

Guide de la science ouverte de l'EDSP2

Principes et définition de la Science Ouverte

Sommaire

Principes et définition de la Science Ouverte

Pourquoi ouvrir la Science ouverte

Les avantages de la science ouverte pour un doctorant

Comment ouvrir la science ?

A) Généraliser l'accès ouvert aux publications	3
1- Déposer ses résultats dans une archive ouverte	3
2- Comment connaître la politique de la revue	5
3- Les revues prédatrices.....	5
4- Les droits en matière d'accès ouvert	5
5- L'identification ORCID	7
6- Open U Journals : une solution bordelaise pour créer une revue	8
7- Le portail HAL institutionnel de l'université de Bordeaux.....	8
B) Structurer et ouvrir les données de la recherche.....	8
1- Préparer et documenter des données.....	9
2- Comment choisir un entrepôt de données ?	9
3- Outil d'aide au choix d'un entrepôt de données.....	9
4- ABDo : l'Atelier bordelais des données de recherche	10
5- L'espace institutionnel Recherche Data Gouv de l'université de Bordeaux	10
6- Lier les données aux publications	11
7- FAIR et DMP	11
C) S'inscrire dans une dynamique durable, européenne et internationale	13
Pour aller plus loin : l'offre de formations dans ADUM et les sources	14

Principes et définition de la Science Ouverte

La science ouverte a pour projet de rendre les connaissances, les données, les méthodes et les résultats de recherche accessibles à tous, de manière transparente et collaborative.

Elle repose notamment sur plusieurs principes :

- Accès ouvert aux publications : les articles scientifiques sont disponibles gratuitement en ligne, sans abonnement ;
- Partage des données de recherche : les données produites par les chercheurs peuvent être réutilisées et vérifiées ;
- Transparence des méthodes et des codes : les protocoles, logiciels et outils utilisés sont ouverts et reproductibles ;
- Collaboration et diffusion élargie : chercheurs, citoyens, entreprises et institutions peuvent participer plus facilement à la production et à l'utilisation des connaissances.

L'objectif principal est de rendre la recherche :

- plus accessible ;
- plus reproductible ;
- plus efficace ;
- et plus utile à la société.

Pourquoi ouvrir la science ?

Trois facteurs principaux doivent inciter à une plus large pratique de la science ouverte. Ils sont d'ordre scientifique, économique et citoyen.

- Scientifique : une science qui partage les produits de la recherche scientifique permet une circulation des protocoles, des données et des résultats. L'interaction entre les chercheurs, la reproductibilité (et donc la transparence et l'intégrité scientifique) et la réinterprétation sont favorisés. Par ailleurs, l'ouverture peut aussi, grâce (entre autres) aux nouveaux outils de traitement des données, ouvrir de nouvelles pistes de recherche et favoriser l'innovation.
- Économique : Une science plus ouverte vise à proposer un modèle alternatif au modèle économique actuel de l'édition scientifique. En effet, La hausse constante des coûts n'est plus soutenable alors même que les outils numériques permettent aujourd'hui aux scientifiques de reprendre la maîtrise du circuit de l'édition.
- Citoyen : Une science plus ouverte répond à l'attente de transparence des citoyens, et permet de lutter contre les fausses informations scientifiques qui circulent aujourd'hui dans les plus hautes sphères des instances décisionnelles.

Les avantages de la science ouverte pour un doctorant

- Visibilité et impact accrus
Les publications et données en accès ouvert sont davantage citées et diffusées. Pour un doctorant qui construit sa réputation scientifique, c'est une opportunité de faire connaître ses travaux bien au-delà de sa discipline ou de son institution, dès le début de sa carrière.

- **Crédibilité et intégrité scientifique renforcées**
Partager ses données, protocoles et codes sources démontre la rigueur de sa démarche. Cela protège aussi le doctorant en cas de contestation : tout est traçable et vérifiable, ce qui renforce la confiance de la communauté envers ses résultats.
- **Accélération de la progression de la thèse**
Accéder librement aux données et publications d'autres chercheurs évite de "réinventer la roue". Le doctorant peut s'appuyer sur des ressources existantes, reproduire des analyses et gagner un temps précieux dans ses propres recherches.
- **Opportunités de collaboration et de réseautage**
Rendre ses travaux accessibles attire l'attention d'autres chercheurs, potentiellement dans des domaines complémentaires. Cela peut déboucher sur des collaborations inattendues, des co-publications ou des invitations à des conférences — autant d'atouts pour une carrière académique ou industrielle.
- **Préparation aux exigences institutionnelles et des financeurs**
De plus en plus d'agences de financement (ANR, Horizon Europe...) et d'universités imposent le dépôt en accès ouvert comme condition à l'obtention de financements. Se former à ces pratiques dès le doctorat, c'est anticiper des obligations qui seront incontournables tout au long de la carrière.

