

Comment monter un service d'accompagnement aux données ?

1. En jeux

La création d'un service d'accompagnement aux données peut être vue comme la réponse aux injonctions des agences de financement et leur volonté d'améliorer le retour sur investissement de la recherche publique financée. Mais les enjeux de la gestion de données de la recherche sont plus complexes et variés :

- « Science **cumulative** » : éviter de refaire ce qui a été déjà fait, gagner du temps, faciliter la coopération et l'échange d'informations entre les partenaires d'un projet ;
- « **Reproductibilité** » : permettre à une équipe différente de reproduire des résultats de recherche grâce à des données structurées ;
- Administration de la « **preuve scientifique** » : lutter contre la fraude et poser l'antériorité de sa recherche ;
- « **Découverte** » : la mise à disposition des données permet l'exploration de thèmes qui n'avaient pas été envisagés par l'équipe de chercheur-e-s à l'origine.

Pour répondre à ces enjeux, il convient de respecter à la fois l'adage « aussi ouvert que possible, aussi fermé que nécessaire » et de promouvoir l'adoption des [principes FAIR](#). Le respect de ces principes confronte les chercheur-e-s à de nouvelles questions qui bouleversent leurs pratiques. Le service d'accompagnement aux données permet de proposer un appui et des ressources pour les guider dans cette démarche.

2. Exemples de questions de chercheur-e-s

- Je vous assure que je n'ai pas été informé-e de l'obligation de rédiger un PGD... ou encore j'ai une ANR flash. Suis-je concerné-e ?
- J'ai besoin de logiciels de traitement, fouille et visualisation de données.
- Je veux que mon projet respecte les attendus de l'agence de financement dans la gestion des données.
- Je suis perdu-e dans la rédaction du PGD.
- Je dois me conformer aux principes FAIR. De quoi s'agit-il ?
- Une agence de financement ou un éditeur me demande de déposer les données d'appui de mon article dans un entrepôt. À qui m'adresser ?
- Mes données ne sont pas anonymisables. Que faire ?
- Quid des données générées par des projets lancés il y a 10 ans ?
- Je veux archiver mes données pour l'éternité.
- Je suis seul-e à pouvoir comprendre et interpréter mes jeux de données, les partager ne servirait à rien !
- Mon jeu de données est incompréhensible tel quel pour quelqu'un qui n'a pas suivi le projet. Que faire ?
- Comment choisir le standard de métadonnées approprié ? Comment s'y prendre ?
- Pourquoi une agence de financement ou un éditeur me demande d'ouvrir mes données ?
- Comment faire pour arrêter de perdre mes données (erreurs de format, logiciel propriétaire qui n'est plus maintenu, perte de clef USB...)?
- Comment faire pour rendre mes données intelligibles pour d'autres (organisation, nommage, documentation) ?

3. Les étapes de création d'un service

Commencer petit !

Une offre complète ne peut être constituée en quelques mois. Il est recommandé de mettre en place des services immédiatement nécessaires. Vous pourrez ainsi créer des liens avec des équipes de recherche et avoir des retours sur les actions mises en œuvre. Afin de mettre en place la série d'actions énoncée ci-dessous, il est recommandé mais pas indispensable de disposer d'un soutien politique.

Actions à mettre en place dans un deuxième temps

- Demander la création d'un espace institutionnel DMP OPIDoR (Inist-CNRS) : mise en place de recommandations et/ou création d'un modèle de PGD, activation de l'option Assistance conseil. Certains établissements vont plus loin en proposant des *data interviews* et de la corédaction.
- Accompagner au *text and data mining* : fournir des ressources numériques pour le traitement, la fouille et la visualisation de données et formation en collaboration avec les collègues des laboratoires.
- Relire les dossiers de financement en lien avec les services d'aide au montage de projets. Introduire la Science Ouverte et les bonnes pratiques de gestion des données dès le montage de projet. [Voir le guide Améliorer son projet ANR grâce à la Science Ouverte](#) ou le [Guide données de la Commission européenne pour les projets Horizon Europe](#).
- Mettre en place une collaboration avec la-e DPO de son établissement : accompagner les équipes de recherche aux questions éthiques et juridiques sur les données.
- Réfléchir sur les chantiers de dépôt rétrospectif de données dans un entrepôt : identification des jeux de données de longue traîne, interopérabilité, sélection...
- Conseiller sur l'archivage numérique pérenne : liens avec le Cines, la BnF (Spar).
- Proposer la relecture de *data papers* et/ou des modèles de *data papers* par discipline.

5. Interlocuteurs et interlocutrices privilégié-e-s

- Les chercheuses et les chercheurs sensibilisés à la Science Ouverte
- Les personnels des bibliothèques qui sont souvent impliqués dans les activités transverses (service à la recherche, données, Science Ouverte...)
- Les personnels IST des laboratoires de votre site / des réseaux locaux
- Les responsables et ingénieurs de plateformes de données (PUD)
- Les ingénieur-e-s dans les laboratoires
- La Direction des Services Informatiques
- Les services d'aide au montage de projets
- Les services de valorisation de la recherche
- Les services juridiques
- La-e délégué-e à la protection des données (DPD ou DPO - *Data Protection Officer*) de l'établissement
- La comitè d'éthique ; la-e référent-e intégrité scientifique
- Les services des archives
- Les réseaux professionnels externes à l'établissement (GTSO Données de Couperin, RDA France...), les listes : [datalibrarians](#), [bibrecherche](#), [liste donnees-inter-reseaux...](#)

6. Liens utiles

- Guide de l'OCLC *Building Blocks: Laying the Foundation for a Research Data Management Program*: <https://www.oclc.org/research/publications/2016/oclcresearch-data-management-building-blocks-2016.html>
- MOOC : *Delivering Research Data Management Services* <https://www.futurelearn.com/courses/delivering-research-data-management-services>
- Page du GTSO Données qui recense les enquêtes d'établissements sur les pratiques de leur communauté scientifique en matière de gestion des données : <https://scienceouverte.couperin.org/donnees-recherche-enquetes/>
- Outils de rédaction de PGD : DMP OPIDoR, DMP Tool, DMP online, Data Stewardship Wizard, Argos.
- Cat OPIDoR (Inist-CNRS), wiki des services dédiés aux données de la recherche : https://cat.opidor.fr/index.php/Cat_OPIDoR_wiki_des_services_d%C3%A9di%C3%A9s_aux_donn%C3%A9es_de_la_recherche
- Le moteur de recherche de revues « [Où publier ?](#) » du CIRAD : choisir sa thématique et filtrer par type d'article (*data papers*).