l'« Enquête EPI » dont les résultats statistiques sont, une fois exploités, mis, en retour, à leur disposition.

L'annexe technique décrit en partie 4 les données mises à disposition par les Régions. En outre et dans le cadre conventionnel les services académiques peuvent selon leurs besoins demander à la Région la transmission d'informations supplémentaires.

Article 4 : Transmisslon des données et règles d'usages

La mise à disposition, selon la complexité des données, se fait par accès direct sur internet (données publiques) ou via un accès réservé, aux bases nationales du MENESR : BCN, RAMSESE, ACCÉ, BCP (univers tels que définis dans l'annexe technique), ICÔTEP, IVAL, APAE, ou par transmission de fichiers.

Les données ne seront pas utilisées pour des opérations de gestion individuelles des élèves.

Les données ne seront pas cédées à un tiers, que ce soit à titre gracieux ou payant, excepté dans le cadre d'une étude ponctuelle confiée à un prestataire. Dans ce dernier cas il sera fait obligation au prestataire de détruire les fichiers concédés à la fin de l'étude.

Hors les données nominatives décrites dans l'annexe technique de la présente convention cadre, les données nominatives ne sont pas a priori exclues des conventions locales, mais devront faire l'objet d'accords spécifiques, conformes aux prescriptions de la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés en matière de communication de données à caractère directement ou indirectement personnel.

Les données individuelles, non nominatives, transmises aux Régions, ne seront pas accompagnées de l'INE, mais pourront être accompagnées d'un identifiant crypté.

Les fichiers transmis seront conformes aux lois et réglementations sur l'obligation et le secret en matière de statistiques et d'informatique actuellement en vigueur : loi n° 51-711 du 7 juin 1951 sur l'obligation, la coordination et le secret en matière statistique, loi n° 78-17 du 6 janvier 1978 relative à l'informatique aux fichiers et aux libertés. Les données ainsi fournies ne pourront en aucun cas permettre d'identifier des individus.

Ces données seront réservées aux agents habilités et leur sécurité sera assurée contre les intrusions ou les accès frauduleux.

Les publications utilisant les données transmises en mentionneront la source, particulièrement en faisant figurer dans chaque tableau ou figure utilisant les données fournies l'origine des données utilisées selon le format défini [sources : MENESR, DEPP ou académie de ..., ou Région ..., Système d'information ...].

Les partenaires s'engagent à ne pas lancer d'enquêtes en direction des établissements scolaires dès lors que les informations recherchées sont disponibles dans les bases de données ouvertes ou transmises.

Article 5 : Outils collaboratifs partagés

Afin de renforcer la qualité du dialogue et de l'exercice des compétences partagées, il est recommandé de construire des partenariats, mis en œuvre aux niveaux national, régional ou mixte, permettant de mettre à disposition des décideurs des informations de qualité, documentées et sélectionnées en commun.

Cela concerne en particulier :

- La formation des agents des collectivités territoriales à l'usage des données échangées
- La participation de représentants des Régions à la définition et à l'amélioration des enquêtes statistiques de la DEPP
- Le développement conjoint et la maintenance conjointe d'outils de pilotage, notamment par la mise en commun possible de ressources humaines et financières
- Le développement de plateformes Régionales d'échange des données géographiques
- Le développement d'études menées en commun portant par exemple sur des études contextualisées des territoires infrarégionaux.

Article 6 : Date d'effet, durée et renouvellement de la convention, mode de résiliation

La présente convention prend effet à sa date de notification pour une durée de validité d'une année renouvelable par tacite reconduction à l'issue d'une réunion de bilan entre les signataires permettant de constater les efforts de partenariats entre les Régions et les académies et de renforcer éventuellement les aspects techniques des échanges de données.

La présente convention peut être dénoncée par l'une ou l'autre des parties au moyen d'une lettre recommandée avec accusé de réception. Cette résiliation prend effet une année après réception de cette lettre.

Article 7 : Conditions financières

a-) Les échanges de données ne donnent pas lieu à des contreparties financières.

b-) Le MENESR, Direction de l'Évaluation de la Prospective et de la Performance et l'ARF peuvent convenir de financer en commun ou de participer au financement du développement d'outils liés à l'information décisionnelle partagée ainsi qu'à des études statistiques utiles à l'ensemble des parties.