Comment ouvrir la science ?

3 axes majeurs :

- A) Généraliser l'accès ouvert aux publications
- B) Structurer et ouvrir les données de la recherche
- C) S'inscrire dans une dynamique durable, européenne et internationale

A) Généraliser l'accès ouvert aux publications

1- Déposer ses résultats dans une archive ouverte

Trois grandes voies permettent de répondre à cet objectif : le dépôt en archive ouverte (*la voie verte*), la publication en accès ouvert (*la voie dorée*), la publication dans des revues en accès ouvert (*la voie diamant*).

La voie verte

La voie verte consiste, pour les chercheurs, à déposer une version de leurs publications scientifiques dans une archive ouverte, indépendamment de leur publication dans une revue scientifique.

- HAL (Hyper Articles en Ligne) : est une archive ouverte pluridisciplinaire développée en 2001 par le Centre pour la Communication Scientifique Directe (CCSD), unité du CNRS.
- arXiv : créée en 1991 par l'Université Cornell, arXiv est considérée comme l'une des plus anciennes archives ouvertes. Elle couvre les domaines de la recherche en physique, mathématiques, informatique et biologie quantitative.

- PubMed Central : lancée en février 2000 par la National Library of Medicine des National Institutes of Health (NLM/NIH) aux États-Unis, PubMed Central permet l'accès au texte intégral d'articles dans le domaine biomédical et des sciences de la vie.

La voie dorée

La voie dorée correspond à la publication d'un article scientifique en accès ouvert dans une revue, rendant son contenu immédiatement accessible gratuitement à tous les lecteurs. Le principe général consiste à transférer les coûts de publication vers un tiers autre que le lecteur ou l'abonné. Trois principaux modèles de revues en accès ouvert, sont distingués et diffèrent surtout par leur mode de financement et l'accès aux contenus.

- Les revues en accès ouvert intégral (ou "gratuites") : Ce sont des revues à comité de lecture dont tous les articles sont accessibles gratuitement dès leur publication. Les lecteurs ne paient pas pour accéder aux contenus, et les auteurs ne paient généralement pas de frais de publication (exemples : OpenEdition, les épi-revues ou Open U Journals). Le Directory of Open Access Journals (DOAJ) recense ces revues selon des critères de qualité, de transparence éditoriale et de respect des droits d'auteur.
- Le modèle auteur-payeur (APC) : Dans ce modèle, la publication en accès ouvert est financée par des frais appelés Article Processing Charges (APC), payés par les auteurs, leur laboratoire ou leur institution. L'article est ensuite immédiatement accessible gratuitement pour tous les lecteurs.
- Le modèle hybride : Il s'agit de revues principalement financées par les abonnements, mais qui proposent aux auteurs de rendre leur article en accès ouvert moyennant le paiement d'APC. Cela signifie que dans une même revue certains articles sont accessibles uniquement par abonnement et d'autres sont en accès ouvert. Ce modèle est souvent critiqué car il peut entraîner un double paiement : les institutions paient à la fois les abonnements pour lire les revues et les APC pour publier en accès ouvert. Pour cette raison, de nombreuses politiques de science ouverte encouragent à limiter ou éviter ce modèle. La publication dans des revues hybrides (abonnement + APC) ne sont plus financées par l'ANR ni Horizon Europe.

La voie diamant

La voie diamant se distingue des voies verte et dorée par son modèle économique. Contrairement à la voie verte, la voie diamant concerne directement la publication dans des revues en accès ouvert. Elle se différencie également de la voie dorée classique car elle ne repose pas sur le paiement de frais de publication (APC) par les auteurs ou leurs institutions.

Ce modèle repose généralement sur le financement par des institutions publiques, des universités, des bibliothèques, des organismes de recherche ou des subventions publiques. Ce financement permet de soutenir directement les revues diamant, c'est-à-dire des revues scientifiques en accès ouvert intégral qui ne facturent ni frais de lecture ni frais de publication. Ces revues reposent sur le travail collaboratif des chercheurs, des bibliothécaires, des universités et des plateformes publiques d'édition scientifique, dans une logique de coopération académique plutôt que commerciale.

