



Liberté  
Égalité  
Fraternité



Normandie Université

# Sensibilisation aux données de la recherche

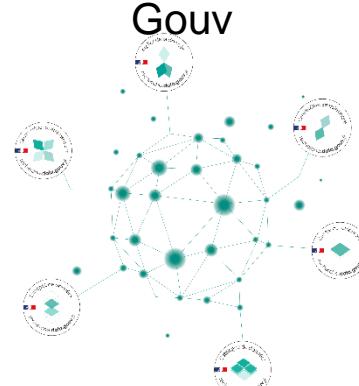
## Gestion et bonnes pratiques

08/01/2026

# Recherche Data Gouv, du national au local

## L'échelon national

L'écosystème Recherche Data Gouv



### Accompagnement des équipes :

- Ateliers de la donnée
- Centres de référence thématiques
  - Centres de ressources
  - Centres de référence établissements

### Dépôt, publication et signalement des données

- Un entrepôt
- Un catalogue

## L'échelon régional

L'Atelier de la donnée en Normandie (ADN)\*



Normandie Université



### Offre de formation normande sur les données de la recherche :

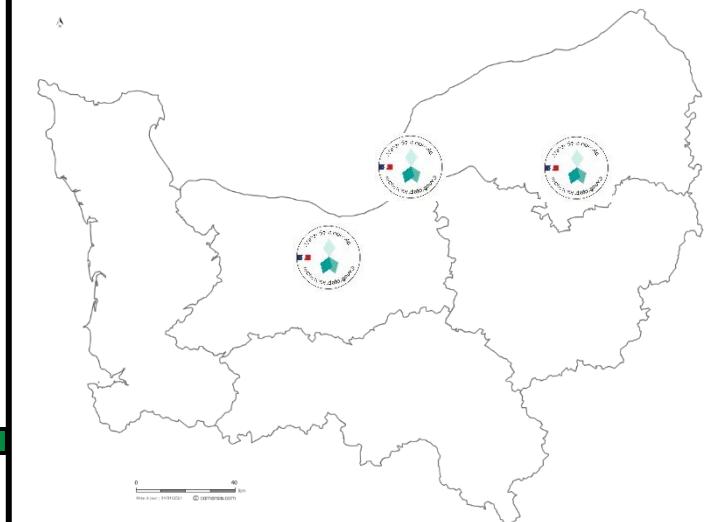
- Sensibilisation aux données de la recherche
- Rédiger un plan de gestion de données
- Diffuser ses données grâce aux *data papers*

### Organisation de journées d'études normandes sur les données de la recherche

Pour contacter l'ADN, [cliquez ici](#)

## L'échelon local

Les guichets uniques



### Accompagnement de proximité :

- Accompagnement dans la bonne gestion des données
- Relecture des plans de gestion de données
- Aide au dépôt et au signalement des données

# Sommaire

- 1. Une introduction aux données de la recherche**
- 2. Les enjeux scientifiques et éthiques de l'ouverture des données**
- 3. Contexte juridique et politique de l'ouverture des données**
- 4. Cycle de vie et principes d'ouverture**
- 5. Gérer ses données au quotidien**
  - a) Documenter ses données : les métadonnées**
  - b) Stockage, partage et archivage**
  - c) Expliquer et décrire la production et la gestion de ses données**
- 6. Rechercher des jeux de données**

# Animation de la séance



Le chat ou le micro pour répondre à des questions ouvertes

Des sondages pour les questions fermées



Données de la recherche,  
de quoi parle-t-on ?



# Sondage

Selon vous, qu'est-ce que  
les données de la recherche ?

Réponses :





# Définitions

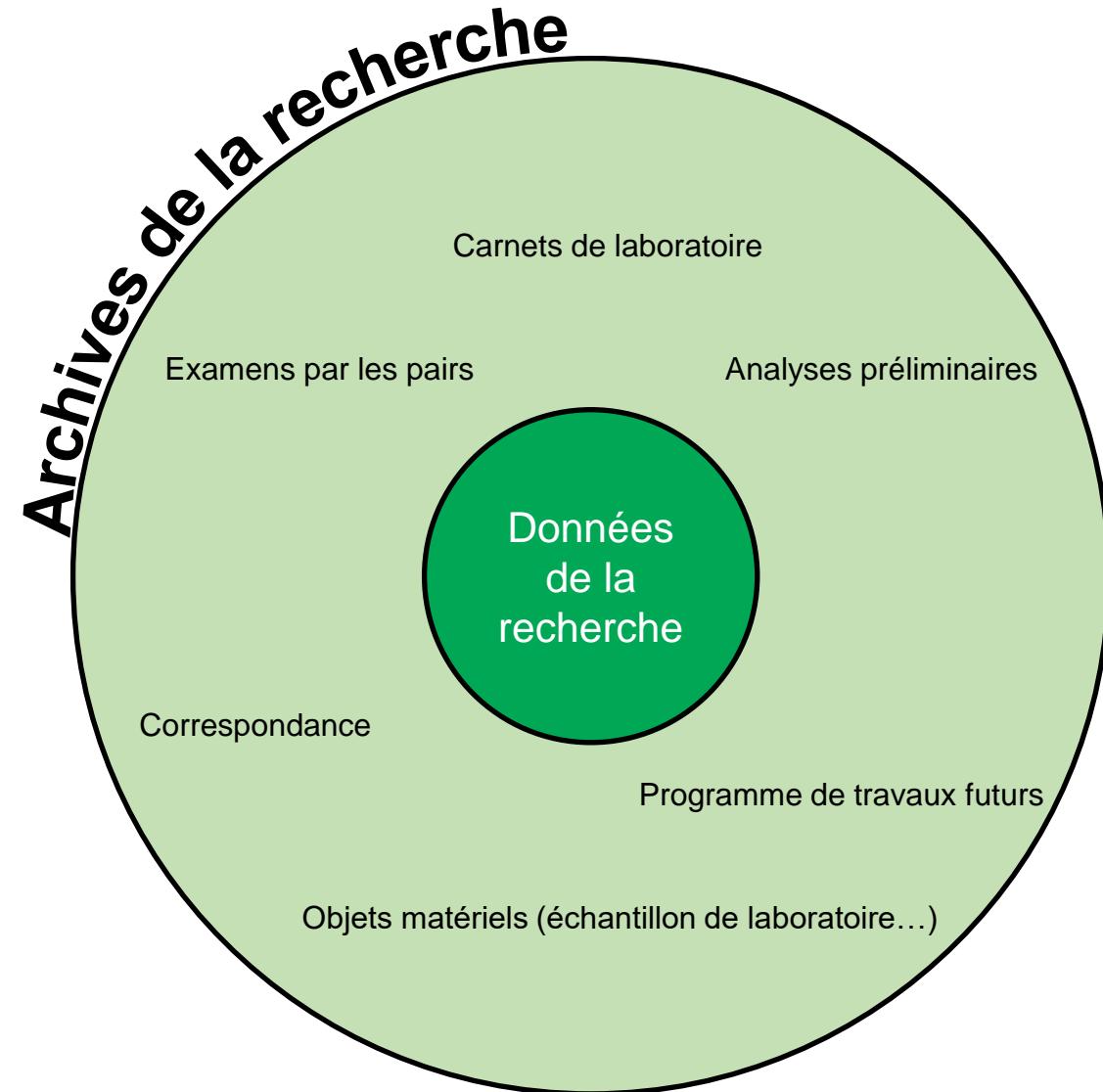
Les données de la recherche sont...

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 des résultats de recherche* » [OCDE](#) (2007)

Des « *informations, spécimens et matériaux produits, recueillis et documentés. Elles sont collectées ou exploitées à des fins de recherche et de preuves par les chercheurs et leurs équipes. À ce titre, elles constituent une partie des archives de la recherche* ». [Association des Archivistes Français](#) (2014)

Des « *documents se présentant sous forme numérique, autres que des publications scientifiques, qui sont recueillis ou produits au cours d'activités de recherche scientifique et utilisés comme éléments probants dans le processus de recherche, ou dont la communauté scientifique admet communément qu'ils sont nécessaires pour valider des conclusions et résultats de la recherche* » (art. 2-9 de la [directive 2019/1024 du Parlement européen et du conseil du 20 juin 2019 concernant les données ouvertes et la réutilisation des informations du secteur public](#))

# Des données aux archives





# Typologie des données

Plusieurs typologies existent, par exemple :

Une typologie issue des textes de lois françaises



Une trentaine d'occurrences dans les textes  
dont certaines sont juridiquement qualifiées :

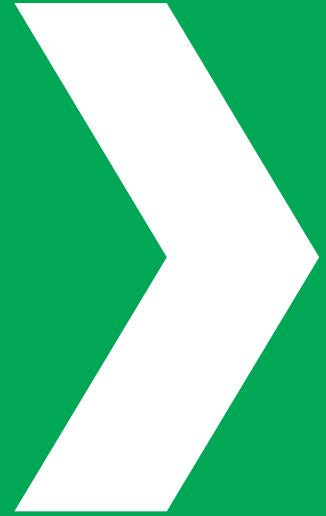
- Données confidentielles
- Données de la recherche
  - Données de santé
  - Données fiscales
- Données personnelles
- Données sensibles

- ...

