

Stratégie nationale pour les données, algorithmes et codes sources : *un défi à relever collectivement*

Mobilisant tous les acteurs concernés, la feuille de route 2021-2024 des données, algorithmes et codes sources du ministère de l'Enseignement supérieur et de la Recherche a pour objectif de répondre à trois enjeux majeurs : innovation, transparence, simplification.

La structuration, la circulation et l'ouverture des données, initiées depuis plusieurs années, doivent désormais se généraliser à tous les types de données, qu'elles soient d'enseignement, de recherche, de gestion ou de pilotage, pour permettre un gain de valeur scientifique, économique et d'efficacité collective. Ce travail de fond mobilise tous les métiers de l'enseignement supérieur et de la recherche : administratifs, informaticiens, bibliothécaires et documentalistes, juristes, techniciens et ingénieurs en soutien à la recherche, enseignants-chercheurs et chercheurs... ainsi que les étudiants que nous devons former à ces enjeux. La feuille de route 2021-2024 des données, algorithmes et codes sources du ministère de l'Enseignement supérieur et de la Recherche (MESR) fixe ainsi un cadre de travail commun au ministère et à ses opérateurs pour répondre à trois enjeux majeurs :

- Favoriser **l'innovation** en permettant la réutilisation des données et codes sources
- Amplifier **la transparence**, pour davantage de confiance, de l'action publique
- **Simplifier** les processus de travail, notamment en allégeant la charge administrative (mise en œuvre du principe du « dites-le nous une fois »), et aider au pilotage à tous niveaux par une bonne circulation des données.

Cette feuille de route¹ comporte 53 actions qui seront, pour la quasi-totalité d'entre elles, achevées en 2024. Une nouvelle version verra le jour en 2024 et prolongera cette dynamique sur les années à venir.

LES BÉNÉFICIAIRES DE CETTE POLITIQUE

Au-delà des profils d'acteurs propres à l'enseignement supérieur et à la recherche, l'ouverture des données, des algorithmes et des codes constitue un vecteur de transparence de notre action et un levier de création de savoirs et de valeurs économique, démocratique, scientifique et politique.

En matière de transparence, on peut citer par exemple l'ouverture du code de Parcoursup. En termes de simplification, la mise en œuvre de dispositifs de circulation de données (via les API attestant du statut étudiant ou de boursier) permet aux étudiants de ne pas devoir justifier de leur statut à chacune de

leur démarche administrative, et à l'administration de bénéficier d'une information certifiée à la source, évitant de nouvelles vérifications. Elle facilite la vie des étudiants dans le cadre du recours à de nombreux services au sein et en dehors de l'université (transports, culture, sport...).

Pour la recherche, RechercheDataGouv constitue aujourd'hui une offre souveraine de dépôt et de signalement des données de recherche, associée à une offre d'accompagnement des chercheurs producteurs de données. Le MESR soutient la conservation des codes sources au travers de l'archive universelle *Software Heritage* et coordonne les différents acteurs de manière à ce que ces deux dispositifs constituent, avec l'archive nationale HAL dédiée au dépôt et à la diffusion d'articles scientifiques, un écosystème reliant codes logiciels, données et publications de recherche.

Au travers de la promotion de l'adoption d'ORCID et des identifiants liés aux structures et aux résultats de recherche (publications, données, codes logiciels...), la feuille de route adresse aussi le sujet de l'allègement de la « charge administrative » des chercheurs. L'interconnexion entre les identifiants de personnes, structures et résultats pour la circulation de l'information déjà connue facilite la constitution de CV et de listes de résultats de recherche notamment dans le cadre de candidatures à des appels d'offres, de la participation à des missions d'expertise, de l'évaluation ou de l'accès aux services des infrastructures de recherche, du dépôt en archives ouvertes des publications, des données et codes sources... Ces PIDs (identifiants uniques et pérennes) constituent ainsi la clé de voûte d'une meilleure circulation des informations et réduisent les ressaisies pour les chercheurs.

Dans le domaine de la gestion, les unités mixtes de recherche sont, pour leur pilotage et lors de leur évaluation par l'Hcéres, confrontées à l'exercice complexe et fastidieux de devoir consolider des données issues des systèmes d'informations de leurs différentes tutelles.

Faciliter cette consolidation des données constitue d'ailleurs la recommandation prioritaire en matière de simplification identifiée par la mission Gillet²

[1] www.esr.gouv.fr/politique-donnee

[2] <https://www.enseignementsup-recherche.gouv.fr/sites/default/files/2023-06/rapport-mission-sur-l-cosyst-me-de-la-recherche-et-de-l-innovation-28193.pdf>

[3] <https://www.appelsprojetsrecherche.fr>

(proposition 7 du rapport). Dans ce même esprit, la feuille de route identifie la nécessité de disposer d'un référentiel national univoque des données décrivant les structures de recherche.