Les revues diamant sont diffusées et accessibles via différentes plateformes de publication, telles que OpenEdition Journals, Episciences ou certaines revues hébergées par des universités et des infrastructures de recherche.

2- Comment connaître la politique de la revue

Plusieurs outils permettent de vérifier la politique d'un éditeur en matière d'accès ouvert avant de publier un travail de recherche.

- SHERPA/RoMEO : ce service développé par l'université de Nottingham recense les politiques des éditeurs en matière d'accès ouvert et les classe selon différents niveaux d'autorisation de diffusion et d'archivage.
- Willo : service développé par l'université de Lille, permettant de connaître les droits liant un auteur à son éditeur en matière d'accès ouvert pour un article donné.
- Directory of Open Access Journals (DOAJ) : le DOAJ recense les revues scientifiques en accès ouvert selon plusieurs critères : accès gratuit au texte intégral, transparence de la politique éditoriale, archivage pérenne, utilisation d'identifiants pérennes, respect des droits d'auteur et recours à des licences Creative Commons.

Le DOAJ permet également d'accéder aux métadonnées des articles, améliorant ainsi la visibilité des contenus scientifiques. Son mode de recherche avancée permet de trouver des revues entièrement en accès ouvert, avec APC ou hybrides.

3- Les revues prédatrices

Les revues prédatrices sont des revues scientifiques qui privilégient des objectifs commerciaux au détriment des standards académiques et de l'éthique scientifique.

Bien que les éditeurs prédateurs soient antérieurs à l'apparition de l'accès ouvert, leur prolifération récente a été favorisée par l'émergence et le succès du modèle auteur-payeur. Ces éditeurs exploitent le modèle des frais de publication facturés aux auteurs afin de générer des profits.

Ces revues publient des articles sans véritable processus rigoureux d'évaluation par les pairs, malgré les garanties affichées sur leurs sites. Elles proposent généralement des délais de publication très rapides, des frais de publication peu élevés, un processus d'évaluation inexistant ou superficiel, ainsi que des comités éditoriaux peu transparents ou composés de chercheurs difficilement identifiables. Leur taux d'acceptation des articles est également très élevé.

Les revues prédatrices démarchent fréquemment les chercheurs par l'envoi massif de courriels non sollicités (spam) et ciblent particulièrement les jeunes chercheurs, notamment après la publication de leur thèse.

4- Les droits en matière d'accès ouvert

La loi pour une République numérique (2016)

L'article 30 de la loi n° 2016-1321 du 7 octobre 2016 pour une République numérique (dite « loi Lemaire »), codé à l'article L. 533-4 du Code de la recherche, crée un droit nouveau pour les auteurs : celui de diffuser librement leurs écrits scientifiques en accès ouvert, quel que soit le contrat signé avec l'éditeur. Ce droit s'applique à tout écrit publié dans une revue parue au moins une fois par an, dès lors que la recherche a été financée au moins à 50 % par des fonds publics (dotations de l'État, collectivités territoriales, établissements publics, agences de financement nationales ou fonds européens).

- Sans embargo lorsque l'article est déjà accessible (il peut y avoir des coûts d'APC) sur le site de la revue ;
- Ce droit autorise le dépôt de la version acceptée pour publication (postprint/AAM), avec embargo maximal de 6 mois (STM) ou 12 mois (SHS), sans que l'éditeur puisse s'y opposer.

Le Plan S et l'accélération de l'accès ouvert

Élaboré en septembre 2018 par cOAlition S, un consortium de 16 agences nationales de financement de la recherche soutenu par la Commission européenne, le Plan S vise à accélérer la diffusion des publications scientifiques en accès ouvert. Il affirme notamment que : À partir de 2021, toutes les publications scientifiques issues de recherches financées par des subventions publiques ou privées accordées par des organismes de financement nationaux, régionaux et internationaux devront être publiées dans des revues ou sur des plateformes en accès ouvert, ou être immédiatement disponibles sans embargo dans des archives ouvertes.