Une typologie plus opérationnelle, issue  
de Recherche Data Gouv

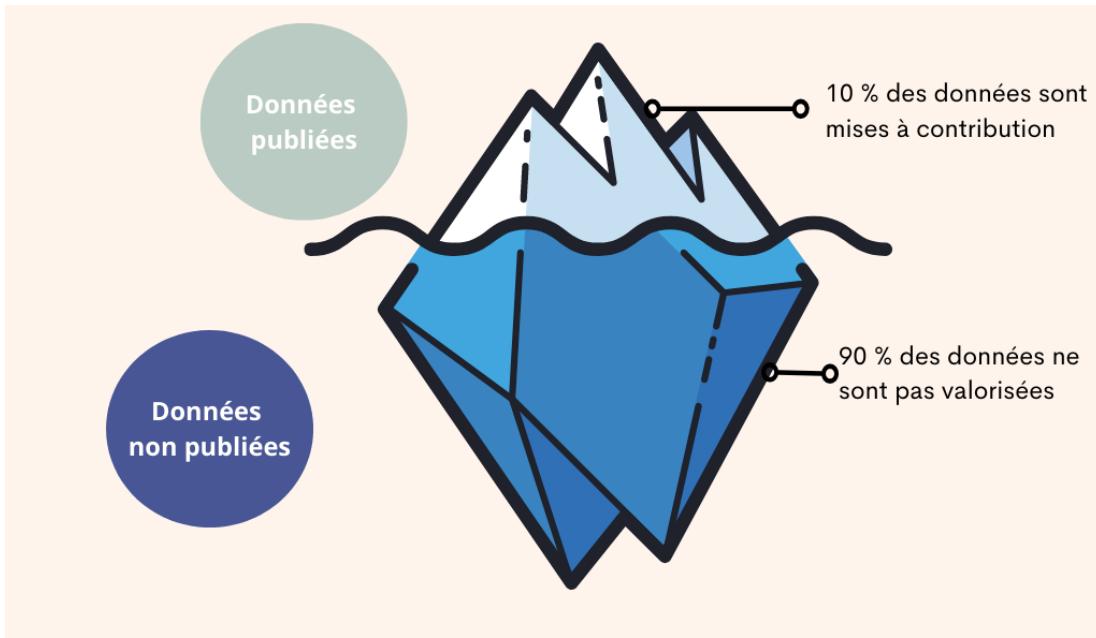


- [Données d'observation](#)
- [Données d'expérimentation](#)
  - [Données agrégées](#)
  - [Données de simulation](#)
    - [Corpus de texte](#)
    - [Corpus audiovisuel](#)



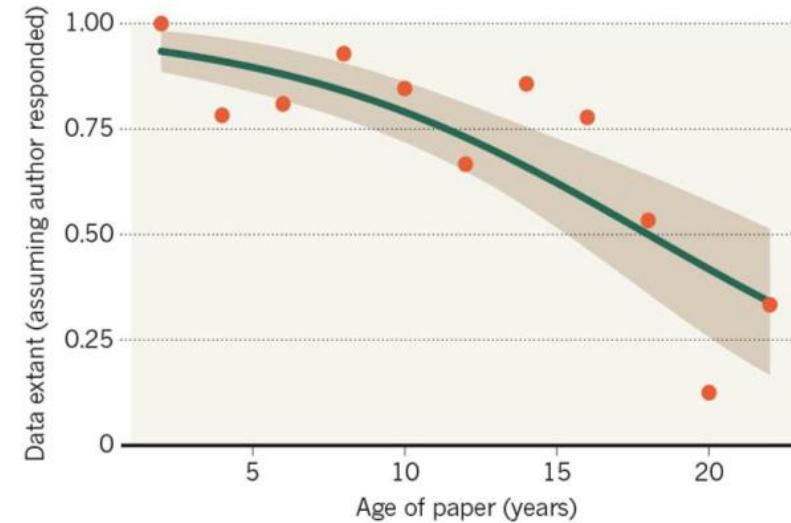
Les enjeux scientifiques et  
éthiques de  
l'ouverture des données

# Pourquoi ouvrir ses données ?



## MISSING DATA

As research articles age, the odds of their raw data being extant drop dramatically.



Elizabeth Gibney, Richard Van Noorden,  
[« Scientists losing data at a rapid rate »](#), *Nature*,  
19 décembre 2013.



# Les enjeux du partage et de l'ouverture des données

## Pour les chercheurs

- Meilleure visibilité (nationale et internationale) du travail de recherche
- Meilleure sauvegarde de ses données et possibilités de les réutiliser plusieurs années après
- Favorisation des collaborations
- Reproductibilité des recherches menées

## Pour la science

- Nouvelles perspectives grâce à la réutilisation des données
- Une meilleure validation des articles
- Renforcer la transparence
- Éviter les efforts redondants

## Pour la société

- Développement de la science citoyenne et des pratiques amateurs
- Stimuler l'innovation économique



# La charte française de déontologie des métiers de la recherche

## 1. Respect des dispositifs législatifs et réglementaires

*« Tout chercheur se tient informé des dispositifs législatifs et réglementaires qui régissent les activités professionnelles et veille au respect des textes correspondants, s'agissant notamment des recherches sur l'être humain, l'animal et l'environnement. »*

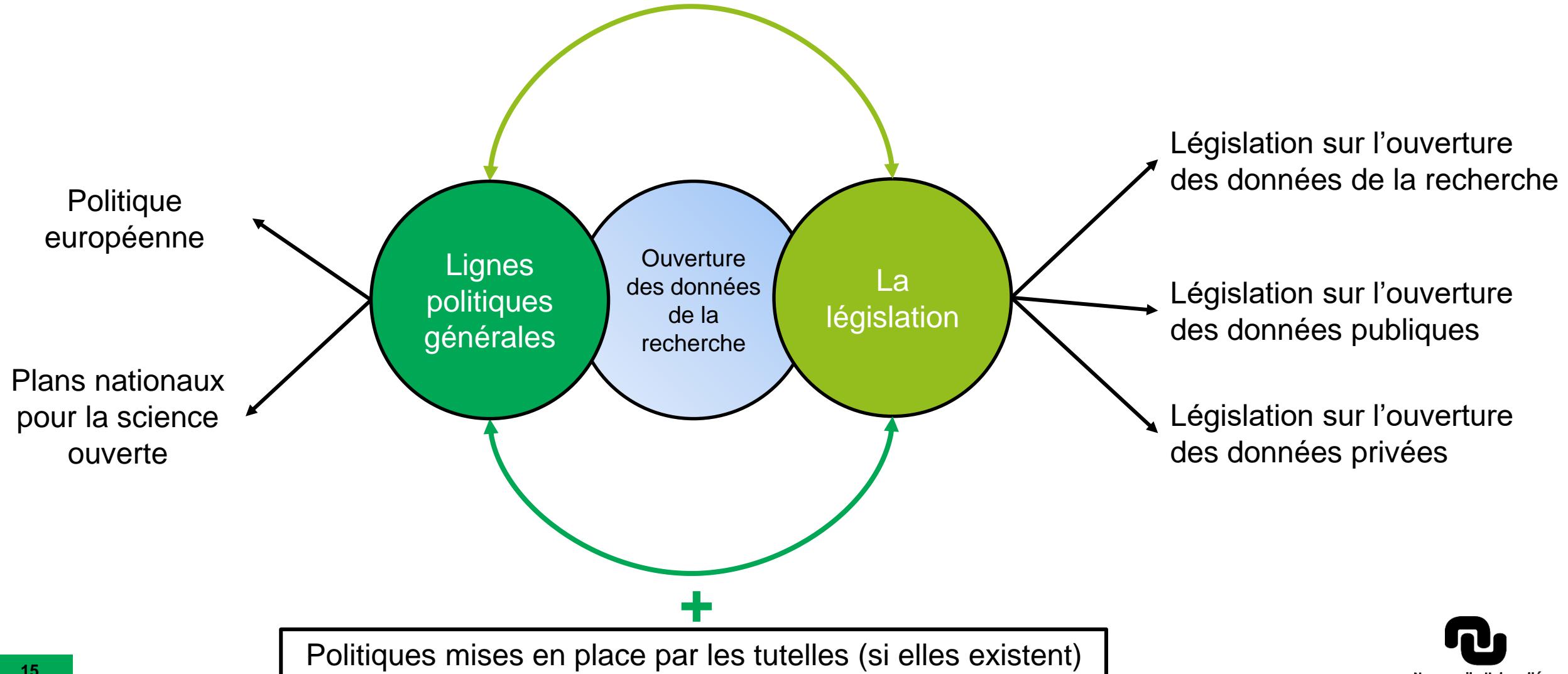
## 2. Fiabilité du travail de recherche

*« Les chercheurs doivent respecter les engagements pris dans le cadre de leur unité de recherche ou dans le cadre de contrats spécifiques. Les méthodes mises en œuvre pour la réalisation du projet de recherche doivent être les plus appropriées. La description détaillée du protocole de recherche, dans le cadre de cahiers de laboratoire ou de tout autre support, doit permettre la reproductibilité des travaux expérimentaux. Tous les résultats bruts (qui appartiennent à l'institution) ainsi que l'analyse des résultats doivent être conservés de façon à permettre leur vérification. [...] »*



# Contexte juridique et politique de l'ouverture des données

# Des contextes politiques et juridiques imbriqués mais indépendants



# Le contexte politique européen

Des obligations liées aux financements...



Depuis 2014

Obligation pour les projets financés par Horizon 2020 de **mettre leurs données en libre accès**

...et une infrastructure dédiée à la science ouverte



Inscription dans la politique de l'UE sur les données et l'IA

Gestion par la Commission Européenne et l'association EOSC

Des services et des initiatives...

Un exemple de projet : Portail d'accès et de mise à disposition de données et de logiciels

Plateforme de services pour le stockage, le traitement et le partage des données durant un projet de recherche ([EOSC EU Node](#))

**Objectif principal :** « *Fournir aux chercheurs, aux innovateurs, aux entreprises et aux citoyens européens un accès continu et une réutilisation fiable des données de la recherche grâce à un environnement de données distribué fiable et ouvert et à des services interopérables connexes* ».

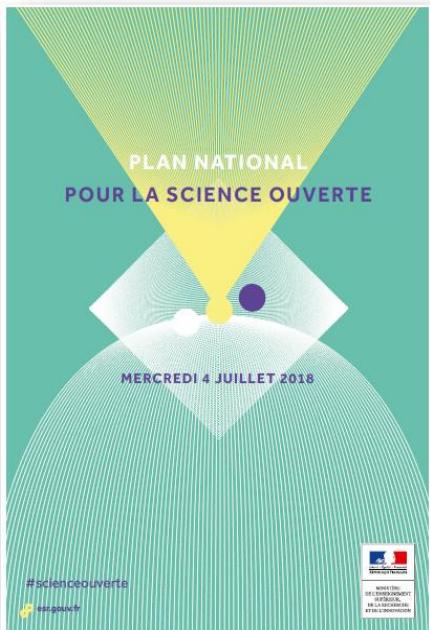


# La législation européenne

Directive (UE) 2019/1024 du Parlement européen et du Conseil du 20 juin 2019  
concernant les données ouvertes et la réutilisation des informations du secteur public  
(Art. 10 « Données de la recherche »)

- 1) Les États membres **encouragent la mise à disposition des données de la recherche en adoptant les politiques et en prenant les mesures nécessaires à l'échelon national afin de rendre librement accessibles les données résultant de la recherche financée au moyen de fonds publics** («politiques de libre accès») qui respectent le principe d'ouverture par défaut et sont compatibles avec les principes FAIR. Dans ce contexte, il y a lieu de tenir compte des préoccupations liées aux droits de propriété intellectuelle, à la protection des données à caractère personnel et à la confidentialité, à la sécurité et aux intérêts commerciaux légitimes dans le respect du principe «aussi ouvert que possible, mais aussi fermé que nécessaire». Ces politiques de libre accès visent les organismes exerçant une activité de recherche et les organisations finançant une activité de recherche.
- 2) Sans préjudice de l'article 1er, paragraphe 2, point c), les données de la recherche sont **réutilisables à des fins commerciales ou non commerciales**, conformément aux chapitres III et IV, dans la mesure où **elles sont financées au moyen de fonds publics** et où des chercheurs, des organismes exerçant une activité de recherche ou des organisations finançant une activité de recherche les ont déjà rendues publiques par l'intermédiaire d'une archive ouverte institutionnelle ou thématique. À cette fin, il est tenu compte des intérêts commerciaux légitimes, des activités de transmission des connaissances et des droits de propriété intellectuelle préexistants.