LES IDENTIFIANTS PÉRENNES INTERNATIONAUX : CLÉ DE VOÛTE DE LA CIRCULATION ET DU PARTAGE DES DONNÉES

Dans le domaine de la recherche par exemple, il s'agit d'identifier de manière univoque les « entités clés » d'un système de recherche et d'établir les liens entre les chercheurs, leurs productions (publications, données, logiciels...), les unités de recherche, leurs tutelles, les sources de financement et personnels de l'unité.

Cette représentation s'entend aux niveaux national et international et elle doit être pérenne et tolérante aux transformations des organisations.

C'est pourquoi le ministère porte une stratégie d'emploi d'identifiants pérennes internationaux et souhaite définir et mettre en œuvre un plan d'actions favorisant leur adoption au niveau national.

Directement ou au travers de ses opérateurs, le ministère est déjà impliqué au niveau des gouvernances des instances internationales chargées d'attribuer ces identifiants que sont Crossref (publications), DataCite (données), ORCID (chercheurs) et est en lien étroit avec ROR (structures).

LES ACTIONS IMPORTANTES À MENER ET LE RÔLE DE L'ABES

Au niveau national, l'emploi de ces identifiants pérennes et internationaux est aujourd'hui naturel pour ce qui touche aux productions de la recherche. L'effort doit être porté sur l'identification des chercheurs et des structures de recherche.

Pour l'entité « chercheur », il s'agit d'agir à différents niveaux pour en faire *in fine* un véritable support de simplification et de visibilité :

- Faire en sorte que les différents systèmes qu'utilisent les chercheurs leur permettent d'employer ORCID pour faire transiter, sous leur contrôle, les informations pertinentes de leur profil : le portail des appels à projet³ le permettra dès 2024 dans le cadre de la réponse à un appel
- Faciliter l'initialisation et l'entretien d'un profil ORCID, au travers de services capables de lier les productions à leurs auteurs, les chercheurs à leurs unités, les unités aux établissements impliqués
- Améliorer le dispositif ORCID lui-même, au niveau international pour qu'il accroisse encore sa couverture au niveau mondial et sa capacité de mise en qualité des informations.

Pour les structures de recherche, un référentiel français univoque aligné avec ROR émergera de façon à consolider l'information de référence et à la propager.



Barques de pêche amarrées au bord de la rivière

Crédit Adobe stock, par Abdul Momin

Un cadre de référence des structures définissant les types de structures (unités, fédérations, tutelles...) et la gouvernance des données associées sera défini en 2024.

Au vu de son expertise sur les identifiants, dont IdRef, et du rôle de coordonnateur du consortium ORCID France qu'elle joue déjà, l'Abes s'est vu confier récemment par le MESR un rôle opérationnel de garant de la qualité et de l'unicité des données de structures dans le respect du cadre de référence qui sera défini.

HUGUES PONCHAUT

Chargé de mission politique des données, des algorithmes et des codes sources au ministère de l'Enseignement supérieur et de la Recherche
hugues.ponchaut@recherche.gouv.fr


**MINISTÈRE
DE L'ENSEIGNEMENT
SUPÉRIEUR
ET DE LA RECHERCHE**
*Liberté
Égalité
Fraternité*

• • • LOI 3 DS : DU « DITES-LE NOUS UNE FOIS » À L'ADMINISTRATION PROACTIVE

La loi du 21 février 2022 dite loi 3 DS (différenciation, décentralisation, déconcentration et simplification) porte des dispositions visant à simplifier le fonctionnement des institutions locales et son article 162 vient renforcer le principe du « dites-le-nous une fois », tel qu'initialement prévu par l'article L.114-8 du code des relations du public avec l'administration¹ : le partage d'informations devient ici obligatoire. La loi autorise également les échanges qui permettent d'informer proactivement les usagers de leurs droits.

LES DÉCRETS :

- Le décret n° 2023-361 du 11 mai 2023² fixe les conditions d'application du nouveau principe d'échange d'informations entre administrations.
- Le décret n° 2023-362 du 11 mai 2023³ détermine, pour chaque type d'informations ou de données, la liste des administrations chargées de les mettre à la disposition d'autres administrations.

[1] https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000045213315
[2] <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000047541339>
[3] <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000047541361>