



**UNIVERSITÉ DE
MONTPELLIER**

Introduction à la Science Ouverte

Webinaire du 01 février 2022

Equipe d'appui à la recherche de l'Université de Montpellier



— Objectifs et plan de la formation

- **Comprendre le contexte de la science ouverte**
- **Maîtriser les enjeux liés aux publications et aux données de la recherche**

Axes du webinaire :

- I. **Science Ouverte : de quoi parle-t-on ?**
- II. **L'Open Access**
- III. **Les données de la recherche**
- IV. **La Science Ouverte à l'UM**

La Science ouverte : définition

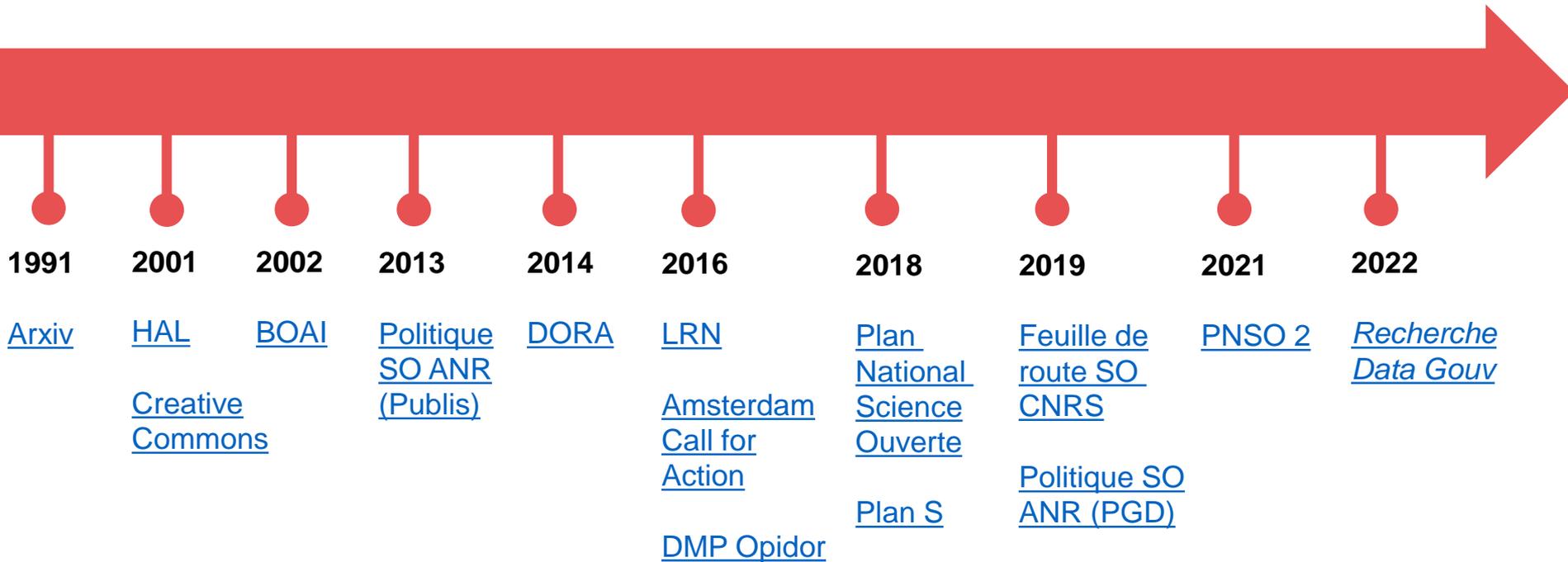
La science ouverte est la diffusion sans entrave des publications et des données de la recherche. Elle s'appuie sur l'opportunité que représente la mutation numérique.

https://www.ouvrirlascience.fr/category/science_ouverte/

La science ouverte est la pratique de la science qui rend possible la collaboration et la contribution avec d'autres, où les données de recherche, les notes de laboratoire et autres processus de recherche sont librement disponibles, dans des conditions permettant la réutilisation, la redistribution et la reproduction de la recherche, ainsi que de ses données et méthodes sous-jacentes.