# Définition de la ligne politique française



2018

1. Généraliser l'accès ouvert aux publications
2. Structurer et ouvrir les données de la recherche
3. S'inscrire dans une dynamique durable, européenne et internationale



2021

1. Généraliser l'accès ouvert aux publications
2. Structurer, partager et ouvrir les données de la recherche
3. Ouvrir et promouvoir les codes sources produits par la recherche
4. Transformer les pratiques pour faire de la science ouverte le principe par défaut

# La législation française

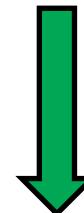
Loi de programmation de la recherche (2013)



Art. L. 112-1 du Code de la Recherche

« La recherche publique a pour objectifs :  
c) Le partage et la diffusion des connaissances scientifiques en donnant priorité aux formats libres d'accès  
[...]  
e) L'organisation de l'accès libre aux données scientifiques

Loi pour une République numérique (2016)



Art. L. 533-4 du Code de la Recherche

« Dès lors que les données issues d'une activité de recherche financée au moins pour moitié par des dotations de l'Etat, des collectivités territoriales, des établissements publics, des subventions d'agences de financement nationales ou par des fonds de l'Union européenne ne sont pas protégées par un droit spécifique ou une réglementation particulière et qu'elles ont été rendues publiques par le chercheur, l'établissement ou l'organisme de recherche, **leur réutilisation est libre** »

La loi de programmation pluriannuel de la recherche (2020) fait de la science ouverte un axe central de la politique de recherche française

# La législation française



Le site [JurisDoR](#) diffuse au public et auprès des communautés scientifiques de l'enseignement supérieur et de la recherche les **ressources juridiques** nécessaires à la gestion des données de la recherche. Ce site est produit par l'Université de Montpellier avec le soutien du Fonds national pour la science ouverte.

Le site n'a pas vocation à proposer du conseil juridique, ni à se substituer à la politique de gouvernance des données de la recherche par les établissements publics d'enseignement supérieur et de recherche.



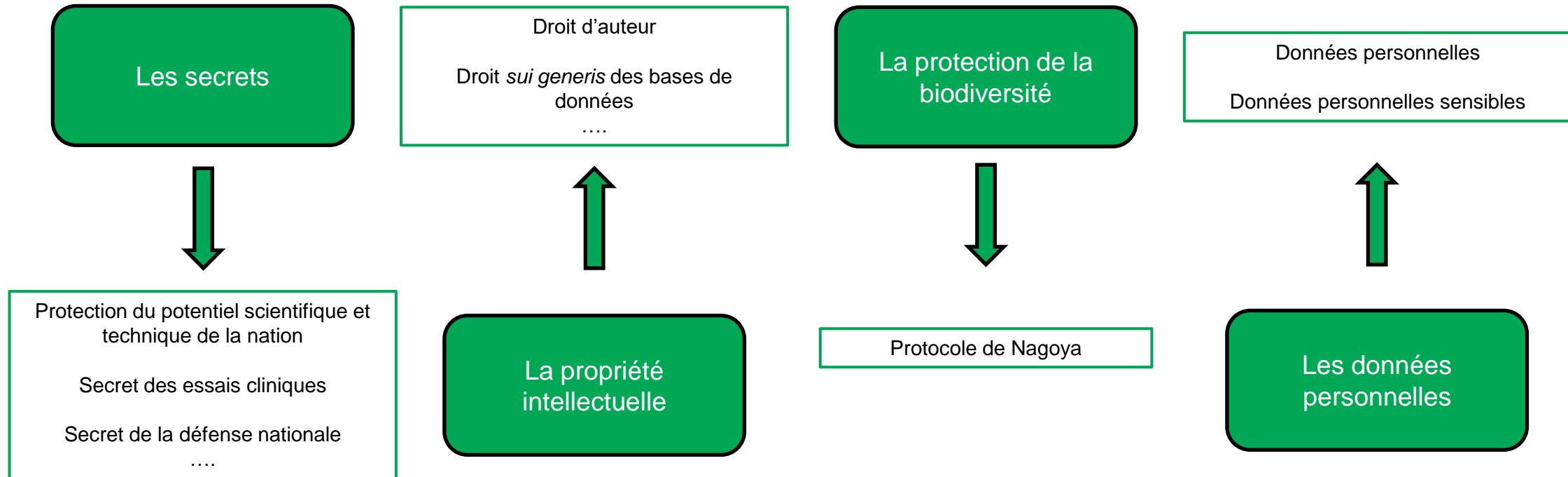
## Principe général de l'ouverture des données

Des lignes politiques et des textes de loi découlent un principe général :

**L'accès aux données doit être aussi ouvert que possible,  
pas plus fermé que nécessaire**

# Les exceptions au principe d'ouverture par défaut

**Principe général pour les exceptions :**  
**Tout ce qui préempte les données en raison d'intérêts légitimement protégés**