• •
•

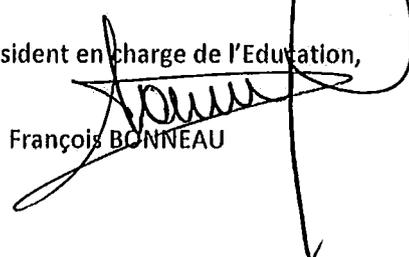
Fait à Paris en trois exemplaires originaux, le 21 août 2015

Pour l'Association des Régions de France

Le Président


Alain ROUSSET

Le Vice-Président en charge de l'Éducation,


François BONNEAU

Pour le Ministère de l'éducation nationale, de
l'enseignement supérieur et de la recherche



La Directrice de l'évaluation, de la prospective et
de la performance

Catherine MOISAN

ANNEXE TECHNIQUE Description des données mentionnées dans la convention
--

1- Données mises à disposition par la DEPP

1.1 - Base centrale des nomenclatures (BCN)

Accès en consultation et chargement de fichiers par internet sans restriction (téléchargement et extraction par web service)

Les nomenclatures constituent le langage commun du système d'information du ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche : nomenclatures propres au ministère ainsi que nomenclatures interministérielles et internationales qui lui sont nécessaires. Elles sont mises à jour régulièrement et historisées au moyen de dates d'ouverture et de fermeture.

1.2 - Répertoire académique et ministériel sur les établissements du système éducatif (RAMSESE)

RAMSESE recense les établissements assurant une activité de formation initiale générale, technologique ou professionnelle, de la maternelle à l'enseignement supérieur, qu'ils soient publics ou privés, sous tutelle ou non du ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche ainsi que les structures d'administration du système éducatif public.

L'accès aux données de RAMSESE reste pour les Régions provisoirement fourni via la BCN (vue BCE_UAI).

L'accès aux données de RAMSESE n'est pas limité à l'aire géographique stricte de la collectivité, ni aux établissements relevant strictement des compétences de la Région.

1.3 - Application de consultation et cartographie des établissements du système éducatif français (ACCÉ)

Accès réservé, sur inscription.

La Région disposera d'un accès à l'application de consultation et cartographie des établissements du système éducatif français, qui est l'interface rénovée de la base centrale des établissements (BCE). ACCÉ est alimentée quotidiennement par RAMSESE. Elle permet de faire des requêtes, de les enregistrer et de les partager avec d'autres utilisateurs.

1.4 - Base centrale de pilotage (BCP)

Accès réservé.

L'ensemble des univers des domaines suivants :

Elèves du premier degré

Elèves du second degré et classe post-bac

Les apprentis

Les étudiants
Les examens
Les établissements
Domaine transversal et contexte
Formation continue universitaire

NB :

Dans le cas d'une agrégation par établissement, les données relatives aux établissements privés hors contrat peuvent ne pas être accessibles.

Dans le cas où la variable « nationalité » est présente dans certains de ces univers, celle-ci est regroupée (UE, hors UE, français).

La cession de listes d'établissements à des tiers n'est pas autorisée (sauf dans le cadre d'une prestation).

L'accès à la BCP n'est pas limité à l'aire géographique stricte de la collectivité.

1.5 - APAE - Aide au pilotage et à l'auto-évaluation des établissements

Accès réservé.

L'application d'aide au pilotage et à l'auto-évaluation des établissements (APAE) met à la disposition de la communauté éducative des informations et des indicateurs pour chaque établissement du second degré, public ou privé sous contrat, en France métropolitaine ou dans les DOM. Ces indicateurs concernent la population scolaire accueillie, les personnels et les moyens ainsi que la performance.

1.6 - IVAL - Indicateurs de valeur ajoutée des lycées

Accès en consultation et chargement de fichiers par Internet sans restriction (téléchargement, web service, open data, etc.)

A côté du taux brut de réussite au baccalauréat, 3 indicateurs de valeur ajoutée dans les lycées permettent de tenir compte du profil de la population accueillie par l'établissement par rapport à des moyennes académiques ou nationales. Le fichier IVAL comprend aussi les valeurs brutes.

1.7 - Prévisions d'effectifs du second degré :

Les prévisions nationales sont publiées tous les ans par la Depp pour les années n et n+1. Pour les Régions comportant plusieurs académies, la Depp réalise les matrices Régionales correspondantes.

2 - Fichiers académiques mis à disposition par chaque Rectorat

2.1 - Annuaire restreint des personnels

Accès réservé.

Mise à disposition d'un fichier

Champ : Chefs d'établissement (établissements relevant de la compétence Régionale) publics et privés sous contrat de l'académie ; Adjoint, chefs de travaux, gestionnaires et agents comptables pour les lycées, LP et EREA publics de l'académie.

Variables :

Année de rentrée scolaire

Code de l'unité administrative immatriculée (UAI)

Fonction : (principal, directeur, proviseur, proviseur-adjoint, gestionnaire, agent comptable, chef des travaux)

Nom, prénom, civilité, mail personnel, téléphone personnel (si possible)

Date de prise de fonction

Un numéro d'identifiant

Le Conseil Régional s'engage à limiter les usages de ces données à de la communication Institutionnelle.