<https://www.fosteropenscience.eu/foster-taxonomy/open-science-definition>

Développement de la Science ouverte



Amsterdam Call for action 2016

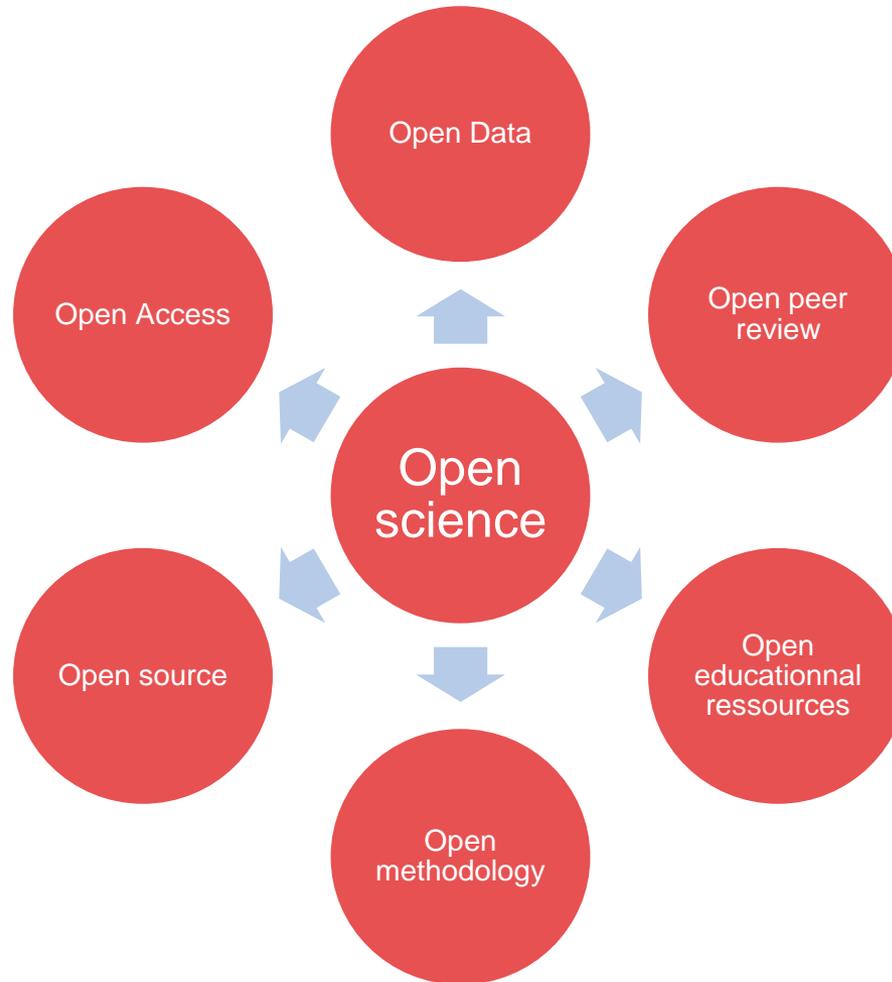
Deux objectifs importants au niveau européen ont été formulés pour 2020 :

1. Libre accès total pour toutes les publications scientifiques
2. Une approche fondamentalement neuve en termes de réutilisation optimale de données de la recherche

Pour les atteindre :

- Nouveaux systèmes d'appréciation, promotion et d'évaluation de la recherche
- Alignement de politiques et échange de bonnes pratiques

Périmètre de la science ouverte



Enjeux de la Science ouverte

Enjeux académiques

- Faciliter l'accès à l'information scientifique et technique
- Faciliter les échanges entre pairs, la circulation des idées
- Garantir l'objectivité des connaissances

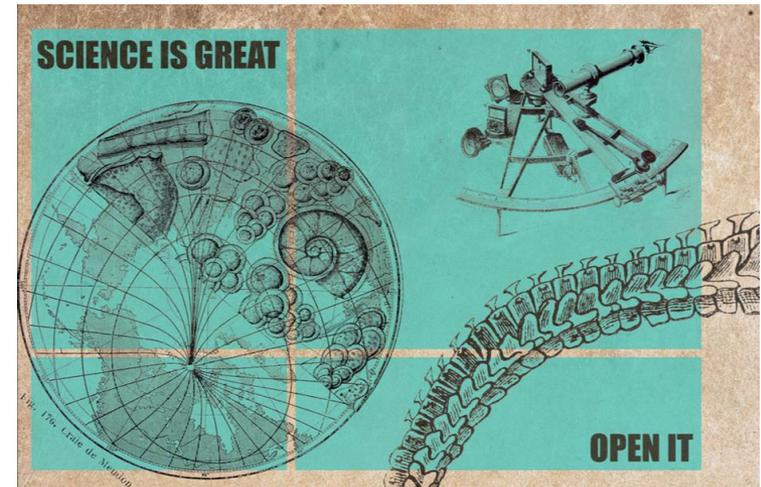
Enjeux politiques

Restaurer la confiance (mouvements anti-science)

- Science au cœur de la démocratie
 - besoin d'information citoyen
 - sciences participatives, forums hybrides
 - pilier des démocraties modernes

Enjeux économiques

- Efficience des dépenses publiques
- Economie de la connaissance
 - partage générateur de croissance, de réduction des coûts



By Martin Clavey – CC-BY-SA

Plan

I. Science ouverte : de quoi parle-t-on ?

II. L'Open Access

III. Les données de la recherche

IV. La Science ouverte à l'UM

Plan

II. L'Open Access

1. Définition et généralités

2. Le contexte national et international

Définition de l'Open Access

Déclaration de Budapest (2002)

« **Mise à disposition gratuite de la littérature scientifique sur l'Internet public**
..., **sans barrière financière, légale ou technique** autre que celles indissociables
de l'accès et l'utilisation d'Internet
...garantir aux auteurs **un contrôle sur l'intégrité de leurs travaux et le droit à
être correctement reconnus et cités.** »

Voie verte (archive ouverte)

Réservoir en ligne où déposer la
production scientifique
Accès gratuit aux documents

Voie dorée (revue en libre accès)

Revue scientifique à comité de lecture
Accès gratuit aux articles publiés

**L'Open Access est un mouvement international visant à rendre les publications
scientifiques accessibles librement et gratuitement pour le lecteur**

Les archives ouvertes (voie verte)



Photo by [Andea Ferrario](#) on [Unsplash](#)

- Site web où les documents sont librement accessibles
- Archivage pérenne et gratuité d'accès
- Productions scientifiques publiées ou non (articles, communications, rapports, thèses, preprints,...)