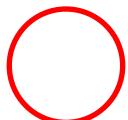




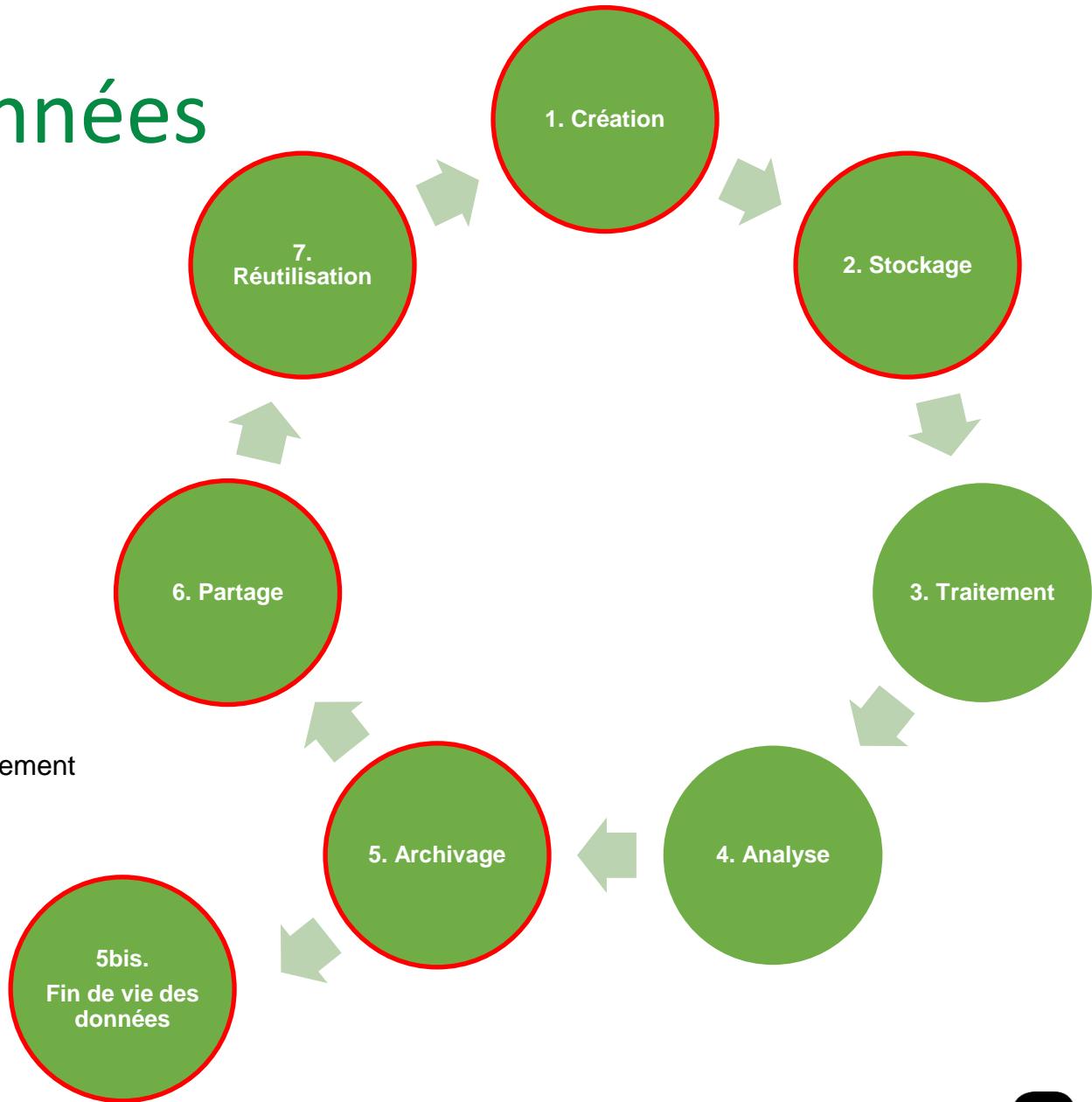
# Cycle de vie et principes d'ouverture

# Le cycle de vie des données

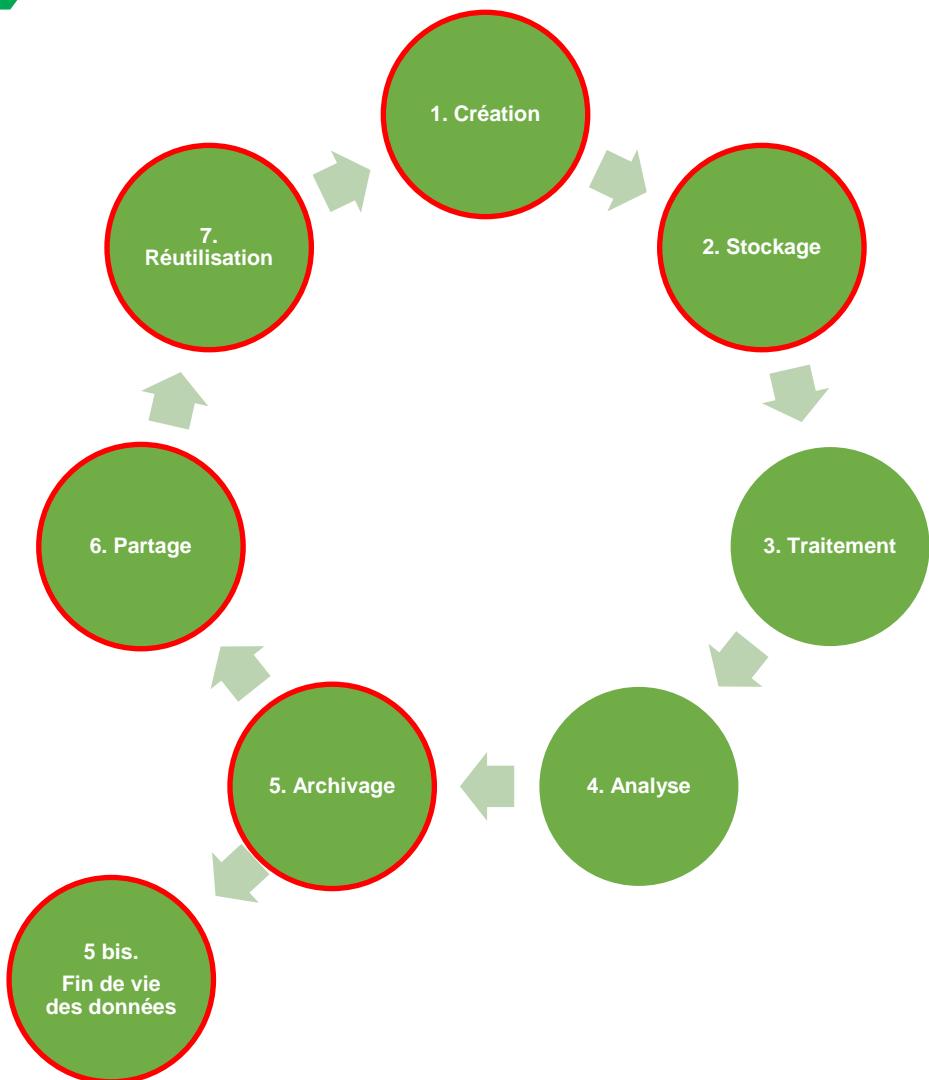
Le **cycle de vie des données** est une représentation du **processus** d'utilisation des données de leur **création**, jusqu'à leur **publication** et éventuelle **réutilisation**.



Étapes du cycle de vie des données pour lesquelles un accompagnement de l'Atelier de la Donnée en Normandie est possible



# Le cycle de vie des données



- 1 • Créer, collecter ou acquérir des données  
• Récupérer des métadonnées associées
- 2 • Définir un espace de stockage
- 3 • Vérifier, organiser, structurer, nettoyer les données  
• Décrire et documenter
- 4 • Analyser et interpréter  
• Produire des résultats de recherche  
• Citer les sources des données
- 5 • Préparer les données à la conservation  
• Migrer des données vers un format et un support adaptés  
• Sauvegarder et décrire les données dans un système d'archivage électronique (SAE)  
• Rendre accessibles et promouvoir les données archivées  
• Supprimer les données non nécessaires
- 6 • Définir des droits d'auteur et d'utilisation  
• Contrôler les métadonnées et la documentation  
• Effectuer des copies de sauvegarde  
• Publier et partager des données
- 7 • Réaliser des analyses secondaires, des suivis de recherche, de nouvelles recherches, des évaluations de la recherche, etc.



# Les principes FAIR





# Les principes FAIR

F

A

I

R



# Les principes FAIR

## Facile à trouver

Données facilement **repérable** par des humains et des machines *via* des **métadonnées** et des **identifiants** uniques et persistants. Objectif : permettre aux chercheurs, institutions ou entreprises de découvrir rapidement des données pertinentes pour leurs besoins

A

I

R



# Les principes FAIR

## Facile à trouver

Données facilement **repérable** par des humains et des machines *via* des **métadonnées** et des **identifiants** uniques et persistants. Objectif : permettre aux chercheurs, institutions ou entreprises de découvrir rapidement des données pertinentes pour leurs besoins

## Accessible

Données accessibles via des **protocoles standards et ouverts**, même si certaines restrictions d'accès (authentification, permissions) s'appliquent. Objectif : garantir que les données restent disponibles et compréhensibles, tout en respectant les exigences légales et éthiques

I

R



# Les principes FAIR

## Facile à trouver

Données facilement **repérable** par des humains et des machines *via* des **métadonnées** et des **identifiants** uniques et persistants. Objectif : permettre aux chercheurs, institutions ou entreprises de découvrir rapidement des données pertinentes pour leurs besoins

## Accessible

Données accessibles via des **protocoles standards et ouverts**, même si certaines restrictions d'accès (authentification, permissions) s'appliquent. Objectif : garantir que les données restent disponibles et compréhensibles, tout en respectant les exigences légales et éthiques

## Interopérable

Données structurées et formatées de manière **standardisée** pour pouvoir être combinées ou utilisées avec d'autres jeux de données ou outils. Objectif : favoriser la collaboration entre disciplines et permettre l'utilisation des données dans différents environnements ou systèmes technologiques

R



# Les principes FAIR

## Facile à trouver

Données facilement **repérable** par des humains et des machines *via* des **métadonnées** et des **identifiants** uniques et persistants. Objectif : permettre aux chercheurs, institutions ou entreprises de découvrir rapidement des données pertinentes pour leurs besoins

## Accessible

Données accessibles via des **protocoles standards et ouverts**, même si certaines restrictions d'accès (authentification, permissions) s'appliquent. Objectif : garantir que les données restent disponibles et compréhensibles, tout en respectant les exigences légales et éthiques

## Interopérable

Données structurées et formatées de manière **standardisée** pour pouvoir être combinées ou utilisées avec d'autres jeux de données ou outils. Objectif : favoriser la collaboration entre disciplines et permettre l'utilisation des données dans différents environnements ou systèmes technologiques

## Réutilisable

Données **bien documentées**, selon les normes disciplinaires, associées à des **licences** claires et conformes aux normes. Objectif : encourager l'innovation et réduire la duplication des efforts en tirant pleinement parti des données existantes

# Les principes FAIR en pratique : FAIR ou pas FAIR ?

Est-ce que la citation de ce jeu de données est FAIR ?

Scherber, Christoph; Andert, Hagen; Niedringhaus, Rolf; Tscharntke, Teja (2018). « Data from: A barrier island perspective on species-area-relationships [Dataset] ». *Dryad*

Non, car il manque le lien avec l'identifiant pérenne (Findable/Accessible)

Scherber, Christoph; Andert, Hagen; Niedringhaus, Rolf; Tscharntke, Teja (2018). « Data from: A barrier island perspective on species-area-relationships [Dataset] ». *Dryad*.  
<https://doi.org/10.5061/dryad.j5j730p>

Réponses :



# Les principes FAIR en pratique : FAIR ou pas FAIR ?

Personnalité, position sociale et pratiques culturelles : pour une articulation de la recherche psychologique et sociologique (ELIPSS 2019)

Version 3.3

BRISSON, Romain; BIANCHI, Renzo; ROLLAND, Jean-Pierre; Sciences Po, Centre de données socio-politiques (CDSP), CNRS, 2022, "Personnalité, position sociale et pratiques culturelles : pour une articulation de la recherche psychologique et sociologique (ELIPSS 2019)", <https://doi.org/10.21410/7E4/REVLRD>, data.sciencespo, V3, UNF:6:DRYy1MU6UGLfpluvWLCKA== [fileUNF]

**Citer l'ensemble de données** Pour en apprendre davantage sur le sujet, consulter le document [Data Citation Standards \[en\]](#).

Description

L'étude "Personnalité, position sociale et pratiques culturelles : pour une articulation de la recherche psychologique et sociologique" (PSYSOC) se fixe comme objectif de lier les champs de la sociologie et de la psychologie de la personnalité pour investiguer les relations entre personnalité, position sociale et pratiques culturelles. D'autre part, elle cherche à caractériser la distribution sociale des profils de personnalité dans la population française. Cette étude utilise le questionnaire NEO-Five Factor Inventory pour évaluer les traits de personnalité, ainsi que le Patient Health Questionnaire-9 pour contrôler l'influence des symptômes dépressifs sur la mesure du trait de personnalité de névrosisme. Elle s'appuie aussi sur des questions relatives aux pratiques culturelles déjà utilisées dans le cadre du programme ELIPSS, auxquelles elle ajoute de nouvelles questions portant sur le rapport à deux formes d'art : la musique et la photographie. Elle s'est déroulée du 23/05/2019 au 01/07/2019, pendant la phase élargie du panel ELIPSS.

Sujet

Sciences sociales

Mot-clé

PERSONNALITE, PSYCHOLOGIE, SOCIOLOGIE, CULTURE, MUSIQUE, PHOTOGRAPHIE, STATUT SOCIAL

Modalités d'accès à l'ensemble de données

Contacter le propriétaire Partager

Statistiques d'utilisation sur l'ensemble de données

16 téléchargements

Cette description de métadonnées est-elle FAIR ?

# Les principes FAIR en pratique : FAIR ou pas FAIR ?

Personnalité, position sociale et pratiques culturelles : pour une articulation de la recherche psychologique et sociologique (ELIPSS 2019)

Version 3.3

The screenshot shows a data portal interface for the ELIPSS 2019 dataset. At the top left is a green arrow icon. The main title is "Personnalité, position sociale et pratiques culturelles : pour une articulation de la recherche psychologique et sociologique (ELIPSS 2019)". Below it is a sub-section title "(ELIPSS 2019)". A small thumbnail image of the dataset is on the left. The dataset details include: "BRISSON, Romain; BIANCHI, Renzo; ROLLAND, Jean-Pierre; Sciences Po, Centre de données socio-politiques (CDSP), CNRS, 2022, "Personnalité, position sociale et pratiques culturelles : pour une articulation de la recherche psychologique et sociologique (ELIPSS 2019)", <https://doi.org/10.21410/7E4/REVLRD>, data.sciencespo, V3, UNF:6:DRYy1MU6UGLjflpuvWLCKA== [fileUNF]". Below this, there are two red links: "Citer l'ensemble de données" and "Pour en apprendre davantage sur le sujet, consulter le document Data Citation Standards [en]". On the right, there's a red button "Modalités d'accès à l'ensemble de données" with options "Contacter le propriétaire" and "Partager". Below that is a section "Statistiques d'utilisation sur l'ensemble de données" with a link "16 téléchargements". At the bottom, there are sections for "Description", "Sujet", and "Mot-clé", and a footer with a "Licence/Accord d'utilisation des données" section showing the CC BY-SA 4.0 license logo.