2.2 - Fichier individuel anonyme des élèves et des formations suivies du 2nd degré à la date du constat de rentrée (F_CARAC_ELEV_aaaa)

➤ Fichier anonyme des élèves et des formations suivies (fourni sur demande).

Champ : enseignement secondaire de la Région sous tutelle du MEN, public et privé sous contrat, tous établissements du champ SCOLARITE, collèges et lycées (Enquête 16-17-18).

Variables :

Caractéristiques des élèves et de leur famille

Identifiant anonymisé

Sexe

Année de naissance

Nationalité regroupée (France, autres pays de l'UE, autres pays d'Europe, continents)

Catégorie socioprofessionnelle du responsable principal (Origine sociale détaillée)

Lien de parenté avec le responsable principal

Département de résidence

Régime scolaire (Interne, demi-pensionnaire, ...)

Boursier (oui/non)

Formation suivie

Année de rentrée scolaire

N° d'établissement

Statut de l'élève (scolaire, apprenti, ...)

Division

Type de contrat de la division (pour le privé)

Codes formation correspondant à la division

MEF

MEFST11

Lib MEF

Options choisies (jusqu'à 12) MATIERE_ENSEIGNE

Formation suivie l'année précédente

N° d'établissement
Statut de l'élève (scolaire, apprenti, ...)
Codes formation correspondant à la division
MEF
MEFST11
Lib MEF

➤ Fichier anonyme de localisation géographique des élèves (fourni sur demande). (Enquête 16-17-18).

Champ : enseignement secondaire sous tutelle du MEN, public et privé sous contrat, pour tous les établissements du champ SCOLARITE

Variables :

Identifiant anonymisé
Année de rentrée scolaire
N° d'établissement
Division
MEF
MEFST11
Lib MEF
Code Insee commune de résidence
Libellé de la commune de résidence

Echéance : Novembre

D'autres envois en cours d'année peuvent s'avérer nécessaires.

➤ Autres fichiers

- Les fichiers THEME et données sur les ERPD (hors enquêtes individuelles)
- Le nombre d'élèves inscrits dans des formations relevant de la MLDS, la formation suivie et l'établissement d'inscription
- Les nombre des élèves inscrits en maternelle et élémentaire par établissement (public et privé) par niveau ou âge

2.3 - Fichier de données locales sur les établissements publics du 2nd degré (F_ETAB_EPI_aaaa) (EPI)

Echéance : Février

2.4 - Fichier de comptage du nombre de boursiers (F_BOURSIERS_aaaa)

Champ : enseignement public de la Région sous tutelle du MEN,

Deux fichiers :

Pré-bac : répartition du nombre de boursiers, par établissement et nombre de part

Post-bac : répartition du nombre de boursiers, par établissement, par type de formation (CPGE/BTS) et échelon

Echéance : Mai

2.5 - Fichier des comptes financiers (F_CPT_FIINANC_aaaa)

Champ : enseignement public de la Région sous tutelle du MEN,

Echéance : mai

2.6 - MEF académiques

Ensemble des MEF académiques : MEF / Libellé / date ouverture / date fermeture / MEFSTAT11 correspondant.

2.7 - Réussite aux examens

Les données figurent dans la BCP.

2.8 - Données de cadrage, prévisions d'effectifs

Les constats d'effectifs années n et n-1, figurent dans la BCP.

Les projections annuelles ou pluriannuelles, départementales et académiques pour le public et le privé sous-contrat, réalisées par les services académiques, (Enquête 41).

Mise à disposition des matrices de passages départementales et académiques du second degré constatée (publics et public plus privé sous contrat) servant à ces projections.

Pour les Régions comportant plusieurs académies, la DEPP réalise les matrices Régionales consolidées correspondantes mais il n'existe pas d'outil spécifique produisant des prévisions Régionales.

2.9 - MIGRACA

Les flux annuels d'élèves, sous statut scolaire, entre les académies de la Région concernée et les autres académies au départ de la Région, à savoir les élèves scolarisés en N-1 dans la Région concernée et dans une autre académie l'année N.

Variables:

- UAI Etablissement Année N
- UAI Etablissement Année N-1
- MEF Année N
- MEF Année N-1
- MEFST Année N
- MEFST Année N-1
- Effectifs concernés.

2.10 - Capacités d'accueil

Les Régions ont besoin de disposer, éventuellement 2 ou 3 fois par an, dans le cadre de la mise en œuvre du PPI et de la préparation des rentrées, des capacités d'accueil par établissement et formation (divisions et places par divisions) fixées par les rectorats pour chaque rentrée, données dynamiques, disponibles dans les applications académiques, pour l'année en cours et l'année à venir.