Archive nationale HAL (+ portails d'établissements)

Archives institutionnelles (Okina, Agritrop, Orbi)

Archives thématiques (BioRxiv, Pubmed, RepeC,...)

bioRxiv
THE PREPRINT SERVER FOR BIOLOGY

PubMed

RePEc



Coronavirus (COVID-19)

Consulter et télécharger toutes les publications disponibles dans HAL

[➤ Accéder aux publications](#)

RECHERCHE

Termes de recherche (* pour tous)



NOMBRE DE DOCUMENTS SCIENTIFIQUES

1 011 773

NOMBRE DE RÉFÉRENCES

2 884 384

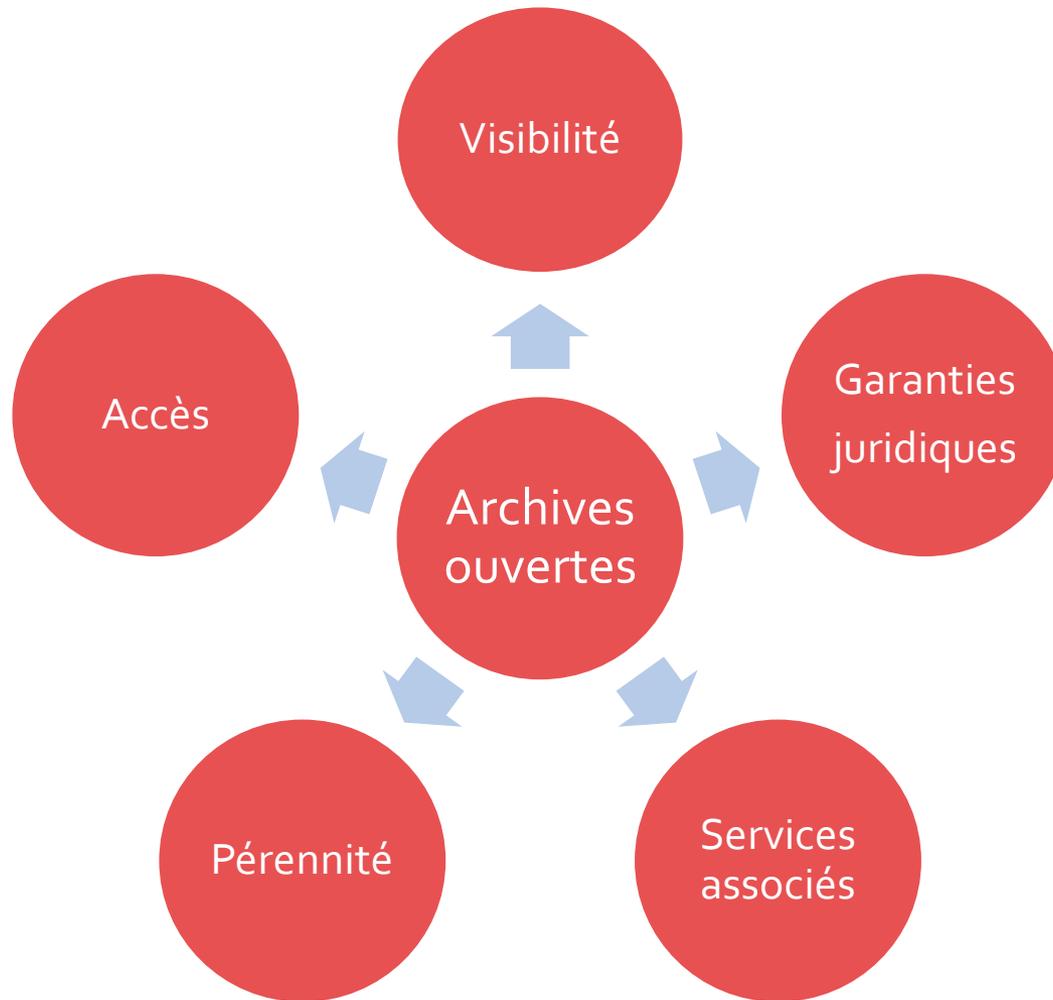
Archive ouverte HAL

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion d'articles scientifiques de niveau recherche, publiés ou non, et de thèses, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

À l'attention du déposant

- Le dépôt du texte intégral est effectué en accord avec les co-auteurs et dans le respect de la politique des éditeurs
- La mise en ligne est assujettie à une modération, la direction de HAL se réservant le droit de refuser les articles ne correspondant pas aux critères de l'archive
- Tout dépôt est définitif, aucun retrait ne sera effectué après la mise en ligne de l'article
- Les fichiers textes au format pdf ou les fichiers images composant votre dépôt sont maintenant envoyés au CINES dans un contexte d'archivage à long terme.

Atouts des archives ouvertes



Publier dans une revue Open Access (voie dorée)

Une revue en libre accès est une revue scientifique avec comité éditorial et processus de relecture, qui donne accès aux articles publiés gratuitement pour le lecteur.

Différents modèles économiques existent pour la financer.

Modèle hybride

accès à la revue via abonnement + facturation de frais de publication aux auteurs (APC), montants très variables
A éviter absolument !

Modèle auteur-payeur

facturation de frais de publication aux auteurs (APC), montants très variables
Attention éditeurs prédateurs => DOAJ

Modèle freemium

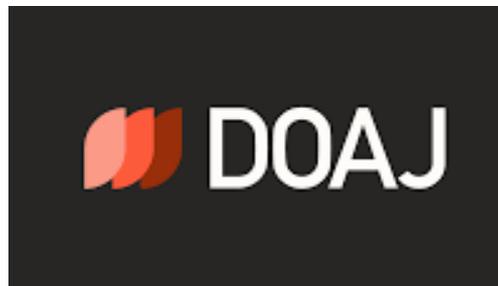
accès au format html gratuit mais des services à valeur ajoutée réservés aux abonnés (pdf, stats, etc)
Ex : OpenEdition

Modèle diamant

financement amont par une institution (société savante, université, bibliothèque, association, etc)

Publier dans une revue Open Access (voie dorée)

[Le DOAJ](#) est une liste blanche internationale des revues en libre accès. Elles y sont ajoutées sur la base de critères de qualité éditoriale.