**Description** L'étude "Personnalité, position sociale et pratiques culturelles : pour une articulation de la recherche psychologique et sociologique" (PSYSOC) se fixe comme objectif de lier les champs de la sociologie et de la psychologie de la personnalité pour investiguer les relations entre personnalité, position sociale et pratiques culturelles. D'autre part, elle cherche à caractériser la distribution sociale des profils de personnalité dans la population française. Cette étude utilise le questionnaire NEO-Five Factor Inventory pour évaluer les traits de personnalité, ainsi que le Patient Health Questionnaire-9 pour contrôler l'influence des symptômes dépressifs sur la mesure du trait de personnalité de névrosisme. Elle s'appuie aussi sur des questions relatives aux pratiques culturelles déjà utilisées dans le cadre du programme ELIPSS, auxquelles elle ajoute de nouvelles questions portant sur le rapport à deux formes d'art : la musique et la photographie. Elle s'est déroulée du 23/05/2019 au 01/07/2019, pendant la phase élargie du panel ELIPSS.

**Sujet** Sciences sociales

**Mot-clé** PERSONNALITE, PSYCHOLOGIE, SOCIOLOGIE, CULTURE, MUSIQUE, PHOTOGRAPHIE, STATUT SOCIAL

**Licence/Accord d'utilisation des données** CC BY-SA 4.0

**Non, car il manque la licence (R = Réutilisable)**



Gérer ses données au quotidien

---

# Documenter ses données



# Comment nommer ses fichiers ?

- Ne pas utiliser un nom trop long (max 25 caractères)
- Inclure des caractères descriptifs pour faciliter l'identification du fichier
- Si le nom du fichier inclut des dates, elles doivent être formatées de manière consistante
- Éviter les caractères spéciaux (& \* % \$ £ ° { ! @ qui peuvent être réservés pour des tâches spécifiques selon les systèmes d'exploitation ainsi que les caractères accentués
- Utiliser autant que possible des extensions (.xls, .ssd, .txt, etc.) pour refléter de manière pertinente l'environnement logiciel dans lequel le fichier a été créé
- Préférer le caractère souligné ou le tiret au lieu du point ou de l'espace, ces deux derniers pouvant être interprétés différemment selon les systèmes d'exploitation
- Ne pas hésitez à utiliser un logiciel de renommage
- Le numéro de la version du fichier ou du dossier

# Les métadonnées

Les boites de conserves = jeux de données



Avec  
métadonnées

VS

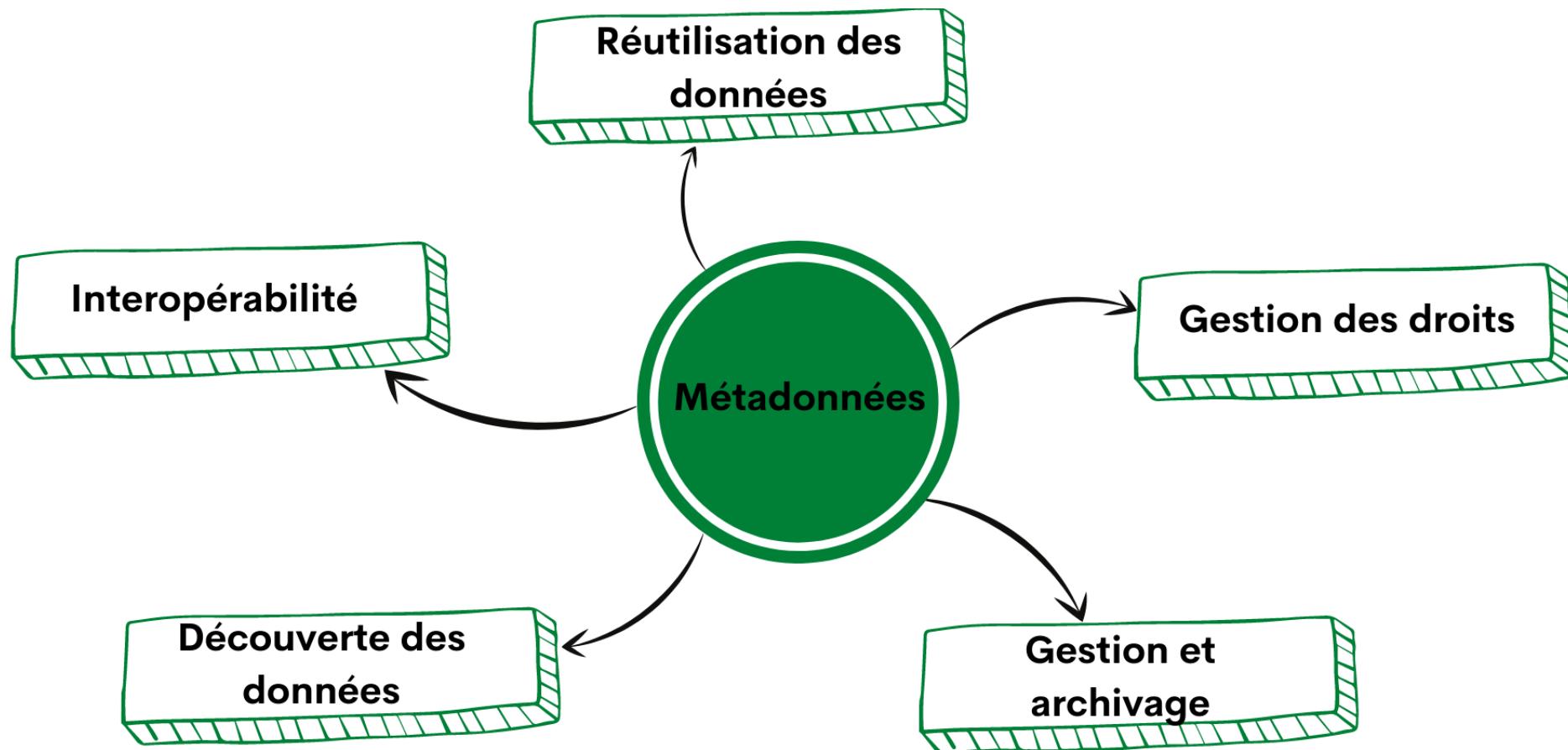


Sans  
métadonnées

Les **métadonnées** (informations sur les étiquettes) décrivent le **contenu du jeu de données** et permettent son **utilisation** ou sa **réutilisation**.

# Les métadonnées

Les métadonnées permettent de faciliter un certain nombre d'actions :





# Les standards de métadonnées

Les standards sont des **modèles** qui préconisent une **manière uniforme de décrire** les données.  
Cette description est effectuée en fonction de différents critères :

Le type de  
ressources  
à décrire

Le champ  
disciplinaire

L'entrepôt  
choisi\*

\* Par exemple Recherche Data Gouv (cf. le Guide de saisie des métadonnées générales)



# Les standards de métadonnées

Le standard de métadonnées doit être déterminé au début du projet de recherche afin de préparer le partage des données et faciliter leur interopérabilité.

Standard interdisciplinaire

*Dublin Core*

*DataCite Metadata*

Besoin de trouver un standard ?

Des répertoires existent :

- [Le répertoire du Digital Curation Centre](#)
- [Le Metadata Standards Catalog](#)
  - [FAIRsharing](#)

Standard disciplinaire

*DDI (sciences sociales)*

*EML (écologie, sciences de la terre,  
sciences environnementales)*

*AVM (astronomie)*

*CIF (chimie)*

*MIBBI (biologie)*



# Les standards de métadonnées : exemples

SciencesPo

Recherche ▾ Guide d'utilisation Support Français ▾ Se connecter

data.sciencespo > Banque de données du CDSP >

## Baromètre Racisme 2018

Version 3.0

Mayer, Nonna; Michelat, Guy; Tiberj, Vincent; Vitale, Tommaso, 2022, "Baromètre Racisme 2018", [http://doi.org/10.21410/7E4/LWW3UA](https://doi.org/10.21410/7E4/LWW3UA), data.sciencespo, V3

Citer l'ensemble de données Pour en apprendre davantage sur le sujet, consulter le document [Data Citation Standards \[en\]](#).

Description	Le baromètre racisme est un sondage annuel réalisé depuis 1990, offrant un véritable baromètre sur les opinions à l'égard du racisme et des discriminations en France. Réalisée en face à face du 6 au 14 novembre 2018, l'édition 2018 du baromètre racisme contient un échantillon de 1 007 personnes, représentatif de la population métropolitaine, âgée de 18 ans et plus, constitué d'après la méthode des quotas (sexe, âge, profession du chef de ménage, après stratification par région et catégorie d'agglomération). Comme tous les ans depuis 2000, une analyse détaillée du sondage a été réalisée par une équipe de chercheurs de Sciences Po et du CNRS et, et présentée dans le rapport 2018 de la CNCNDH sur la lutte contre le racisme, l'antisémitisme et la xénophobie.
Sujet	Sciences sociales
Mot-clé	RACISME, DISCRIMINATIONS
Publication liée	Nonna Mayer, Guy Michelat, Vincent Tiberj, Tommaso Vitale, « Évolution et structures des préjugés : le regard des chercheurs », in CNCNDH, La lutte contre le racisme, l'antisémitisme et la xénophobie. Année 2018 (Paris: La Documentation française, 2019), p. 75-161.
Licence/Accord d'utilisation des données	 CC BY-SA 4.0

Un jeu de données : <https://doi.org/10.21410/7E4/LWW3UA>

Sa description en DDI-C : <https://data.sciencespo.fr/api/datasets/export?exporter=html&persistentId=doi%3A10.21410/7E4/LWW3UA>

Title	Baromètre racisme 2018
Creator	Mayer, Nonna ; Michelat, Guy ; Tiberj Vincent ; Vitale Tommaso
Subject	Social Sciences ; racisme ; discriminations
Description	Le baromètre racisme est un sondage annuel réalisé depuis 1990, offrant un véritable baromètre sur les opinions à l'égard du racisme et des discriminations en France. Réalisée en face à face du 6 au 14 novembre 2018, l'édition 2018 du baromètre racisme contient un échantillon de 1 007 personnes, représentatif de la population métropolitaine, âgée de 18 ans et plus, constitué d'après la méthode des quotas (sexe, âge, profession du chef de ménage, après stratification par région et catégorie d'agglomération). Comme tous les ans depuis 2000, une analyse détaillée du sondage a été réalisée par une équipe de chercheurs de Sciences Po et du CNRS et, et présentée dans le rapport 2018 de la CNCNDH sur la lutte contre le racisme, l'antisémitisme et la xénophobie.
Publisher	IPSOS
Contributor	IPSOS ; CNCNDH ; SIG
Date	2022-06-29
Type	Nombre
Format	
Identifier	<a href="https://doi.org/10.21410/7E4/LWW3UA">doi:10.21410/7E4/LWW3UA</a>
Source	
Language	fr.
Relation	Nonna Mayer, Guy Michelat, Vincent Tiberj, Tommaso Vitale, « Évolution et structures des préjugés : le regard des chercheurs », in CNCNDH, La lutte contre le racisme, l'antisémitisme et la xénophobie. Année 2018 (Paris: La Documentation française, 2019), p. 75-161.
Coverage	France
Rights	CC-BY-SA 4.0