Ce plan comporte 10 principes disponibles sur <https://www.coalition-s.org/addendum-to-the-coalition-s-guidance-on-the-implementation-of-plan-s/principles-and-implementation/>

Les obligations des financeurs : ANR et Horizon Europe

Au-delà du droit ouvert par la loi numérique de 2016, les principaux financeurs de la recherche ont rendu le dépôt en accès ouvert obligatoire dans le cadre de leurs conditions de financement.

- Agence Nationale de la Recherche (ANR) : depuis les appels à projets (AAP) 2022, toutes les publications issues de projets ANR doivent être disponibles en accès ouvert immédiat sous licence CC-BY. Les auteurs doivent déposer le texte intégral (version éditeur ou manuscrit accepté) dans l'archive ouverte nationale HAL en indiquant le code décision du projet ANR (ex. ANR-22-CE64-0001), afin de permettre le suivi automatique des publications. La Stratégie de Non-Cession des Droits (SNCD) doit être appliquée lors de la soumission. Par ailleurs, un Plan de Gestion des Données (PGD) est exigé dans les 6 mois suivant le démarrage du projet, et une version finale est attendue à la clôture.
- Horizon Europe : le programme-cadre européen de financement impose l'accès ouvert immédiat (sans embargo) à toutes les publications scientifiques évaluées par les pairs, sous licence ouverte (CC-BY pour les articles). Une copie de la version éditeur ou du manuscrit accepté doit être déposée dans un entrepôt de confiance dès la publication. La gestion des données de recherche selon les principes FAIR est obligatoire, et un Plan de Gestion des Données (DMP) doit être fourni dans les 6 premiers mois du projet, puis régulièrement mis à jour. Les coûts liés à l'accès ouvert (APC pour des revues entièrement ouvertes) sont éligibles au financement à condition d'avoir été budgétés dans la proposition.

Quelle version déposer ?

- Version preprint : Il s'agit de la version initiale de l'article rédigée par l'auteur avant évaluation par les pairs, ou avant soumission à une revue, ou encore après refus de publication. Cette version peut souvent être déposée librement dans une archive ouverte. Les règles de chaque éditeur peuvent être vérifiées sur SHERPA/RoMEO.
- Version postprint (Author Accepted Manuscript – AAM) : Il s'agit de la version finale du manuscrit acceptée par l'éditeur après évaluation par les pairs, mais sans la mise en forme de la revue. La loi pour une République numérique autorise son dépôt en accès ouvert, sous réserve du respect des éventuels embargos disciplinaires.
- Version éditeur : il s'agit de la version finale publiée avec la mise en page de l'éditeur. La loi n'autorise pas son dépôt dans une archive ouverte, sauf si l'éditeur l'a explicitement autorisé dans sa politique.

Les licences Creative Commons

Les travaux scientifiques déposés dans des archives ouvertes sont protégés par le droit d'auteur, conformément au Code de la propriété intellectuelle. Leur réutilisation reste soumise aux conditions définies par l'auteur. Il est donc recommandé d'associer une licence Creative Commons au dépôt afin de préciser les conditions de diffusion et de réutilisation.

Les licences Creative Commons sont des licences libres qui définissent les usages autorisés par l'auteur pour toute œuvre de l'esprit. L'auteur peut autoriser la copie, la diffusion et, selon les cas, la modification de l'œuvre. Creative Commons recommande notamment la licence CC BY 4.0 comme standard pour l'accès ouvert.

La stratégie de non-cession des droits (Rights Retention Strategy, RRS)

Cette stratégie permet aux auteurs de diffuser légalement leur manuscrit dans une archive ouverte comme HAL, avant et après l'évaluation par les pairs, quelle que soit la revue de publication.

Recommandée par l'Agence Nationale de la Recherche depuis les appels à projets 2022, elle consiste à informer l'éditeur, dès la soumission de l'article, qu'une licence ouverte CC BY est appliquée aux différentes versions du manuscrit jusqu'à la version acceptée. Il est également possible d'y associer la licence Etalab 2.0, recommandée par l'État français et juridiquement équivalente à la CC BY. Ce dispositif permet un dépôt immédiat en accès ouvert dans HAL dès la publication, sans embargo. Cette stratégie est soutenue par France Universités et le CNRS.

5- L'identification ORCID

ORCID est une organisation à but non lucratif qui attribue des identifiants numériques pérennes aux chercheurs et contributeurs scientifiques à l'échelle mondiale.

