

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

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 2nd 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 décroissant 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.

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

1. Enjeux : améliorer l'action publique grâce à la donnée

1.1. 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.education.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 Ramses. 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.

1.2. Favoriser la transparence et stimuler la démocratie

Le ministère de l'Education 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 (education.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 Education" (IDEE) du Programme d'investissement d'avenir (PIA 4, Stratégie d'accélération « Numérique et enseignement »). IDEE 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'Economie 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 IDEE 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.

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

Une plateforme des données d'éducation (Education 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 essayent 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 de services globale 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.

1.4. Bénéficier 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'Education 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 ou 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 : Etre 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.

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

2.1. 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.education.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).

A 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'Education, 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 oeuvre :

- 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 ses besoins en terme 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 (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 « Elève 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.

2.2. 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, leurs lieux d'activités, 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'activité, qui pourront alors être exploitée 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 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.

2.3. 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.

A 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 Gaïa-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 « Education 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 Education 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 université, 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 Education 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.

2.4. 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 2nd 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.education.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'Education 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 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 soutiens ;
- 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 au 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).

3. Développer le capital humain : acculturation et gestion des compétences

3.1. 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 au 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.

3.2. 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'action 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 (pilote par la donnée, data analyst, dataviz par exemple) dans le Plan national de formation de l'Education 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'Education nationale comptait 27 000 inscrits en juin 2021 (ouverture en avril 2020) et près de 2000 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).

3.3. 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 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.

3.4. 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, fonctionnelle), bénéficieront des benchmark, kits, conseils, de leurs communautés. 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.).

4. Mobiliser et évaluer : gouvernance et pilotage

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

Placé auprès du Directeur du numérique pour l'Éducation, l'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.

4.2. 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, DNE).

A 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.

4.3. 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.

4.4. Entretien 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'Education 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'Education, 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).

5. 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). A 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 produite 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 1er 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 (education.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	Etre 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
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 Education 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 (1er et 2nd 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'Education 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
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'Education, 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'Education, 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