Sa description en Dublin Core

# Les standards de métadonnées

Mixed layer depth climatology computed with a density threshold criterion of 0.03kg/m<sup>3</sup> from 10 m depth value

DATE 2022-11-23

AUTHORS de Boyer montégut Clément <sup>1</sup>

AFFILIATIONS 1. Ifremer, LOPS (University of Brest – Ifremer – Cnrs – Ird, UMR6523), Plouzané, France

DOI 10.17882/91774

PUBLISHER SEANOE

The dataset made available here is the monthly climatology (i.e. 12 months) of ocean surface Mixed Layer Depth (MLD) over the global ocean, at 1 degree x 1 degree spatial resolution. It is provided as a netccdf data file embedded in a tar file at bottom of this page. A short documentation about the product is also available as a pdf readme file. The climatology is based on about 4.5 million casts/profiles of temperature and salinity measurements made at sea between January 1970 and November 2021. Those profiles data come from the ARGO program and from the NCEI-NOAA World Ocean Database (WOD). The MLD is computed on each individual cast/profile using a threshold criterion. The depth of the mixed layer is defined as the shallowest depth where the surface potential density of the profile is superior to a reference value taken close to the surface added with the chosen threshold. Here we take a threshold value for the density of 0.03kg/m<sup>3</sup>, and a surface reference depth fixed at 10m (de Boyer Montégut et al. JGR 2004, see reference details at bottom of page). This mixed layer is by definition homogeneous in density (up to 0.03 kg/m<sup>3</sup> variations) and can also be called an isopycnal layer.

DISCIPLINES Physical oceanography

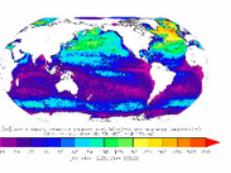
KEYWORDS ocean surface Mixed Layer Depth, MLD, global ocean

LOCATION 90N, -90S, -180E, 180W

LICENCE 

ACKNOWLEDGEMENTS The production of this MLD climatology relies on individual profile measurements of pressure/depth, temperature and salinity made at sea. Those data come from several sources. The Argo data were collected and made freely available by the International Argo Program and the national programs that contribute to it. (<https://argo.ucsd.edu>, <https://www.ocean-ops.org>). The Argo Program is part of the Global Ocean Observing System. All other kind of casts/profiles comes from the World Ocean Database (WOD, Boyer et al. 2018). First released in 1994, and updated regularly since then, the World Ocean Database represents the world's largest collection of vertical profile data of ocean characteristics available internationally without restriction.

 DOWNLOAD DATA

  
MLD climatology for the month of January



[Download data](#)

TXT RIS XLS RTF BIBTEX

Exemple d'un schéma de métadonnées développé par l'IFREMER pour Archimer et utilisé par Seanoe

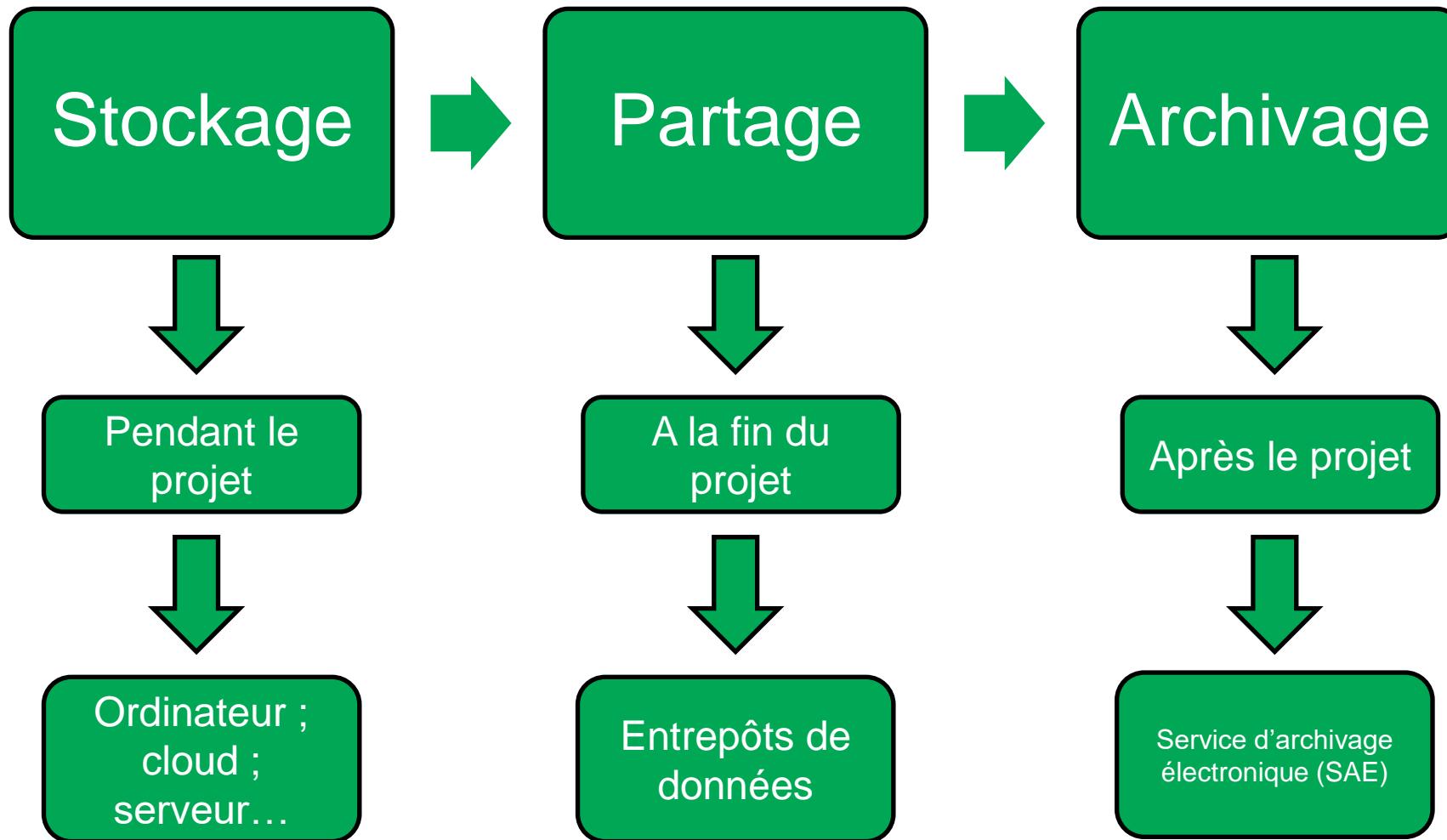
Lien vers le jeu de données décrit :  
<https://doi.org/10.17882/91774>

---

# Stockage, partage et archivage

---

# Stockage, partage et archivage : 3 étapes ; 3 objectifs ; 3 outils...





# Sauvegarde et stockage sécurisé pendant le projet

## Les questions à se poser :

- Comment allez-vous sauvegarder vos données ?
- À quelle régularité ferez-vous des sauvegardes ?
- Sauvegarderez-vous toutes vos données (sauvegarde complète) ou uniquement celles qui ont été ajoutées / modifiées depuis la dernière sauvegarde (sauvegarde incrémentielle) ?
- Combien de temps garderez-vous vos sauvegardes ?
- Combien d'espace-disque sera nécessaire pour ces sauvegardes ?
- Comment garder trace des différentes sauvegardes effectuées, notamment si elles sont faites sur des supports différents ?
- Comment identifier les sauvegardes les plus récentes / complètes ?

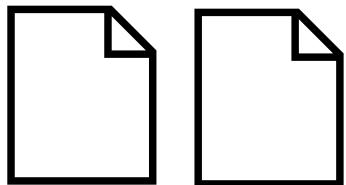
# Sauvegarde – la règle à suivre

Après avoir perdu sa thèse dans un TGV,  
elle lance un appel pour la retrouver

Par Figaro Etudiant

Publié le 04/04/2019 à 16:46, mis à jour le 04/04/2019 à 16:46

De l'importance de suivre la règle suivante...



Deux sauvegardes



Deux supports  
différents



Un autre support  
(si possible)

*En multipliant les copies de vos données sur des supports différents, vous êtes sûr d'avoir toujours une sauvegarde disponible en cas de problème sur l'un des supports. De plus, la sécurité des données peut être renforcée grâce :*

- À un mot de passe
- À un chiffrement
- À l'utilisation d'un antivirus
- L'utilisation d'un support de stockage (ordinateur) qui ne soit relié à aucun réseau

Source : [https://doranum.fr/wp-content/uploads/stockage\\_sauvegarde\\_donnees.pdf](https://doranum.fr/wp-content/uploads/stockage_sauvegarde_donnees.pdf)

Les supports de stockage utilisés doivent être gérés par votre université ou la tutelle de votre laboratoire (INSERM ; CNRS...)