+ 17000 revues répertoriées

+ 12000 gratuites pour l'auteur (modèle diamant)

Plan

II. L'Open Access

1. Définition et généralités

2. Le contexte national et international

Loi pour une République Numérique (2016)

Objectifs :

- Protéger les auteurs
- Favoriser la libre circulation des idées



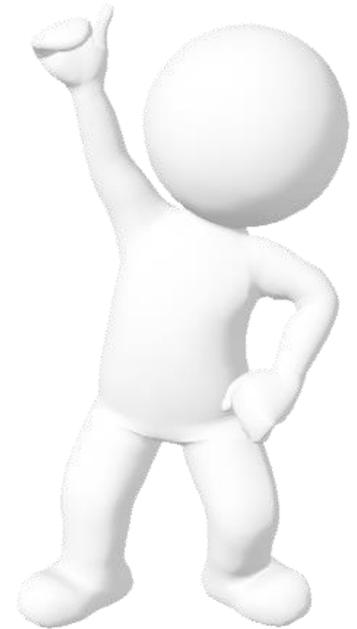
Moyen : Introduire un droit d'exploitation secondaire au bénéfice des auteurs

L'article 30 permet aux auteurs de mettre à disposition gratuitement leurs écrits scientifiques dans un format ouvert

Loi pour une République Numérique (2016)

Loi république numérique :

Un auteur dont la recherche est financée (au moins 50%) par des fonds publics peut déposer la version postprint de son article gratuitement en ligne, immédiatement ou après embargo de 6 mois pour les STM et 12 mois pour les SHS, même s'il a cédé tous ses droits à l'éditeur !



Loi pour une République Numérique (2016)

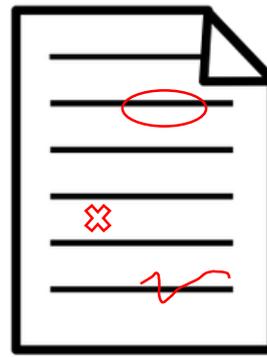
Quelle version déposer ?

- PDF éditeur si l'article est librement accessible depuis le site de l'éditeur
- Postprint ou « manuscrit accepté pour publication » ou « Author Accepted Manuscript (AAM) » si l'article est accessible sous abonnement



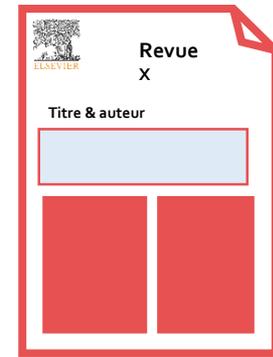
Preprint

Relecture



Postprint

*Mise en
forme*



PDF éditeur

Un contexte national favorable 1/2

Plan National Science Ouverte 2018



1. Généraliser l'accès ouvert aux publications

- Obligation de publication en OA des articles et **monographies**
- Reconnaissance institutionnelle des pratiques de science ouverte

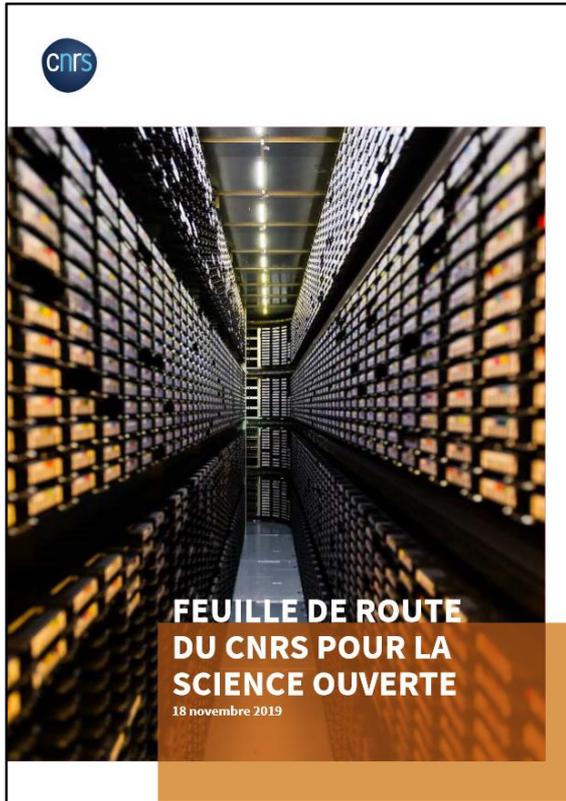
2. Structurer et ouvrir (si possible) les données de la recherche

- Obligation d'ouverture des données issues de projets financés sur fonds publics
- Généraliser la rédaction de plan de gestion de données (ANR 2019)

3. Favoriser le développement des pratiques de science ouverte

- Généraliser les compétences « Science Ouverte » (labellisation écoles doctorales)

La feuille de route Science Ouverte du CNRS



<https://www.science-ouverte.cnrs.fr/>

- Les publications
- Les données de la recherche
- La fouille et l'analyse des textes et des données
- **L'évaluation individuelle** des chercheurs et des chercheuses et la science ouverte
- La refondation de l'Information scientifique et technique pour la science ouverte
- La formation et les compétences
- Le positionnement international

Un contexte national favorable 2/2

Plan National Science Ouverte 2021

Généraliser la science ouverte en France 2021-2024



1e axe : Généraliser l'accès ouvert aux publications

2e axe : Structurer, partager et ouvrir les données de la recherche

3e axe : Ouvrir et promouvoir les codes sources produits par la recherche

4e axe : **Transformer les pratiques pour faire de la SO le principe par défaut**