L'identifiant ORCID permet d'éviter les confusions liées aux homonymes entre chercheurs. Il regroupe également les différentes formes d'identité sous lesquelles un auteur peut publier (variations du nom, initiales, changements de nom, etc.).

Il relie entre eux les travaux scientifiques produits par un chercheur, même lorsqu'ils sont répartis dans différentes bases de données ou systèmes d'information. Il facilite également l'interopérabilité entre plateformes en évitant la ressaisie répétitive des informations personnelles et des références bibliographiques.

L'ORCID peut également être relié à l'IDHAL, l'identifiant auteur utilisé dans HAL. Cette connexion facilite la communication entre les deux systèmes, permet la synchronisation des informations et évite aux chercheurs de ressaisir plusieurs fois les mêmes données.

6- Open U Journals : une solution bordelaise pour créer une revue

Open U Journals est une plateforme numérique proposée par l'université de Bordeaux depuis 2016 sur un modèle Diamond.

La plateforme permet la création et la gestion de revues scientifiques en accès ouvert. Elle offre aux chercheurs du site bordelais, ainsi qu'à d'autres équipes scientifiques, la possibilité de créer et de publier des revues en assurant l'ensemble du processus éditorial. Celui-ci comprend la soumission des articles, leur évaluation par les pairs et leur mise en ligne en accès ouvert.

Un comité scientifique examine chaque demande de création de revue afin de vérifier son intégration dans le dispositif. Les revues doivent respecter les critères d'éligibilité recommandés par le Comité pour la science ouverte.

7- Le portail HAL institutionnel de l'université de Bordeaux

L'université de Bordeaux dispose de son propre portail HAL institutionnel, accessible à l'adresse <https://u-bordeaux.hal.science>.

Ce portail permet aux membres de la communauté scientifique de l'université de déposer, valoriser et rendre accessibles leurs publications en accès ouvert. Les bibliothèques de l'université proposent un accompagnement personnalisé pour les dépôts (vérification des droits, contrôle qualité des notices, ateliers HAL) via le contact : info@oskar-bordeaux.fr pour les publications.

B) Structurer et ouvrir les données de la recherche

Rendre les données de recherche accessibles et réutilisables le plus largement possible répond à plusieurs enjeux majeurs.

- Favoriser l'innovation et la collaboration : permettre la réutilisation des données pour de nouvelles applications potentielles et encourager la création de collaborations entre producteurs et utilisateurs des données.
- Assurer la préservation des données et réduire les coûts : préserver des données uniques ou coûteuses à produire, éviter les coûts liés à la duplication de données existantes et limiter les risques de perte de données.

- Garantir la transparence scientifique : faciliter la reproduction et la réplication des recherches, permettre la vérification et la précision des résultats, et contribuer à décourager la fraude scientifique.
- Valoriser le travail des chercheurs : générer une reconnaissance directe pour les chercheurs ayant produit les données, notamment par la publication de *data papers*¹ ou par l'adaptation des pratiques de citation. Certaines revues encouragent non seulement la citation des jeux de données, mais les intègrent également dans la bibliographie principale des articles scientifiques.
- Répondre aux obligations institutionnelles et réglementaires : appliquer les exigences des revues scientifiques ou des agences de financement, telles que celles de la Commission européenne.

1- Préparer et documenter des données

La préparation et la documentation des données de recherche, notamment à travers les métadonnées, sont essentielles pour garantir leur compréhension et leur réutilisation. Le Plan de gestion des données (PGD) permet d'organiser ce travail tout au long du projet.

Le partage des données doit tenir compte des contraintes juridiques et des politiques des établissements. Il est recommandé d'éviter les *supplementary data* proposés par les éditeurs, souvent peu adaptés à une bonne documentation et à une réutilisation efficace.

Il est préférable de déposer les données dans des entrepôts institutionnels, disciplinaires ou généralistes, qui garantissent leur pérennité, leur accessibilité et leur réutilisation. Enfin, le lien entre données et publications doit être anticipé dès le début du projet.

2- Comment choisir un entrepôt de données ?