# Quel(s) support(s) de stockage utilisez-vous ?

A) Ordinateur

B) Clé USB

C) Disque dur externe

D) Cloud institutionnel

E) Cloud privé

Autre

Réponses : 

# Stockage, partage et sécurité au cours du projet

Support de stockage	Sécurité	Accès	Coût	Remarque d'utilisation
Ordinateur professionnel	★★★★☆ Sujet au piratage informatique, aux détériorations et pannes	★★★★☆ Pas adapté au partage, nécessite l'utilisation d'un support externe ou d'Internet (mail, cloud...)	★★★★☆ Pas de coût supplémentaire ou coût peu important	- Pour un stockage temporaire - Nécessité de crypter les données confidentielles et sensibles
Support externe	★★★★☆ - Sujet au vol, à la perte du support - Durée de vie limitée (dégradation du matériel)	★★★★☆ Facilement transportable, il permet de transférer les données vers un autre ordinateur	★★★★☆ Pas de coût supplémentaire ou coût peu important	- Pour un stockage temporaire - Nécessité de crypter ou de sécuriser physiquement les données confidentielles et sensibles
Serveur institutionnel	★★★★☆ Stockage fiable, durable et sécurisé (contre le vol, le piratage, les incendies...)	★★★★☆ La connexion au serveur institutionnel ne facilite pas le travail avec des personnes extérieures	★★★★☆ Coût assez important mais pas forcément répercuté sur l'usager	- Pour un stockage plus pérenne - Adapté pour le stockage de données sensibles et des versions « stables » de vos données - Toutes les institutions ne proposent pas ce service
Serveur Cloud	★★★★☆ On ne sait pas vraiment où sont stockées les données, ni ce qu'elles deviennent	★★★★☆ Permet un travail synchronisé avec toutes les personnes ayant été autorisées au partage	★★★★☆ Payant à partir d'une certaine limite de stockage	- Pour un partage avec des personnes externes à l'institution - Ne pas y mettre de données sensibles ou confidentielles - Pas de contrôle sur la procédure de sauvegarde des données



Déconseillé



Déconseillé



# Les solutions normandes de stockage en ligne

Institution	Nom	Espace disponible	Lien vers l'outil
Normandie Université	NC Normandie Université	50 MO	<a href="https://nc.normandie-univ.fr/">https://nc.normandie-univ.fr/</a>
Université de Caen Normandie	<ul style="list-style-type: none"><li>• Unicloud</li><li>• Disque Z</li></ul>	100 GO Disque Z de 30 GO	<a href="https://unicloud.unicaen.fr/">https://unicloud.unicaen.fr/</a> <a href="https://gedemande.unicaen.fr/">https://gedemande.unicaen.fr/</a>
Université Le Havre Normandie	<ul style="list-style-type: none"><li>• Disque Z</li></ul>	20 GO (pour les personnels)	
Université de Rouen Normandie	<ul style="list-style-type: none"><li>• Drive</li><li>• Disque Z</li></ul>	20 GO (pour les personnels) / 1 GO (pour tous les autres) Disque Z de 100 GO	<a href="https://drive.univ-rouen.fr">https://drive.univ-rouen.fr</a>



# Les entrepôts de données

Qu'est-ce qu'un entrepôt de données ?

Un entrepôt de données est « une infrastructure de stockage et de services facilitant le **dépôt**, la **description**, le **partage en accès ouvert**, la **découverte** et la **réutilisation**, par des humains ou des machines, de **jeux de données**\* [...]. Ces jeux de données sont associés à des métadonnées et sont conservés à moyen ou long terme [...] Un entrepôt se distingue d'un catalogue, par sa capacité à assurer l'**hébergement**, la **gestion** et la **curation** des données et pas uniquement le système d'information (catalogage et exposition des métadonnées moissonnées à partir d'autres structures).»

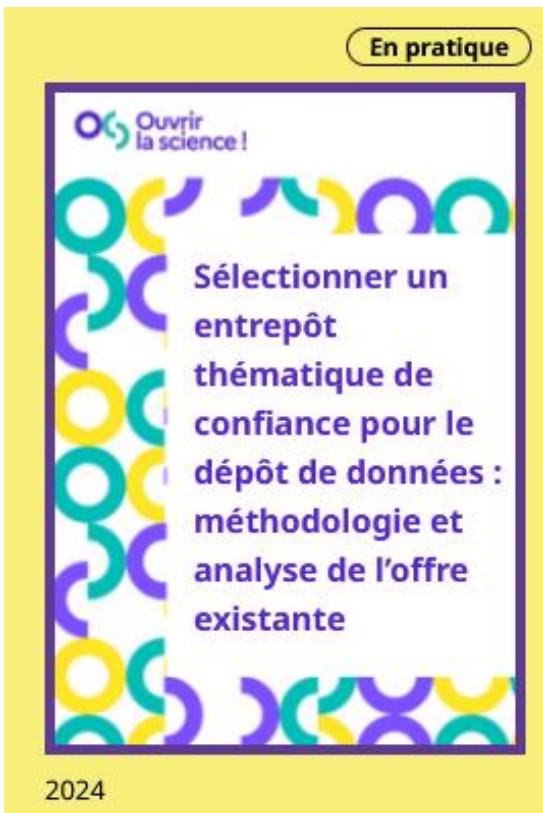
([Comité pour la science ouverte, 2024](#))

\* Un jeu de données (ou *dataset*) est une « agrégation (...) de données brutes ou dérivées présentant une certaine « unité », rassemblées pour former un ensemble cohérent ».

[Rémi Gaillard, De l'Open Data à l'Open Research Data \(2014\)](#)

# Les entrepôts de données

Comment choisir son entrepôt ?



Liste obtenue à partir de critères d'exclusion

- Absence de **modération** des dépôts
- Absence d'**identifiant pérenne**
- Absence de garantie sur la **pérennité de l'infrastructure**
- Entrepôts pratiquant la **cession de droits**
- **Politique tarifaire** excessive
- Localisation des données **hors Union européenne** pour certains types de données
- **Dépôt restreint** par l'affiliation institutionnelle



**Liste évolutive !**

# Les entrepôts de données

Quelques exemples d'entrepôts thématiques de confiance



PANGAEA.

Data Publisher for Earth & Environmental Science



Normandie Université

# Recherche Data Gouv



RÉPUBLIQUE  
FRANÇAISE

Liberté  
Égalité  
Fraternité

recherche.data.gouv.fr

Lors du lancement du deuxième plan national pour la science ouverte, il a été annoncé la mise en place d'un entrepôt de données au niveau national « [Recherche Data Gouv](#) ».

L'entrepôt a été inauguré le 8 juillet 2022.

« Recherche Data Gouv a été pensé pour soutenir les équipes de recherche dans leur travail de structuration des données pour les rendre Faciles à trouver, Accessibles, Interopérables, Réutilisables, c'est-à-dire conformes aux principes FAIR ».



<https://entrepot.recherche.data.gouv.fr/dataverse/normandie-univ>



<https://entrepot.recherche.data.gouv.fr/dataverse/unicaen>



<https://entrepot.recherche.data.gouv.fr/dataverse/univ-lehavre>



<https://entrepot.recherche.data.gouv.fr/dataverse/univ-rouen>

Devez-vous utiliser un entrepôt de données disciplinaire ou Recherche Data Gouv ?  
Pour savoir où publier vos données, [cliquez ici](#)



Normandie Université

# Les entrepôts de données

## Les entrepôts de données d'éditeurs commerciaux : quelle stratégie adopter ?



Elsevier Data Repository

IEEEDataPort™

« Les éditeurs d'écrits scientifiques ne peuvent valablement obtenir de cessions exclusives de droits sur des données de recherche liées à la publication, qu'elles soient déposées sur un entrepôt de données ou qu'elles figurent comme *supplementary materials* de l'article »

# Protéger ses données lors de leur ouverture



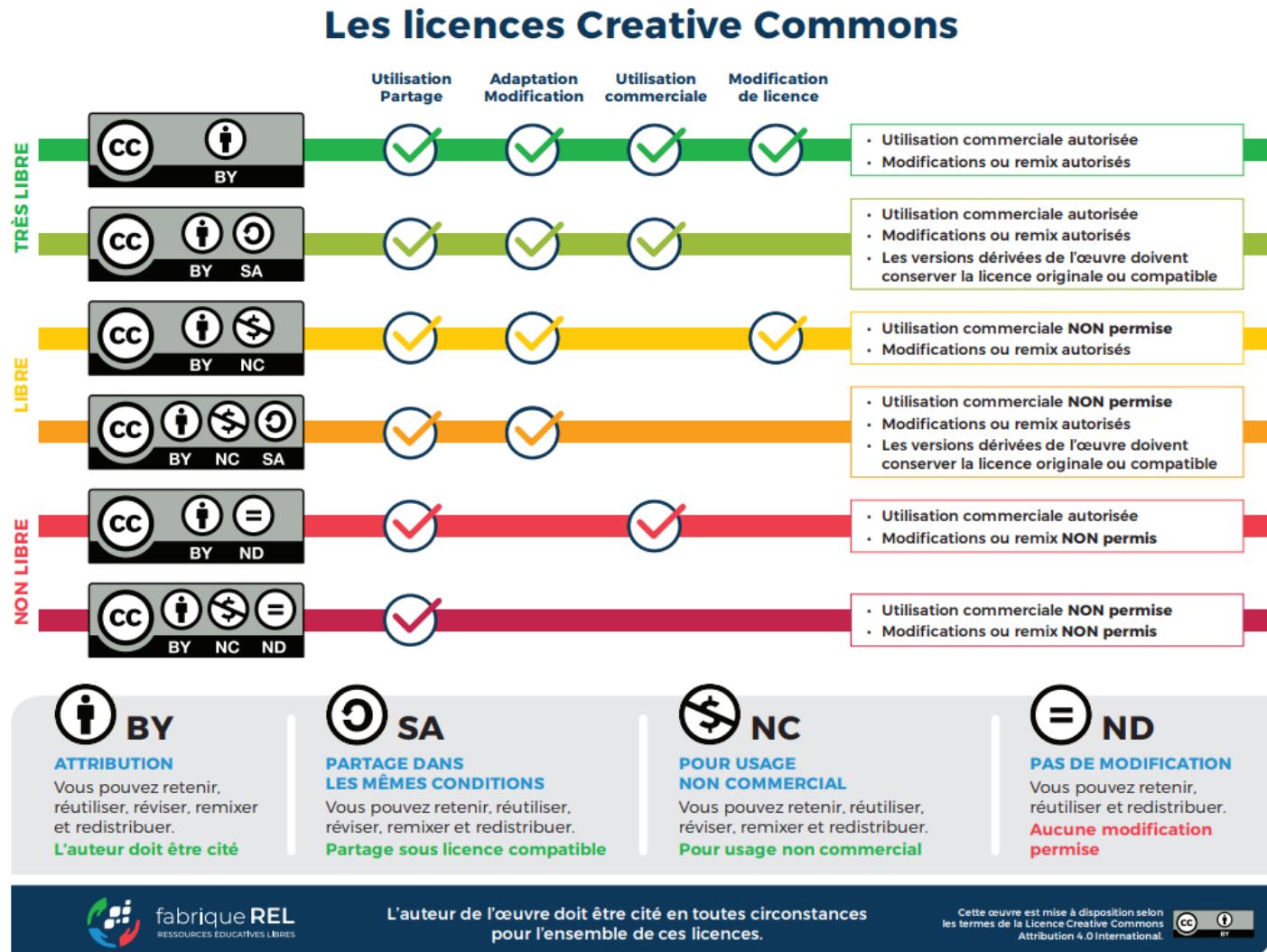
LICENCE OUVERTE  
OPEN LICENCE



En 2011, **Etalab** présente la « **Licence ouverte** » qui permet la diffusion des données publiques françaises. Elle autorise la réutilisation, la reproduction, la modification, la redistribution des données et leur exploitation à titre commercial sous réserve de mentionner a minima le nom du producteur et la date de dernière mise à jour.