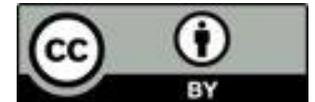
<https://www.ouvrirlascience.fr/>

Plan S de cOAlition S

« Le Plan S exige qu'à **partir de 2021**, les publications scientifiques résultant de recherches financées par des subventions publiques soient publiées dans des revues ou des plateformes en libre accès. »



Open Access immédiat = licence CC-BY et pas d'embargo



Plan

I. Science ouverte : de quoi parle-t-on ?

II. L'Open Access

III. Les données de la recherche

IV. La Science ouverte à l'UM

Plan

III. Les données de la recherche

1. Définition et généralités

2. Le plan de gestion des données (PGD)

3. Les entrepôts de données

Définition des données de la recherche

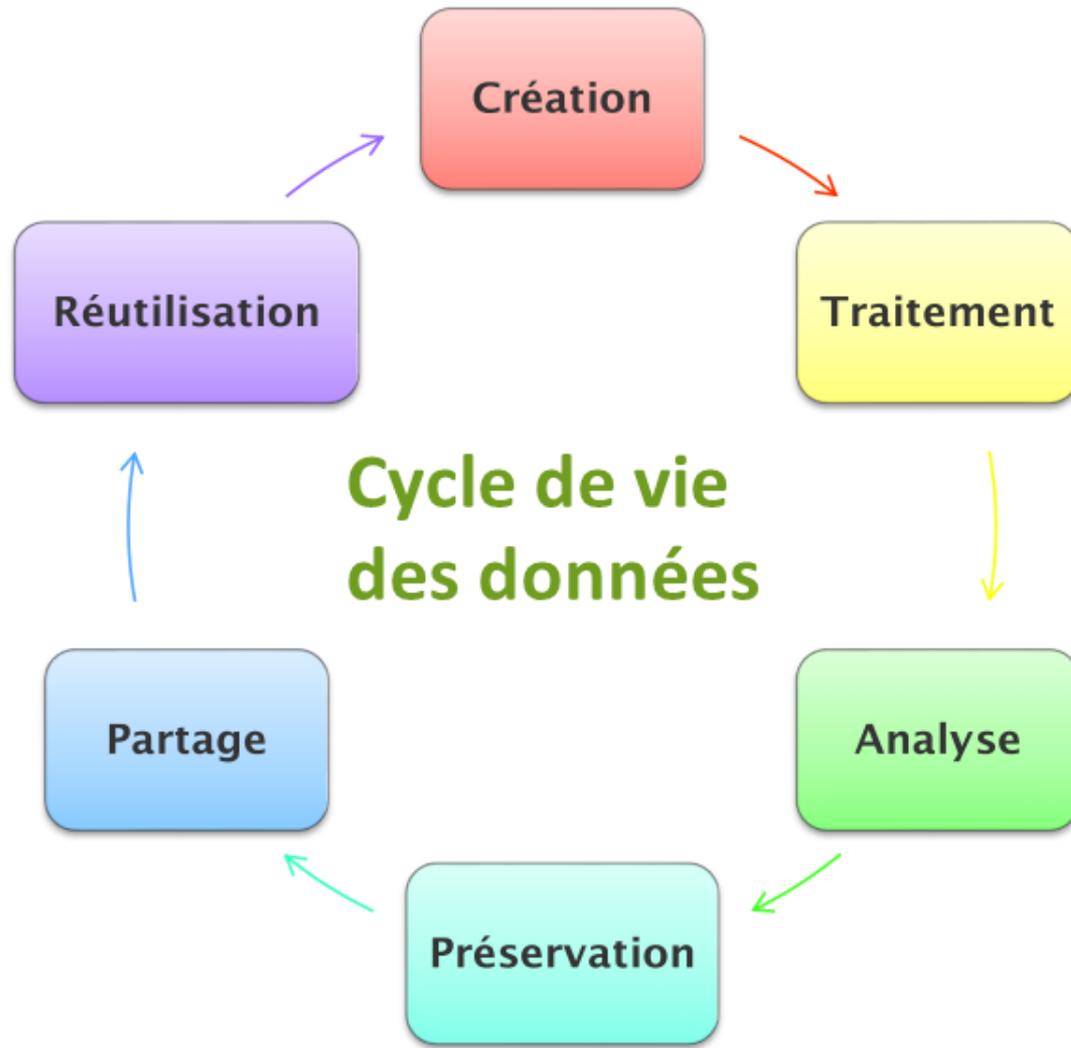
« Les données de la recherche sont définies comme des enregistrements factuels (chiffres, textes, images et sons), qui sont utilisés comme sources principales pour la recherche scientifique et sont généralement reconnus par la communauté scientifique comme nécessaires pour valider les résultats de la recherche. »

Définition de l'OCDE, rapport 2017 - *Principes et lignes directrices pour l'accès aux données de la recherche financée sur fonds publics*

Elles peuvent avoir différents :

- **supports** : carnets de laboratoire, documents électroniques, logiciels, programmes informatiques...
- **types** : archives, audio, vidéo, bases de données, codes sources, données géospatiales, images, photographies, numérisations, scans, qualitatives, quantitatives, statistiques...
- **natures** : brutes, expérimentales, produites ou collectées, issues d'observation ou de simulation...