La démarche recommandée pour le partage des données de recherche consiste à :

- Privilégier un entrepôt de données disciplinaire lorsqu'il en existe un, par exemple PANGAEA, 4TU.ResearchData, NAKALA ou Harvard Dataverse ;
- Recourir, à défaut, à un entrepôt généraliste tel que Zenodo ou B2SHARE ;

Le choix d'un entrepôt de données repose sur plusieurs critères, notamment les formats pris en charge, les possibilités d'indexation ainsi que les modalités d'accès aux données (accès libre, restreint ou sous embargo). Le recours à des entrepôts certifiés constitue également un critère important afin de garantir la qualité, la pérennité et la fiabilité du service proposé. Le portail re3data.org facilite l'identification de ces entrepôts grâce à son filtre « Certificates ».

3- Outil d'aide au choix d'un entrepôt de données

La direction de la documentation de l'université de Bordeaux met à disposition un outil d'aide à la décision librement accessible en ligne intitulé « Trouver un entrepôt de données ». Cet

¹ Un *data paper* est une publication scientifique dont l'objectif principal est de décrire un jeu de données de recherche. Contrairement à un article scientifique classique, il propose une description détaillée d'un jeu de données et des métadonnées associées, ainsi que des conditions et des méthodes de collecte, sans présenter de résultats scientifiques fondés sur leur analyse.

outil permet d'identifier l'entrepôt de données généraliste le plus adapté aux besoins des chercheurs à partir de quatre critères principaux :

- La nécessité de restreindre l'accès aux données ;
- L'existence d'options de dépôt gratuites ;
- La possibilité de définir librement les conditions d'utilisation des données ;
- La relecture du dépôt avant sa mise en ligne.

Une fiche descriptive synthétique accompagne chacun des huit entrepôts de données comparés : Zenodo, Figshare, Mendeley Data, Dryad, B2SHARE, 4TU.ResearchData, Open Science Framework et Harvard Dataverse.

4- ABDo : l'Atelier bordelais des données de recherche

ABDo (Atelier Bordelais des Données de recherche) est un service de proximité dédié aux chercheurs, doctorants, ingénieurs et personnels d'appui à la recherche de l'ensemble des établissements du site bordelais. Labellisé par le Ministère de l'Enseignement Supérieur et de la Recherche, ABDo propose un guichet unique (adresse : donnees-recherche@u-bordeaux.fr) pour : la formation et la sensibilisation à la gestion des données de recherche et aux principes FAIR ; l'appui au dépôt, au stockage, au partage et à l'archivage des données ; le conseil juridique et éthique (propriété intellectuelle, licences, RGPD) ; l'aide à la rédaction des Plans de Gestion des Données (PGD/DMP) ; et la mise en réseau pour des échanges de bonnes pratiques entre pairs.

5- L'espace institutionnel Recherche Data Gouv de l'université de Bordeaux

L'université de Bordeaux dispose d'un espace institutionnel dédié au sein de l'entrepôt national pluridisciplinaire Recherche Data Gouv, la plateforme souveraine française dédiée au partage des données de recherche (<https://entrepot.recherche.data.gouv.fr/dataverse/univ-bordeaux>). Basé sur le logiciel libre Dataverse et développé notamment par l'INRAE, cet entrepôt constitue une solution privilégiée pour le dépôt de données à caractère institutionnel, pour des unités souhaitant une solution nationale souveraine. Accessible à tous les membres de l'université via leur identifiant UBx (SSO), il permet de déposer ou de regrouper des jeux de données par structure, équipe ou projet. (contact : donnees-recherche@u-bordeaux.fr).

Quel que soit l'entrepôt de confiance retenu pour le partage des données, celui-ci doit garantir un ensemble minimal de fonctionnalités assurant la diffusion, la traçabilité, la conservation et la réutilisation des données de recherche. Ces fonctionnalités contribuent également au respect des principes FAIR².

Un entrepôt de données doit notamment permettre :

² Les principes FAIR reposent sur quatre dimensions : Findable (faciles à trouver), Accessible (accessibles), Interoperable (interopérables), Reusable (réutilisables).

- L'assignation d'un identifiant pérenne (*Persistent Identifier* – PID) de type DOI, permettant d'assurer la citation, la traçabilité et l'interconnexion des données avec d'autres productions scientifiques, telles que les publications ;
- Une description suffisamment détaillée des données afin d'en faciliter la découverte, la compréhension et la réutilisation, grâce à des métadonnées standardisées et à l'utilisation de vocabulaires disciplinaires contrôlés ;
- L'utilisation de licences et la définition de règles d'accès permettant d'encadrer juridiquement la réutilisation des données, dans le respect du droit français et européen ;
- Une durée minimale de conservation des données cohérente avec la politique de conservation définie par l'établissement de rattachement.

