

Mettre en œuvre une gestion des données de la recherche dans une institution en 6 points

1. **S'appuyer sur le contexte international et national**
2. **Associer les scientifiques et les services d'appui**
3. **Tirer parti des expériences des autres**
4. **Intégrer un large panel de contributeur.rice.s**
5. **Communiquer en interne et à l'extérieur**
6. **Former les personnels et décrire les emplois nouveaux**

1. S'appuyer sur le contexte international et national

L'intérêt public porté aux données de la recherche date de 2013, à la faveur d'évènements internationaux, notamment :

- le lancement de la [Research Data Alliance](#) (RDA) par la Commission européenne, la National Science Foundation (NSF, USA), le National Institute of Standards and Technology (NIST, USA) et le gouvernement australien. En juillet 2019, la RDA comprenait 8 600 membres dans 137 pays et une vingtaine de groupes nationaux dont le groupe France créé en 2018 ([RDA France](#)) ;
- la coalition internationale SPARC (Scholarly Publishing and Academic Resources Coalition), issue de l'association des bibliothèques de recherche (Association of Research Libraries, ARL), publiait en mars 2013 le guide [Implementing an Open Data Policy: A Primer for Research Funders](#) pour aider les agences de financement de la recherche à mettre en œuvre une politique d'ouverture des données ;
- au niveau politique, les chefs d'Etat du G8 signaient en juin 2013 [une charte sur l'ouverture des données publiques](#).

En France, entre 2012 et 2013, le ministère de l'enseignement supérieur et de la recherche lançait le chantier des données de la recherche dans le cadre de la Bibliothèque scientifique numérique (BNS) et [le site d'informations sur les données de la recherche](#). En 2019, ceux-ci ont été prolongés par le Comité national pour la science ouverte (CoSO) et par le site web [Ouvrir la Science](#).

Le contexte international et les politiques nationales sont un levier pour les institutions, à la fois émulateur et contributeur de dynamiques locales. Les initiatives mondiales permettent aux institutions de recherche de se saisir des concepts de données de recherche, données publiques, politiques de données, ouverture des données, et de lancer leurs propres réflexions et projets.

En inscrivant leurs démarches dans ces mouvements, les institutions s'assurent que leurs choix ne contreviennent pas aux réglementations et qu'ils répondent à des problématiques partagées ne relevant pas de leur seule responsabilité, telles que la propriété intellectuelle autour de la coproduction ou de la réutilisation de données, le développement de compétences et l'évolution des métiers, l'offre de services et de produits pour la gestion et la valorisation des données, etc.

2. Associer les scientifiques et les services d'appui

Le sujet des données de la recherche est pluridisciplinaire, plus complexe que celui des publications. Il recouvre l'ensemble des étapes et des activités associées de gestion, de conservation, de diffusion et de réutilisation des données scientifiques — voir le poster associé à cette fiche [Le cycle de vie des données de la recherche](#) (Cirad, 2016). Il fait appel à la fois à des **savoirs scientifiques** et à des **expertises professionnelles dans des secteurs d'appui à la recherche** comme l'information scientifique, les systèmes d'information, la propriété intellectuelle, la valorisation, la formation, la communication.

Pour mener un chantier institutionnel sur les données de recherche, le copilotage entre une composante scientifique et une composante d'appui est essentiel. Faire travailler ensemble de façon constructive et efficiente les communautés scientifiques et les services d'appui à la recherche est un défi mais c'est aussi un facteur de réussite : ce copilotage permet d'ancrer le chantier dans l'institution et d'optimiser ses ressources — humaines, financières, technologiques.

3. Tirer parti des expériences des autres

Dans une institution de recherche, le chantier de gestion des données de la recherche nécessite deux formes de préparation :

- définir en amont les objectifs génériques, les objectifs spécifiques, les livrables attendus, les actions à mener et le calendrier associé ;
- identifier et distinguer les démarches sur le temps long et les initiatives sur le temps court.

Le temps long peut être celui de la **constitution d'une communauté** disciplinaire ou sectorielle vivante sur le sujet des données de la recherche. Par exemple, créer une activité de veille sur ce sujet peut être traité sur un temps long avec des objectifs génériques et des contraintes minimales : identifier les animateurs ainsi que les personnes qui contribuent à la veille, réfléchir aux modalités de la veille — rôles et activités de chacun.e, choix des outils et du produit en sortie.