Le cycle de vie des données

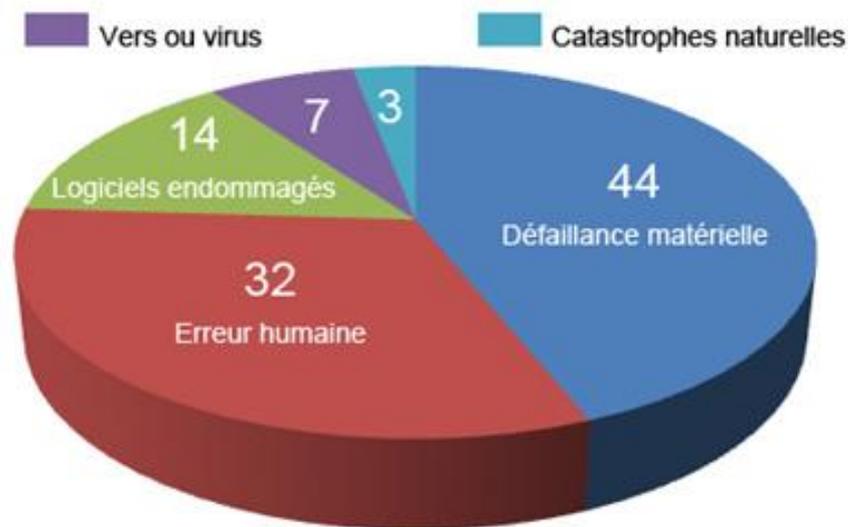


Source : <https://data.ird.fr/utiliser-des-donnees/>

Avant toute chose, pourquoi gérer les données ?

20 ans après publication, 80 % des données scientifiques sont perdues...

- Pour les auteurs
- Pour les équipes, organismes, institutions
- Pour la recherche mondiale
- **Causes :**
 - Destruction des supports
 - Problème de version, obsolescence matérielle ou logicielle
 - Lieu de stockage indéfini
- **Conséquences :**
 - Perte de temps, d'argent (fonds publics...)
 - Pas de vérification des résultats possible
 - Pas de comparaison des résultats dans le temps ou l'espace
 - Pas de réutilisations (différentes)

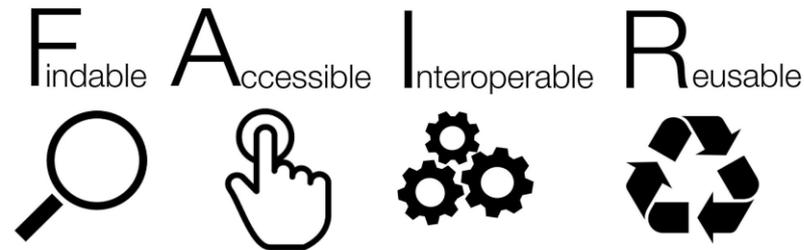


Enjeux du partage et de l'ouverture des données

- Enjeux **patrimoniaux**
- Enjeux **économiques**
 - Valeur économique de la donnée
 - Réutilisation des données et exploitation des résultats de recherches
 - Innovation et retour sur Investissement dans la R&D
- Enjeux **scientifiques**
 - Pour la communauté : science ouverte, « data driven science », interdisciplinarité, encourager la collaboration, éviter la duplication des efforts, augmenter les possibilités de financements
 - Pour le scientifique : gain en termes de visibilité et de citations
- Enjeux **sociétaux**
 - Participation des citoyens et de la société civile : "Citizen Science"
 - Plus de transparence du processus scientifique, plus de confiance en la recherche

Source : <http://www6.inra.fr/datapartage/Partager-Publier/Les-enjeux-pour-l-Inra>

Les principes FAIR de partage des données



- **Faciles à trouver** : dotées d'identifiants uniques et pérennes (ex. DOI), décrites par des métadonnées précises, déposées dans un entrepôt de données
- **Accessibles** : récupérables par des protocoles de communication standardisés, libres et ouverts, avec authentification en cas de données sensibles
- **Interopérables** : décrites avec un vocabulaire contrôlé permettant l'interopérabilité et les liens entre les données
- **Réutilisables** : publiées avec une licence d'utilisation claire et accessible, avec indication de leur provenance

Plan

III. Les données de la recherche

1. Définition et généralités

2. Le plan de gestion des données (PGD)

3. Les entrepôts de données

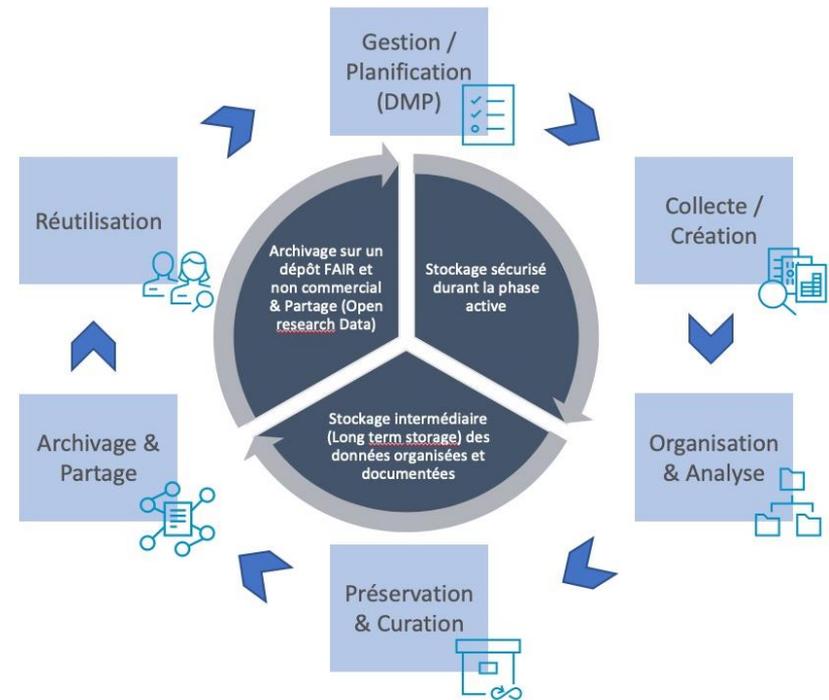
Le Plan de Gestion de Données (PGD) : définition

Le PGD ou Data Management Plan (DMP) est un document **formalisé** rédigé de préférence au démarrage d'un projet de recherche **qui décrit la façon dont les données seront produites ou obtenues, (ré)utilisées, traitées, organisées, stockées, sécurisées, préservées, documentées, partagées, au cours et à l'issue** d'un projet, explicitant notamment leur mise à disposition.