2.11 - Notes au DNB des entrants en lycées

Les Régions souhaitent estimer le niveau scolaire des entrants en lycées. L'indicateur retenu sera les moyennes des notes et/ou la part des mentions. A définir.

(Les résultats au DNB, collège par collège, sont téléchargeables depuis 2013 sur ETALAB (data.gouv.fr). A terme ils seront intégrés dans APAE).

3 - Fichiers académiques susceptibles d'être mis à disposition par les Rectorats :

3.1 - Orientation

1 - Champ : académique.

Elèves des établissements publics et privés du second degré sous statut scolaire dans les formations concernées par Affelnet.

2- Structure du fichier :

Identifiant crypté
Année scolaire
Numéro UAI de l'établissement fréquenté
Code INSEE de la commune (ou nom de la commune)
Académie
MEF de la formation correspondant à un « palier d'orientation »
MEFSTAT (11 positions) et MEF Académique
Libellé MEF (spécialité de formation si possible)

Caractéristiques de la demande d'orientation : totaux par établissement, distinction filles/garçons

Demande des familles au 2^{ème} trimestre (total)

- Nombre de demande de doublement
- Nombre de demande de passage dans une classe supérieure
- MEF pour la formation souhaitée (MEFSTAT11ou MEF Académique)
- Libellé MEF (spécialité de formation) pour la formation souhaitée
- Nombre d'autres vœux dont apprentissage

Décisions du chef d'établissement au 3^{ème} trimestre (total)

- Nombre de décisions de doublement
- Nombre de décisions de passage dans une classe supérieure
- MEF pour la formation proposée (MEFSTAT11ou MEF Académique)
- Libellé MEF (spécialité de formation si possible) pour la formation proposée

Appel de la décision du chef d'établissement

- Nombre d'appels formulés
- Nombre d'appels satisfaits

NB : Les statistiques de l'appel sont du ressort de chaque département et nécessiteront d'être centralisées au rectorat

3.2 - Demande sociale (affectation et préinscription)

Les données sur l'affectation des élèves sont disponibles, pour la plupart, à partir des applications AFFELNET et APB mais peuvent nécessiter des compléments.

1 - Champ : Académie

Elèves des établissements publics et privés du second degré sortant d'une classe dite « palier d'orientation » ou souhaitant une réorientation et demandant une affectation en formation professionnelle, générale ou technologique en lycée public (ou une préinscription dans l'enseignement supérieur).

2 - Structure du fichier :

Identifiant crypté

Effectifs détaillés par Sexe

- nombre de candidats ayant exprimé un vœu de rang 1
- nombre de candidats ayant exprimé un vœu de rang 2
- nombre de candidats ayant exprimé un vœu de rang 3
- nombre de candidats ayant exprimé un vœu de rang 4
- nombre de candidats affectés sur vœu de rang 1
- nombre de candidats affectés sur vœu de rang 2
- nombre de candidats affectés sur vœu de rang 3
- nombre de candidats affectés sur vœu de rang 4
- nombre de candidats affectés inscrits

- *Caractéristiques de la formation demandée :*

- Année de rentrée scolaire
- Académie
- Code INSEE de la commune (ou nom de la commune)
- Numéro UAI
- Nom de l'établissement d'accueil
- Statut de l'élève
- MEF (MEFSTAT11 et MEF Académique)
- Libellé MEF (spécialité de formation)
- Capacités d'accueil

- *Caractéristiques de la formation suivie l'année précédente (origine scolaire) :*

- Année scolaire
- Académie
- Code INSEE de la commune (ou nom de la commune)
- Numéro UAI
- Nom de l'établissement d'origine
- Statut de l'élève
- MEF (MEFSTAT11 et MEF Académique)
- Libellé MEF (spécialité de formation)

4 - Fichiers mis à disposition par la Région

4.1 - Fichier des ressources humaines et données Immobilières (F_ETAB_RH_aaaa.csv)

Les Régions renseignent l'enquête EPI, sous la forme prévue par les textes ; sous réserve, en cas d'évolutions, d'un délai d'adaptation pour les Régions.

Echéance : 1er trimestre année civile

4.2 - Fichier des capacités d'accueil dans les formations sanitaires et sociales

Champ : Formations hors MEN établissements publics et privés.

La Région s'engage à transmettre aux Académies les tableaux des capacités d'accueil potentielles des formations sanitaires et sociales, par formation et par établissement.

Pour les formations sanitaires, les capacités indiquées sont celles autorisées par la Région.

Pour les formations du travail social, les capacités correspondent aux effectifs constatés pour les formations agréées par la Région.