Le temps court est celui des **projets** de durée limitée et dont le résultat et l'échéance sont précisément définis. Par exemple, créer un inventaire des données ou un entrepôt institutionnel de données de recherche fait appel à un fonctionnement en mode projet avec un arbitrage institutionnel via des instances de décision, un portage, un livrable et des moyens sur une échéance à moyen terme.

Quels que soient l'ampleur et le temps des démarches et des initiatives, il est essentiel de les formaliser et de les jalonner par des bilans associés à une évolution des orientations si nécessaire. La formalisation et les bilans sont indispensables pour éviter toute confusion, frustration et démobilitation des personnes sollicitées au départ.

Avant de lancer le chantier, il est utile de voir ce qui se passe ailleurs :

- bénéficier du retour d'expérience des établissements homologues dans son pays, quel que soit l'avancement de leur projet. Il est souvent plus avantageux de s'appuyer sur des méthodes et des organisations du travail qui ont fait leur preuve, afin d'éviter les erreurs vécues par d'autres ;
- s'inspirer d'un fonctionnement éprouvé. La Research Data Alliance (RDA, <https://www.rd-alliance.org/>) a formalisé des principes d'organisation, de fonctionnement et de communication, qui sont une base solide pour organiser le chantier institutionnel.

Pour aborder les sujets de gestion des données de recherche, des groupes de personnes intéressées seront constitués dans l'institution, sous la forme de groupes de travail ou d'intérêt selon le sujet :

- un **groupe de travail** a pour objectif un livrable à produire en quelques mois à un an (temps court), par exemple des préconisations pour identifier de façon pérenne un jeu de données, pour citer un jeu de données, l'élaboration d'un modèle de plan de gestion de données (PGD), la définition de métriques sur l'usage des données, l'élaboration d'un programme de formation.
- un **groupe d'intérêt** est plus pérenne (temps long). Il peut mener une réflexion autour d'une discipline (sciences humaines et sociales, données génomiques, biodiversité, etc.) ou d'un secteur d'activités (information scientifique et technique, juridique, ressources humaines, archivage, etc.). Il peut aussi rédiger des guides de bonnes pratiques. Selon les conclusions auxquelles il aboutit, il peut soumettre des propositions à un groupe de travail ou lancer un nouveau groupe de travail.

4. Intégrer un large panel de contributeur.rice.s

Dans une institution, la gestion des données de la recherche concerne l'ensemble des secteurs, que ce soit les activités scientifiques et les activités d'appui — information scientifique et technique, juridique, valorisation, ressources humaines, etc.

Cela implique de mener le chantier de gestion des données de la façon inclusive et participative, c'est-à-dire d'intégrer un large panel de contributeurs et de contributions. L'émulation collective, l'appel à la créativité et à l'esprit d'initiative, la collaboration et le partage sont les clés du succès dès lors que les rôles et responsabilités sont clairement identifiés.

Sur un sujet sectoriel très spécialisé (juridique, information scientifique et technique, système d'information), on peut toutefois imaginer que le service d'appui concerné s'organise en interne pour réaliser une étude ou mener une action. Par exemple :

- le service juridique peut rédiger un guide sur les lois et règlements en vigueur dans le domaine des données ;
- le service d'information scientifique et technique peut adhérer pour le compte de l'institution à des initiatives nationales ou internationales, comme le consortium [DataCite](#) pour attribuer des DOI aux jeux de données produits par l'institution (DOI, identifiant numérique d'objet, *Digital Object Identifier* – Voir la fiche CoopIST : [Identifier et rechercher une publication ou un jeu de données par son DOI](#))
- , ou s'abonner à des sources d'information dédiées comme la base de données [Data Citation Index](#) (base de jeux de données déposés dans des entrepôts publics dans le monde) ;
- le service informatique peut mener une étude d'outils libres de gestion de données de la recherche (logiciels d'entrepôts par exemple), et installer certaines versions pour les tester.