En parallèle des PGD, on note l'émergence des **Plans de Gestion de Logiciels (PGL)**.

Le Plan de Gestion de Données (PGD) : définition

- Le PGD aide à mettre en place de **bonnes pratiques de gestion des données**. Il interroge le chercheurs sur ses pratiques à chaque étapes du cycle de vie des données.
- Il vise à garantir **des données fiables** tout au long du projet, **compréhensibles, disponibles et préservées** sur le long terme pour une réutilisation future.



Source : [Université de Lausanne](#)

Le Plan de Gestion de Données : un document administratif ?

C'est un **livrable obligatoire** qui est de plus en plus demandé par les financeurs (**Horizon Europe, ANR**) pendant les projets de recherche, voire en amont, pendant le montage du projet.

On peut également le voir comme un **atout** dans la construction et le développement **de son projet de recherche** :

- Il offre la possibilité de mettre en place de nouvelles pratiques en adéquation avec les principes FAIR pour gérer ses données
- C'est un document évolutif et collaboratif qui vise à devenir un automatisme

Le Plan de Gestion de Données en pratique

Il existe de nombreux modèles de PGD établis par les financeurs (ANR) et les institutions (UM). La plupart des modèles suivent la même structure en **6 sections**.

On y retrouve les mêmes éléments :

- Informations administratives et responsabilités
- **Description des données et des métadonnées**
- **Stockage et partage** des données durant le projet
- Stockage et partage des données dans un **entrepôt**
- **Archivage pérenne**
- **Aspects juridiques et éthiques**

Pour simplifier la rédaction d'un PGD, il existe deux outils collaboratifs et gratuits en ligne :

- [DMP Opidor](#) (développé par l'Inist-CNRS)
- [Argos](#) (développé par Openaire, utile pour les projets européens)



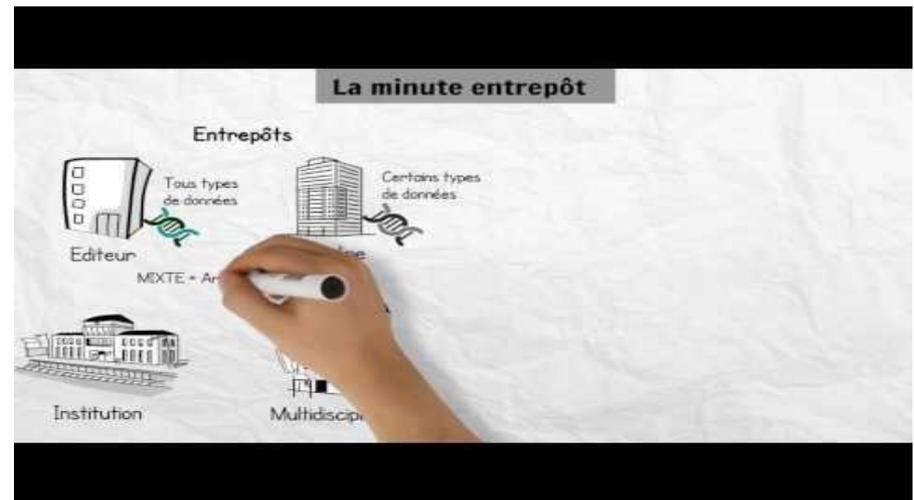
Plan

III. Les données de la recherche

1. Définition et généralités
2. Le plan de gestion des données (PGD)
- 3. Les entrepôts de données**

Rechercher dans les entrepôts de données

- Un entrepôt est un service en ligne permettant le dépôt, la description, la conservation, la recherche et la diffusion des jeux de données et des codes sources.
- Les entrepôts peuvent être généralistes, thématiques, publics ou propriétaires.
- Fonctionnalités :
 - Les entrepôts de données intègrent des moteurs de recherche avec des **filtres** pour faciliter l'accès aux données qu'ils contiennent.
 - Les interfaces de recherches peuvent être plus avancées suivant les entrepôts.