4.3- Fichier de l'enquête détaillée sur les apprentis (pour les Régions en charge de son élaboration)

Les Régions transmettent l'enquête SIFA et le constat rapide INCA, sous la forme prévue par les textes ; sous réserve, en cas d'évolutions, d'un délai d'adaptation pour les Régions.

Echéance : Mai

4.4 - Parc de logements de fonction (F_ParcLF_aaaa)

Champ : enseignement secondaire public sous tutelle du MEN

Variables : à définir au niveau local

4.5 - Etat d'occupation des logements de fonction des EPLE relevant de la Région (F_OccupLF_aaaa)

Champ : enseignement secondaire public sous tutelle du MEN, Personnels relevant de l'Etat

Variables : à définir au niveau local

4.6 - Autres champs pouvant faire l'objet de transmission de données Régionales (non exhaustif, à définir localement, par EPLE lorsque cela est possible) :

- Subventions exceptionnelles de fonctionnement aux EPLE
- Equipements (renouvelés, nouveaux, à prévoir) : nature et chiffre dont TICE
- Aides à la scolarité attribuées aux lycéens

Questionnaire passé en académie

Toutes les questions sont à réponse obligatoire sauf Q32

Q1 Académie répondante :

[Choix multiple, réponse unique, liste déroulante]

Aix-Marseille
Amiens
Besançon
Bordeaux
Clermont-Ferrand
Corse
Créteil
Dijon
Grenoble
Guadeloupe
Guyane
La Réunion
Lille
Limoges
Lyon
Martinique
Mayotte
Montpellier
Nancy-Metz
Nantes
Nice
Normandie
Orléans-Tours
Paris
Poitiers
Reims
Rennes
Strasbourg
Toulouse
Versailles

Q2 Appartenance ou fonction du porteur des réponses académiques :

[Choix multiple, réponse unique]

SSA

DSI

DPD

Autre

[SdP]

Q3 Dans les trois dernières années, avez-vous été sollicités par la collectivité territoriale régionale pour formaliser des échanges de données afin de mettre en œuvre les politiques éducatives académiques ?

[Choix multiple, réponse unique]

OUI, car cela est mis en place mais pourrait être amélioré

OUI, car cela est devenu essentiel pour une efficacité territoriale autour des politiques publiques

NON, car le sujet ne paraît pas opportun dans l'académie

NON, car les expériences n'ont pas été concluantes antérieurement

[SdP]

Q4 Dans les trois dernières années, avez-vous été sollicités par les équipes des DASEN pour formaliser des échanges de données avec les Collectivités départementales et/ou locales afin de pour mettre en œuvre les politiques éducatives départementales ou locales ?

[Choix multiple, réponse unique]

OUI, car cela est mis en place mais pourrait être amélioré

OUI, car cela est devenu essentiel pour une efficacité territoriale autour des politiques publiques

NON, car le sujet ne paraît pas opportun dans l'académie

NON, car les expériences n'ont pas été concluantes antérieurement

[SdP]

Q5 Avez-vous connaissance de la convention cadre de 2015 construite au national par la DEPP avec l'association « Régions de France » (ARF) pour organiser les échanges de données ?

OUI

NON

Q6 Cette convention cadre a-t-elle fait l'objet d'une déclinaison dans votre académie (convention d'application) ?

[Affichage si Q4 = OUI]

OUI

NON

Q7 Existe-t-il dans votre académie d'autres conventions d'échange de données infra-académiques (au niveau local ou département)

OUI

NON

[SdP]

Q8 Les données sollicitées par les collectivités concernent quels niveaux prioritairement ?

[Consigne : Hiérarchiser du plus sollicitant au moins sollicitant en glissant chaque réponse dans la partie droite de l'écran]

[Choix multiple, réponse multiple ordonné, 3 choix exactement]

Premier degré

Second degré

Niveau post bac

[SdP]

Q9 Les données sollicitées par les collectivités sont-elles en lien avec les temps suivants ?

[Consigne : Hiérarchiser du plus sollicitant au moins sollicitant en glissant chaque réponse dans la partie droite de l'écran]

[Choix multiple, réponse multiple ordonné, 3 choix exactement]

Temps scolaire

Temps périscolaire

Temps extrascolaire

[SdP]

Q10 Quelle place accordez-vous à l'échange de données dans le cadre de vos suivis de l'académie ?

[Choix multiple, réponse unique]

Jamais abordé

Abordé une fois l'an pour chaque sujet spécifique

Abordé au moins une fois par trimestre voire plus

Autres

[SdP]

Q11 Quelle(s) place(s) jouent les services académiques dans le travail autour des échanges de données ?