6- Lier les données aux publications

Plusieurs solutions permettent d'établir un lien entre un article scientifique et les données qui lui sont associées avant la publication de l'article. L'utilisation d'identifiants pérennes facilite ainsi l'interconnexion durable entre publications et données de recherche. De même, le référencement des publications associées aux données (y compris les *data papers*) est généralement possible dans la plupart des entrepôts de données, même après le dépôt initial des jeux de données.

Après la publication d'un article, l'ajout d'un lien vers les données associées peut être difficile, voire impossible, sur certaines plateformes éditoriales. Une alternative consiste à renseigner ces données dans la version de l'article déposée dans une archive ouverte telle que HAL, qui permet d'ajouter des identifiants pérennes dans le champ « Données associées ». Cela facilite ainsi la mise en relation entre les publications et les jeux de données.

7- FAIR et DMP

Les principes FAIR

Les principes FAIR ont pour objectif de favoriser la découverte, l'accès, l'interopérabilité et la réutilisation des données de recherche. Ils définissent un ensemble de caractéristiques que doivent présenter les données et leurs métadonnées afin de faciliter leur exploitation par les humains comme par les machines.

Le site de l'INRAE propose des ressources détaillées en français sur les principes FAIR : <https://science-ouverte.inrae.fr/fr/le-numerique-pour-la-science-et-les-donnees-scientifiques>

10 actions pour rendre ses données plus FAIR

Des guides élaborés de manière collaborative proposent des recommandations concrètes pour améliorer la FAIRisation des données et des logiciels. La ressource intitulée *Top 10 FAIR Data & Software Things* présente, pour différents domaines disciplinaires, dix actions clés permettant de rendre les données et le code plus faciles à trouver, accessibles, interopérables et réutilisables. Chaque recommandation est accompagnée d'exemples pratiques.

Ces guides sont disponibles en accès libre sous forme de document PDF, notamment via la plateforme Zenodo, sous la référence suivante : Christopher Erdmann et al. (2019), *Top 10 FAIR Data & Software Things*.

Comment évaluer le niveau FAIR des données ?

Plusieurs outils permettent d'évaluer dans quelle mesure des données respectent les principes FAIR :

- How FAIR are your data ? : une liste de contrôle simple sous forme de cases à cocher permettant une auto-évaluation rapide ;
- FAIR Self-Assessment Tool : un outil en ligne développé par l'Australian Research Data Commons, qui génère un score global ainsi que des scores détaillés pour chacun des principes FAIR.

Ce dernier propose une série de questions pour chaque principe, avec des réponses sous forme de menus déroulants, permettant une évaluation structurée et reproductible.

Le DMP

Le DMP (Data Management Plan), ou PGD (Plan de Gestion de Données), est un document synthétique qui permet d'organiser et d'anticiper l'ensemble du cycle de vie des jeux de données de recherche. Il décrit, pour chaque jeu de données, les modalités de gestion mises en œuvre depuis leur création ou collecte jusqu'à leur partage et leur archivage.

Selon l'INIST-CNRS, il s'agit d'un document « qui aide à organiser et anticiper toutes les étapes du cycle de vie de la donnée » et qui précise « comment seront gérées les données d'un projet, depuis leur création ou collecte jusqu'à leur partage et leur archivage » (INIST-CNRS, 2018, Plan de gestion des données : fiche synthétique, DoRANum).

Le contenu d'un DMP varie selon le modèle utilisé, qu'il soit imposé par une institution ou choisi librement. Il couvre généralement les éléments suivants :

- La description des jeux de données et de leur organisation ;
- Les méthodes et processus de traitement des données tout au long du projet ;
- Les modalités de diffusion et de partage ;
- Les stratégies d'archivage et de conservation.

DMP OPIDoR est un outil d'aide à la planification et à la mise en œuvre de la gestion des données de recherche. Il permet de produire des plans de gestion de données structurés et exploitables par les machines. Développé par l'Inist-CNRS, il est mis à disposition de l'ensemble de la communauté scientifique française de l'enseignement supérieur et de la recherche, ainsi que de ses partenaires. Cet outil constitue l'une des composantes de Recherche Data Gouv, l'écosystème national dédié à la gestion, au partage et à l'ouverture des données de recherche.