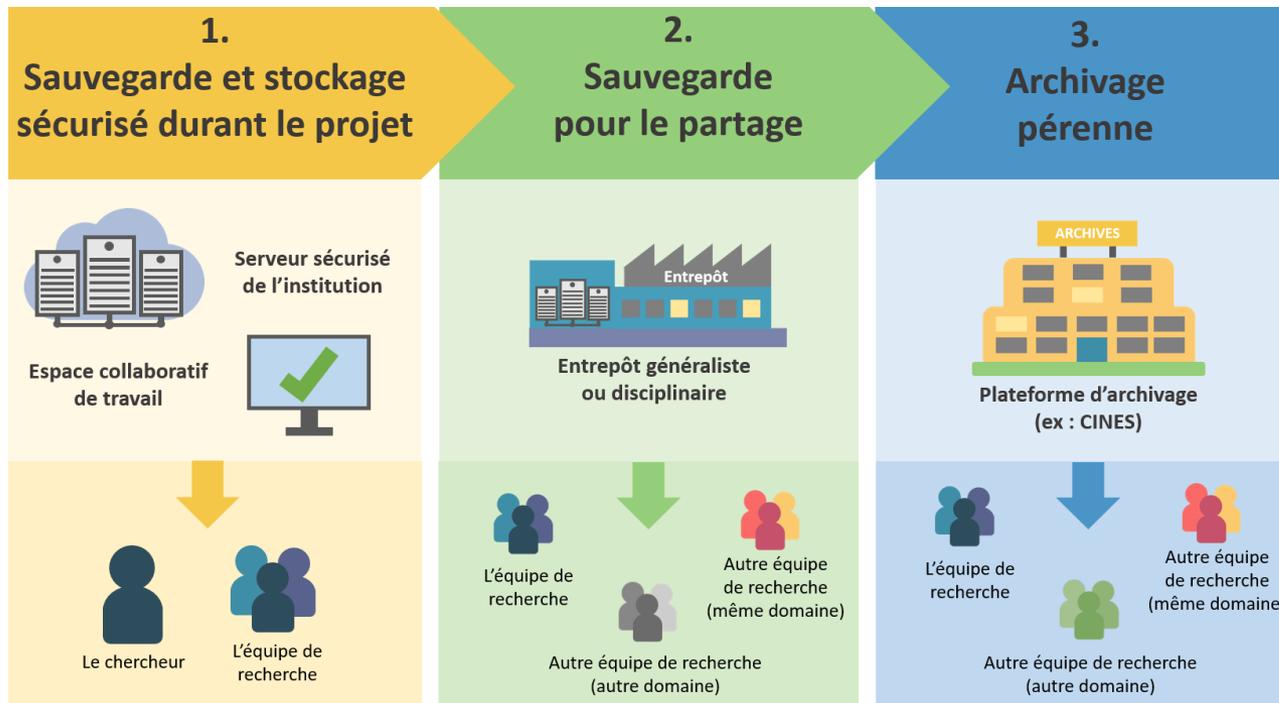


source : <https://doranum.fr/depot-entrepots/minute/>

Conserver et partager avec les entrepôts de données

Pour minimiser les risques de perte :

- **Obsolescence technologique**
- Disparition des **formats**
- Données non accompagnées de **métadonnées nécessaires** à la compréhension du contexte de la production des données



Du côté de l'université de Montpellier **Meso@LR** est un entrepôt de stockage pendant le projet de recherche (à des tarifs concurrentiels pour les chercheurs de l'université)

Typologie des entrepôts de données

Disciplinaires

- **NAKALA: entrepôt national français (SHS)**
- SEANOE: entrepôt national français (sciences marines)
- PANGAEA (sciences de la Terre et de l'environnement)
- Réseau Quetelet: entrepôt national français (sciences sociales)

Multidisciplinaires

- **ZENODO: entrepôt mixte (publications scientifiques et jeux de données) – tous types de données**
- Figshare: tous types de données
- Dryad: données sous-jacentes aux publications
- **Recherchedatagouv : tous types de données et codes sources, entrepôt national français**
- Github : pour les codes sources

Institutionnels

- Edinburgh DataShare
- Merritt
- Enlighten : Research Data
- Dataverse CIRAD

Les répertoires d'entrepôts de données

Les annuaires d'entrepôts permettent de chercher des entrepôts :

- Annuaires multidisciplinaires
 - [Re3data](#)
 - [OAD](#) (Open Access Directory)
- Annuaires spécialisés
 - [FAIRsharing](#) (sciences de la vie et biomédecine)
 - [NIH Data Sharing Repositories](#) (Biomédecine)
 - [CAT Opidor](#) (Entrepôts de données français)
- Annuaires mixtes
 - [OpenDOAR](#)
 - [ROAR](#)
- Annuaires d'entrepôts certifiés
 - [CoreTrustSeal](#)



OpenDOAR



Plan

I. Science ouverte : de quoi parle-t-on ?

II. L'Open Access

III. Les données de la recherche

IV. La Science ouverte à l'UM

La Science Ouverte à l'UM

- VP déléguée à la Science Ouverte et aux données de la recherche : **Anne Laurent**
- **Equipe appui à la recherche** composée de :
 - Equipe HAL-assistance et [portail HAL-UM](#) depuis 2017
 - Equipe données et accompagnement PGD depuis 2020
- Mini [site SO](#) depuis 2021
- Création d'un guichet d'assistance via le centre de services de l'ENT depuis 2021
- Institut de Science des Données de Montpellier ([ISDM](#))
- [Meso@LR](#) et son offre de stockage depuis septembre 2021

A venir :

- Mettre en place une solution de **cahier de laboratoire électronique**
- Proposer un hébergement de **données sensibles**

Webinaires à venir

- **Principes FAIR et ouverture des données de la recherche**

15 février 2022 de 13h à 14h

- **Déposer sa production scientifique dans HAL**

8 mars 2022 de 13h à 14h

Pour se tenir informé des webinaires thématiques
bimensuels, c'est **ici** !

Merci pour votre écoute !

Gardons contact pour la suite :

- Pour des questions autour des publications ou de l'utilisation de HAL : hal-assistance@umontpellier.fr

- Pour des questions autour de la gestion des données de la recherche ou pour toute relecture de PGD :
donnees-recherche@umontpellier.fr

- Pour toute autre question : laure.lefrancois@umontpellier.fr

Chargée de mission Appui à la recherche pour l'université

Webographie récapitulative

- [Dorandum](#) (le site incontournable d'auto-formation pour les données de recherche)

Pour en savoir plus sur :

- les [pratiques FAIR](#)
- les [entrepôts de données](#)
- le [stockage et l'archivage des données](#)
- les [PGD](#) (plan de gestion de données)
- les [PGL](#) (plan de gestion de logiciels)
- le [cycle de vie des données](#)