[Choix multiple, réponse multiple]

[Consigne : Choix multiple possible]

En ont une délégation du recteur pour répondre à la demande au cas par cas

Sollicités à la demande du recteur

Sollicités à la demande du DASEN

Autres

[SdP]

Q12 Existe-t-il dans votre académie des instances régulières qui permettent l'échange de données avec les collectivités territoriales ?

OUI

NON

[SdP]

Q13 Quelle est la nature des données ou outils échangés ?

[Choix multiple, réponse multiple]

[Consigne : Choix multiple possible]

Données de gestion

Données statistiques

Des codes sources

Des applications

Autres

[SdP]

Q14 Sur quels sujets des échanges de données sont-ils pratiqués ?

Données en lien avec les élèves :

[Choix multiple, réponse multiple]

[Consigne : Choix multiple possible]

Recensement des enfants soumis à l'obligation scolaire

Restauration scolaire et extrascolaire

Santé

Politique inclusive

Données quantitatives sur la scolarisation à chaque niveau, de l'école primaire à l'université

Données démographiques élèves (constats)

Données démographiques élèves (prévisions)

Décrochage scolaire

Les activités périscolaires et extrascolaires

L'accueil de la petite enfance au sein d'établissements et services d'accueil des enfants de moins de 6 ans.

Boursiers

Séjours liés à des activités sportives

Autres

[SdP]

Q15 Sur quels sujets des échanges de données sont-ils pratiqués ?

Données en lien avec les structures

[Choix multiple, réponse multiple]

[Consigne : Choix multiple possible]

Carte scolaire

Sectorisation et mixité sociale

Transports scolaires
Accueils collectifs de mineurs avec ou sans hébergement
Internats
Installations sportives
Autres

[SdP]

Q16 Sur quels sujets des échanges de données sont-ils pratiqués ?

Données en lien avec les ressources pédagogiques
[Choix multiple, réponse multiple]
[Consigne : Choix multiple possible]
ENTs mobilisés pour des utilisations collectives de ressources pédagogiques (au-delà des usages de restitution des notes des élèves)
Autres outils numériques (par exemple, des classes virtuelles)
Orientation (Données relatives aux formations et aux métiers)
la participation à l'organisation matérielle et financière de sorties scolaires, séjours scolaires courts et classes de découverte dans le premier degré
Autres

[SdP]

Q17 Sur quels sujets des échanges de données sont-ils pratiqués ?

Données en lien avec les ressources RH
[Choix multiple, réponse multiple]
[Consigne : Choix multiple possible]
Personnels de maintenance
Adjointes gestionnaires
Accompagnants des enfants en situation de handicap (MDPH)
Autres

[SdP]

Q18 Sur quels sujets des échanges de données sont-ils pratiqués ?

Données en lien avec la formation
[Choix multiple, réponse multiple]
[Consigne : Choix multiple possible]
Formation professionnelle
Parcoursup
Développement économique et métiers
Autres

[SdP]

Q19 Quels sont les interlocuteurs décisionnels de l'échange de données en dehors des SSA ?

[Choix multiple, réponse multiple]
[Consigne : Choix multiple possible]
Recteur
DASEN
Services académiques
Services départementaux

Autres

[SdP]

Q20 Quels sont les outils utilisés pour transférer des données ?

[Choix multiple, réponse multiple]

[Consigne : Choix multiple possible]

Open data (et mise en visibilité publique de la donnée transférée)

Transmissions de fichiers par Internet

Données partagées sur une plate-forme commune locale

Autres

[SdP]

Q21 [Tableau]

Dans le cadre d'un processus mis en place pour de l'échange de données, quelles sont les étapes qui vous semblent : réalisées, à construire, à proscrire ?	Réalisée / À construire / À proscrire
Ouvrir l'accès à ses données	
Se mettre en conformité avec le RGPD	
Construire une « fonction <i>data</i> » au sein de son organisation	
Se doter d'une doctrine en matière de diffusion de données	
Faire des choix techniques	
Commencer par un prototype ou une expérimentation puis passer à l'échelle	
Coordonner la gestion des données du territoire	
Formaliser une stratégie globale qui identifie les données, les personnes concernées et l'échéancier	

[SdP]

Q22 Conclusion : propos libre (et optionnel) sur l'échange de données en territoire

[Texte]

**Convention relative à la mise à disposition de l'annuaire fédérateur
dans le cadre du déploiement de l'ENT académique et régional normand**

entre

l'autorité académique

la collectivité représentée par son Président

et

l'EPLE représenté par son proviseur

Après avoir rappelé :