5. Communiquer en interne et à l'extérieur

Les données de la recherche sont un sujet récent et stratégique. L'enjeu est de sensibiliser à la fois les scientifiques, les personnels d'appui, les responsables (directeur.rice.s de laboratoires de recherche) et la direction de l'établissement (direction générale, direction scientifique, conseil scientifique, etc.). Un groupe de travail peut être constitué avec les objectifs suivants :

- l'organisation d'évènements, par exemple des **journées thématiques** ou des **ateliers de travail**, pour faire un état des réflexions en cours, présenter les initiatives lancées, montrer les premiers résultats obtenus, faciliter les discussions et les interactions au sein de l'institution ;

- la **communication interne et externe** sur les données, en veillant à ce que les messages diffusés soient faciles à comprendre et à retenir, de manière à être repris par tous dans l'institution et à l'extérieur ;
- l'animation d'une **liste de diffusion** interne, la création d'un **site intranet**, voire d'un **site internet**, dédiés aux données de la recherche. Les rôles et modalités de modération de la liste de diffusion et de coordination éditoriale du site doivent être précisés ;
- la rédaction de documents institutionnels sur les données de la recherche, comme une **charte de bonnes pratiques** pour faciliter la sensibilisation de tous au sein de l'établissement.

6. Former les personnels et décrire les emplois nouveaux

Pour se former à la gestion des données de la recherche, les possibilités offertes par des dispositifs ouverts de formation continue, comme le dispositif **DORANum** en France, sont à étudier.

Des démarches individuelles de formation peuvent être couplées à des démarches collectives, voire mutualisées entre différents établissements, afin de s'adapter aux besoins des personnes et de favoriser les interactions entre apprenants.

Une attention particulière doit être portée aux points suivants :

- dans un premier temps, il n'est pas raisonnable ni efficient de demander aux personnes formées de former à leur tour leurs collègues — cela impliquerait qu'elles aient les capacités et la disponibilité pour remplir le rôle et la fonction de formateur. Il est préférable de confier le programme de formation à des organismes de formation certifiés. L'animation interne de sessions de formation, sous la forme d'ateliers participatifs par exemple, peut être réservée à des problématiques spécifiques à l'institution, comme utiliser un entrepôt institutionnel de données de la recherche, ou rédiger un plan de gestion de données (PGD) d'un projet de recherche ;
- l'enchaînement et la cohérence des ateliers doivent être minutieusement réfléchis afin qu'un même sujet soit traité de façon intelligible et suivi d'un atelier à un autre : par exemple, le sujet de la propriété intellectuelle des données sera abordé dans l'atelier sur les PGD et repris dans l'atelier sur le dépôt et les conditions d'accès d'un jeu de données dans un entrepôt de données.

Pour une institution, l'effort de formation est aussi l'occasion d'identifier les compétences, les activités et les métiers à développer dans le domaine des données de la recherche. Les référentiels métiers existants (Voir la fiche CoopIST [Découvrir de nouveaux métiers liés aux données de la recherche](#)) sont des sources précieuses pour les actions suivantes :

- rédiger les lettres de mission pour les personnels en place ;
- élaborer des fiches de postes pour de futurs recrutements ;
- actualiser son référentiel des métiers si l'institution en a l'ambition et les moyens.

Marie-Claude Deboin

Délégation à l'information scientifique et technique, Cirad - 23 août 2019

(avec l'appui éditorial de Cécile Fovet-Rabot)

Comment citer ce document :

Deboin M. C. 2019. Mettre en place une gestion de données de la recherche dans une institution en 6 points (FRA) : CIRAD, 4 p.

<https://doi.org/10.18167/coopist/0064>

Cette œuvre est mise à disposition selon les termes de la Licence Creative Commons : Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions 4.0 International, disponible en ligne : <http://creativecommons.org/licenses/by-nc-sa/4.0/deed.fr>

ou par courrier postal à : Creative Commons, 171 Second Street, Suite 300, San Francisco, California 94105, USA.

Cette licence vous permet de remixer, arranger, et adapter cette œuvre à des fins non commerciales tant que vous créditez l'auteur en citant son nom et que les nouvelles œuvres sont diffusées selon les mêmes conditions.