Afin de favoriser l'adoption des principes FAIR et de répondre aux recommandations de la Research Data Alliance, DMP OPIDoR propose un modèle commun structuré de plan de gestion de données. Recommandé par le réseau de la science ouverte des agences de financement françaises, ce modèle est de plus en plus largement adopté par les organismes

de recherche. Il est extensible et personnalisable, ce qui permet d'intégrer les spécificités disciplinaires ainsi que la gestion des codes et logiciels associés aux projets de recherche.

C) S'inscrire dans une dynamique durable, européenne et internationale

La science ouverte est aujourd'hui un mouvement mondial en pleine structuration, qui transforme en profondeur les modes de production, de diffusion et d'évaluation des connaissances scientifiques.

Dans ce paysage, de nombreuses infrastructures européennes et internationales jouent un rôle structurant. L'une des initiatives majeures est l'European Open Science Cloud (EOSC), qui vise à construire un environnement européen interopérable pour le partage et la réutilisation des données de recherche. L'EOSC ne se limite pas à une plateforme technique : il s'agit d'un écosystème fédéré destiné à connecter les infrastructures existantes et à faciliter l'accès aux ressources scientifiques produites en Europe.

D'autres initiatives complètent ce dispositif. GO FAIR promeut l'adoption des principes FAIR à l'échelle internationale, en encourageant la mise en œuvre de pratiques harmonisées pour la gestion et la publication des données. La Research Data Alliance (RDA) constitue une communauté mondiale de chercheurs, d'experts et d'institutions qui travaillent à l'élaboration de recommandations et de bonnes pratiques pour le partage des données. Elle joue un rôle essentiel dans la définition de standards ouverts, favorisant l'interopérabilité entre disciplines et entre pays.

Dans le domaine de la publication scientifique, OpenAIRE contribue à structurer un écosystème ouvert en Europe en facilitant la découverte, l'accès et la liaison entre publications, données et projets de recherche. De même, le Directory of Open Access Journals (DOAJ) joue un rôle central dans la visibilité des revues scientifiques en accès ouvert, en proposant un répertoire de référence garantissant des critères de qualité et de transparence éditoriale. Dans le champ des infrastructures de publication alternatives, OPERAS soutient le développement des sciences humaines et sociales en promouvant des services ouverts et durables adaptés à leurs besoins spécifiques.

Enfin, des initiatives comme SCOSS (Global Sustainability Coalition for Open Science Services) contribuent à la pérennisation des infrastructures de science ouverte. Elles soutiennent financièrement des services essentiels à l'écosystème, souvent portés par des communautés académiques, afin d'assurer leur stabilité et leur développement à long terme.

Dans ce contexte international, la France s'inscrit pleinement dans cette dynamique de coopération. Elle participe activement à de nombreuses initiatives et soutient le développement d'infrastructures ouvertes à l'échelle européenne et mondiale. Parmi ces engagements figure la création d'une fondation franco-néerlandaise dédiée au Directory of Open Access Books (DOAB), une plateforme internationale de référencement des ouvrages scientifiques en accès ouvert.

L'ensemble de ces dispositifs illustre la construction progressive d'un écosystème international de la science ouverte, fondé sur la coopération, l'interopérabilité et la

mutualisation des ressources. Au-delà des infrastructures techniques, il s'agit d'un véritable projet scientifique et politique visant à rendre la connaissance plus accessible, plus équitable et plus durable à l'échelle mondiale.

Pour aller plus loin :

Des formations à la science ouverte sont proposées dans le catalogue des formations du Collège des écoles doctorales via ADUM (avec validation des heures) :

2 sessions en français à distance :

- Découvrir la science ouverte (3h)
- Valoriser ma thèse et mes publications scientifiques avec l'accès ouvert (3h)
- Gestion des données de recherche : fondamentaux (3h)

1 session en anglais à distance :

- Open for you! An introduction to open science (3h)
- Promote my thesis and my scientific publications with open access (3h)
- Research Data Management: Fundamentals (3h)

Sources :

<https://www.u-bordeaux.fr/recherche/ambition-scientifique/science-ouverte>

<https://www.ouvri.la-science.fr/plan-national-pour-la-science-ouverte>

<https://bibliotheques.u-bordeaux.fr/appui-aux-chercheurs>

<https://intranet.u-bordeaux.fr/personnels/appuis-aux-missions/appui-la-recherche/gerer-ses-donnees-de-recherche>