- la loi n°78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés,
- la réglementation en vigueur applicable au traitement de données à caractère personnel et, en particulier, le règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 applicable à compter du 25 mai 2018 (ci-après, « le règlement européen sur la protection des données » – dit RGPD)
- la convention cadre sur le déploiement d'un ENT normand conclue entre le Rectorat de l'Académie de Normandie, la DRAAF, la DIRM, la Région Normandie, les Départements du Calvados, de la Manche et de l'Orne et Manche Numérique,
- la délibération n°... dudu Conseil d'Administration de l'EPLE... autorisant le chef d'établissement à signer la présente convention,

Il est convenu ce qui suit

Article 1 objet de la convention

Dans le cadre de la généralisation de l'ENT dans les EPLE normands, la collectivité a fait le choix d'un annuaire fédérateur mis au point par les services académiques du ministère en charge de l'Education nationale. Alimenté par des bases académiques de gestion des élèves et parents ainsi que des personnels, cet annuaire contient des informations numériques à caractère personnel définies par l'arrêté du 30 novembre susvisé, concernant les usagers nécessaires au fonctionnement de l'ENT.

La présente convention a pour objet d'en définir les modalités de mise à disposition de l'annuaire fédérateur au sein de l'EPLE..... dans le cadre du déploiement de l'ENT ainsi que les responsabilités des différentes parties à la convention.

Article 2 obligations du chef d'établissement

Intégrer obligations RGPD

Au sens de la loi du 6 janvier 1978, le responsable du traitement des données à caractère personnel est le chef d'établissement. A ce titre, il s'engage à :

- Respecter la réglementation en vigueur applicable au traitement de données à caractère personnel et, en particulier, le règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 applicable à compter du 25 mai 2018 (ci-après, « le règlement européen sur la protection des données » – dit RGPD)

- intégrer dans la charte informatique de l'établissement, les mentions prévues par l'article 32 de la loi « informatique et libertés », afin d'informer les représentants légaux des élèves mineurs et tous les autres utilisateurs de données, de l'identité du responsable du traitement, de la finalité poursuivie par le traitement, des destinataires des données, de la possibilité d'exercer son droit d'opposition, d'accès et de rectification et auprès de qui les faire valoir,
- assurer les mesures de sécurité nécessaires à la protection des données à caractère personnel et sensibiliser les utilisateurs de l'ENT aux mesures élémentaires de sécurité,
- mettre à jour, chaque début d'année scolaire, les données à caractère personnel traitées dans le cadre d'un compte ENT et à supprimer ces données dans un délai de trois mois dès que la personne concernée n'a plus vocation à détenir un compte.

Article 3 mise à disposition de l'annuaire fédérateur par l'Académie de Normandie

L'académie de Normandie s'engage à :

- mettre à disposition du chef d'établissement l'annuaire fédérateur contenant les données à caractère personnel des utilisateurs désignés par ce dernier dès lors que ce dernier aura satisfait aux obligations décrites au précédent article,
- mettre à jour l'annuaire chaque début d'année scolaire ainsi que, lors de suppression ou modification de comptes utilisateurs qui lui seront notifiées,
- effectuer de manière sécurisée tout transfert de données personnelles relatives à l'annuaire fédérateur.

Les services régionaux de l'état se concerteront pour la mise à disposition des données des utilisateurs de l'ensemble des établissements scolaires concernés.

Article 4 responsabilités respectives des signataires

Le chef d'établissement est responsable de la protection des données à caractère personnel traitées dans l'ENT et contenues dans l'annuaire fédérateur.

Le Rectorat est en charge de la vérification du respect des exigences de sécurité destinées à protéger les données à caractère personnel.

Dans le cadre du marché passé pour le déploiement de l'ENT, les collectivités vérifieront l'existence et la portée des clauses de confidentialité relatives à la protection des données à caractère personnel ainsi que la non utilisation par le prestataire de ces données à une fin autre que le déploiement et la maintenance de l'ENT. Ces exigences sont formalisées au travers d'un engagement de confidentialité signé par le prestataire.

Elles fourniront une copie de ces clauses à l'académie et au chef d'établissement qui pourront les compléter ou modifier le cas échéant.

Article 5 durée de la convention

Elle est conclue pour une durée de un an renouvelable par tacite reconduction. Elle peut être dénoncée par une des parties avec un préavis de 6 mois.

Fait à le

Le chef d'établissement
collectivité

la Rectrice de l'Académie

le Président de la

Liste des personnes rencontrées

Administrations centrales (MENJ et MESR)

- DGESCO
 - Édouard Geffray, directeur général de l'enseignement scolaire
 - Christophe Gehin, chef de service du budget et des politiques éducatives territoriales
 - Bruno Chocchia, sous-directeur de la performance et des politiques éducatives territoriales
- Secrétariat général du MENJ
 - Céline Kerenflec'h, adjointe à la secrétaire générale
 - Nicolas Babut, conseiller stratégie numérique auprès de la secrétaire générale
- DAJ
 - Guillaume Odinet, directeur
 - Catherine Joly, cheffe de service, adjointe au directeur
 - Victor Lespinard, sous-directeur des affaires juridiques de l'enseignement scolaire, de la jeunesse et des sports
- DEPP
 - Fabienne Rosenwald, directrice
 - Jean-Michel Quenet, chef de service, adjoint à la directrice
 - Pascal André, responsable de l'animation des services statistiques académiques (SSA)
- DNE
 - Audran le Baron, directeur
 - Agnès Duwer, sous-directrice des services numériques
 - Marie Deroide, cheffe du bureau des services et outils numériques pour l'éducation
 - Philippe Ajuelos, administrateur ministériel de la donnée, des algorithmes et des codes sources (AMDAC)
- DJEPVA
 - Pierre Oudot, sous-directeur
- INJEP
 - Augustin Vicard, directeur
- DGESIP - DGRI
 - Pierrette Schuhl, sous-directrice des systèmes d'information et des études statistiques (SIES)
 - Clothilde Lixi, cheffe du département des études statistiques de l'enseignement supérieur (SIES)
 - Margot Perben, cheffe du département des études statistiques de la recherche (SIES)
 - Sébastien Chevalier, chef de service de la coordination des stratégies de l'ESR, DGESIP - DGRI A
 - Isabelle Blanc, administratrice ministérielle de la donnée, des algorithmes et des codes sources (AMDAC)

Région Auvergne-Rhône Alpes

- Sophie Hemery, directrice des enseignements et des lycées, conseil régional
- Stéphane Cabanis, direction du numérique éducatif, conseil régional
- Patrick Darnaud, direction sports, jeunesse et seniors, conseil régional
- Marlene Lesteven, cheffe du service actions éducatives en lycée, conseil régional
- Yazid Louab, direction de l'enseignement et des lycées, conseil régional

Académie de Lyon – Service statistique académique

- Pascal Arros, chef du service statistique interacadémique régional
- Christine Domenge-Vigneron, cheffe du service statistique académique de Lyon
- Paola Serries, adjointe à la cheffe du service statistique académique de Lyon
- Éric Lousteau, chef du service statistique académique de Clermont-Ferrand

Région Île-de-France

- Sarah Kowal, directrice générale adjointe
- Catherine Duboscq, adjointe DGA pôle lycées
- Marie-Cécile Demaison, directrice de l'administration et de la gestion des établissements

Académie de Paris

- Jonathan Duval, chef du service statistique académique

Collectivité de Corse

- Jean-François Cubells, directeur général adjoint des services, en charge de l'éducation, de l'enseignement, de la formation et de la langue corse
- Alexandre Alliaud, chef de la direction adjointe de l'enseignement secondaire, mission de coordination transversale des politiques d'enseignement et de formation
- Paul-Félix Benedetti (CARIF-OREF), direction éducation et recherche
- Stéphanie Ruault, direction de la formation tout au long de la vie, CARIF-OREF pour 40 %
- Éric Ferrari, direction de la transformation et de l'aménagement numérique de Corse (dont les ENTs)
- Alain Fichou, directeur des infrastructures d'enseignement
- Charles Mazzacami, direction des infrastructures d'enseignement dont le numérique éducatif dans les EPLE

Académie de Corse

- Jean-Philippe Agresti, recteur
- Blandine Brioude, SGA
- Jean-Paul Quilichini, directeur de cabinet du recteur
- Virginie Franz, IA-DASEN Corse-du-Sud
- Bruno Benazech, IA-DASEN Haute-Corse
- Jean-Christophe Olivi, chef de la division de la prospective et des statistiques académiques
- Stéphanie Marcelli adjointe à la secrétaire générale, directrice de la prospective et de l'organisation scolaire
- José Guidicelli, directeur régional académique au numérique éducatif
- Julie Albertini, cheffe du département systèmes d'information et numérique éducatif
- Jean-Guy Avelin, délégué à la protection des données, chargé de projets modernisation de l'action publique
- Pierre-Antoine Nesi, délégué académique à la formation professionnelle initiale et continue

Autres institutions

- Jules Nyssen, délégué général de l'association régions de France (ARF)
- Emmanuel Georges-Picot, conseiller en charge de l'éducation (ARF)
- Frédéric Thomas, IGÉSR
- Brigitte Dubois, directrice éducation et jeunesse, conseil départemental du Cantal
- Valérie Fillion, cheffe du service collège éducation, conseil départemental du Cantal
- Stéphanie Combes, directrice du *Health Data Hub* (HDH)