Elle est compatible avec toute licence qui exige a minima la mention de paternité, notamment avec les licences CC-BY et ODC-BY.

# Protéger ses données lors de leur ouverture



## Les licences Creative Commons

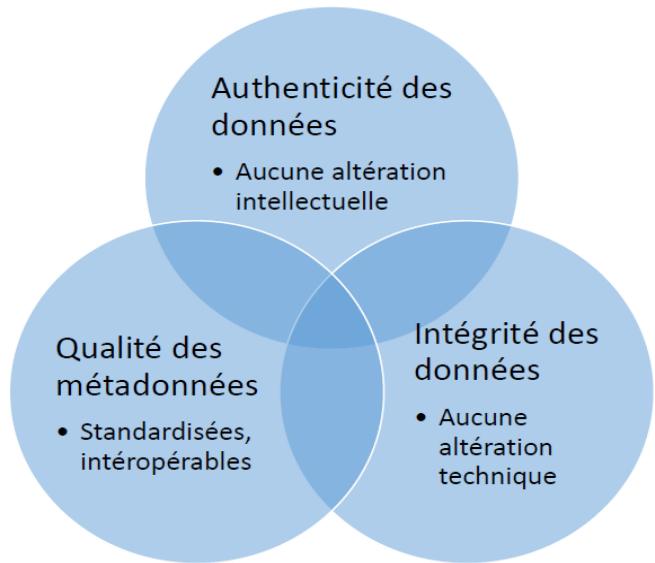
permettent la diffusion de contenus numériques (texte, image, œuvre audiovisuelle, etc.). Les auteurs des productions peuvent ainsi fixer les droits de réutilisation par 4 clauses :

- 1 clause permanente**
  - attribution (= Paternité)
- 3 clauses optionnelles**
  - pas d'utilisation commerciale
  - pas de modification (pas d'œuvres dérivées)
  - partage dans les mêmes conditions



# Archivage pérenne

## Ses buts



Les archivistes de vos universités sont vos premiers interlocuteurs pour vous accompagner dans ce processus

# L'impact écologique



Source : <https://opendatafrance.gitbook.io/greendata-pour-un-impact-maitrise-des-donnees/greendata/1.1-contexte>

---

# Expliquer et décrire la production et la gestion de ses données

---

# Introduction au plan de gestion de données (Data Management Plan)

## Définition

« Le Plan de Gestion de Données (PGD) ou Data Management Plan (DMP) est un **document synthétique qui aide à organiser et anticiper toutes les étapes du cycle de vie de la donnée**. Il explique pour chaque jeu de données comment seront gérées les données d'un projet, depuis leur création ou collecte jusqu'à leur partage et leur archivage »

Inist-CNRS. [“Plan de gestion de données : fiche synthétique”](#), 2024.

Un outil pour rédiger un PGD...



## Code la recherche

Depuis le 27 décembre 2023, le Code de la recherche précise que « *les établissements publics et fondations reconnues d'utilité publique [...] définissent les conditions de conservation, de communication et de réutilisation des résultats bruts des travaux scientifiques menés en leur sein* » (art. [D211-2](#)).

## Obligations contractuelles



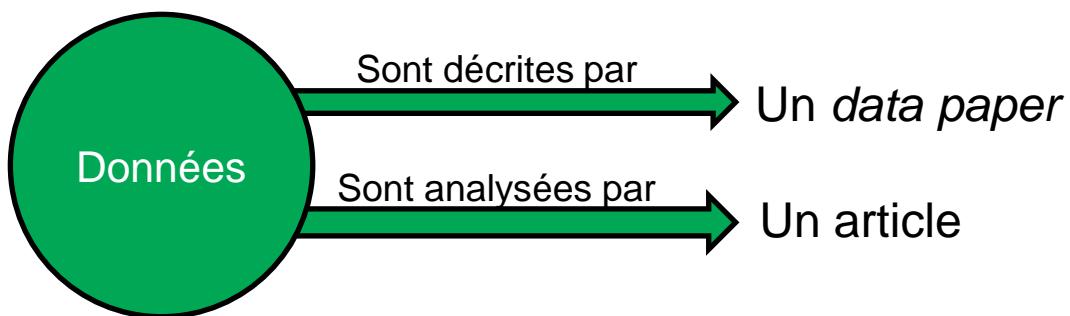
**2 versions du PGD** : T0+6 mois ;  
PGD final

**3 versions du PGD** : T0+6 mois ;  
PGD intermédiaire ; PGD final

# Les data papers

## Principes des *data papers*

- Décrit un jeu de données déposé, au préalable, dans un entrepôt : contexte de production ; méthode de production ; choix techniques...
- Il ne comprend pas d'analyses ni de conclusion (pas de concurrence avec l'article scientifique « classique »)





Rechercher  
des jeux de données

# Rechercher des jeux de données

Deux exemples d'outils pour rechercher des jeux de données dans son domaine de recherche...

Un annuaire d'entrepôts de données



Un moteur de recherche académique



# Rechercher des jeux de données

Un annuaire d'entrepôts de données



Catalogue porté par...



Permet de découvrir des **entrepôts de données** dans son domaine de recherche

3403 entrepôts de données  
(juillet 2025)

Recherche possible par :  
- Pays  
- Domaines  
+ de 30 filtres de recherche

<https://www.re3data.org/>

# Rechercher des jeux de données

Moteur de recherche porté par...



Un moteur de recherche académique



Permet de découvrir des **jeux de données** dans son domaine de recherche

## Recherche large :

11.850 sources

428.387.748 documents

44.062.532 jeux de données

(possibilité de filtrer pour ne rechercher que  
des données de recherche)

Sources continuellement  
enrichies

<https://www.base-search.net/>



Après la formation



# Nos autres formations

Données de la recherche

Rédiger un plan de gestion de données

Diffuser ses données grâce aux *data papers*

Open Access et identifiants chercheurs

Open access et dépôt dans  
HAL Normandie Université

Identifiants chercheurs.  
Panorama et focus ORCID

Ateliers thématiques HAL  
(questions juridiques ; IDHAL/CV ; compte/profil/dépôt)

---

Science ouverte dans les projets financés

Faire entrer la science ouverte dans son projet ANR

Projets financés – Respecter les obligations science ouverte



# Nous contacter



Normandie Université

<https://science-ouverte.normandie-univ.fr/>



UNIVERSITÉ  
CAEN  
NORMANDIE

[donnees-recherche@unicaen.fr](mailto:donnees-recherche@unicaen.fr)



[learningcenter@em-normandie.fr](mailto:learningcenter@em-normandie.fr)



[donnees-recherche@univ-lehavre.fr](mailto:donnees-recherche@univ-lehavre.fr)



[scd-chercheurs@univ-rouen.fr](mailto:scd-chercheurs@univ-rouen.fr)



[eugenie.gourichon@insa-rouen.fr](mailto:eugenie.gourichon@insa-rouen.fr)  
[christelle.vandepitte@insa-rouen.fr](mailto:christelle.vandepitte@insa-rouen.fr)



Normandie Université



**Merci de votre attention !**

# Ressources importantes



Géré par l'INIST-CNRS et le GIS « Réseau URFIST », [DoRANum](#) est une plateforme de formation en ligne sur la gestion et le partage des données de la recherche selon les principes FAIR. Depuis juillet 2022, elle s'inscrit dans l'écosystème national Recherche Data Gouv en tant que [Centre de ressource](#) supports pédagogiques, e-formation. Vous y retrouverez des ressources d'autoformation classées en 9 thématiques

## ENJEUX & BÉNÉFICES

Pourquoi partager les données ?  
Qu'est-ce que l'Open Science ?



## MÉTADONNÉES

Comment décrire les données ?



## STOCKAGE & ARCHIVAGE

Quelles données conserver à long terme et comment ?



## ASPECTS JURIDIQUES, ÉTHIQUES, INTÉGRITÉ SCIENTIFIQUE

Que puis-je partager, réutiliser ?  
Quelles pratiques devrais-je respecter ?



## IDENTIFIANTS PÉRENNES

Comment associer durablement des données à son auteur ?



## DATA PAPERS & DATA JOURNALS

Comment publier la description de ses données comme un article scientifique ?



## PLAN DE GESTION DE DONNÉES

Pourquoi et comment rédiger un plan de gestion des données ?



## DÉPÔT & ENTREPÔTS

Comment et où déposer mes données ?



## ACCÈS & VISUALISATION

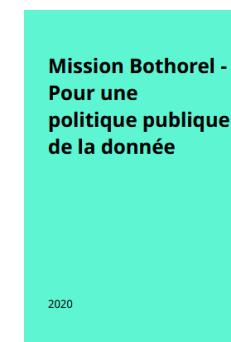
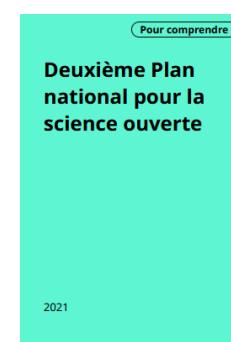
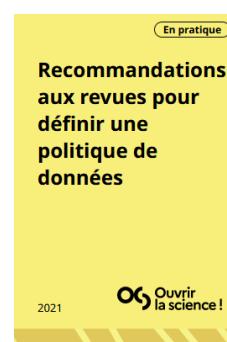
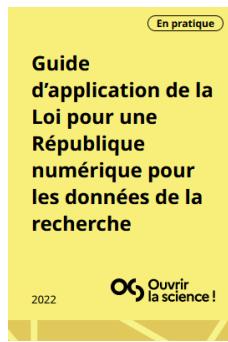
Où et comment extraire et visualiser les données qui m'intéressent ?



# Ressources importantes



Le [Comité pour la science ouverte](#) assure la mise en œuvre de la politique nationale de science ouverte. Dans le cadre de ses activités, il publie un certain nombre de guides ou de rapports concernant la science ouverte et plus particulièrement les données de la recherche. Dans la [Bibliothèque de la science ouverte](#), vous y trouverez notamment :



# Ressources importantes



Le consortium [Couperin](#) (Consortium unifié des établissements universitaires et de recherche pour l'accès aux publications numériques) est une association à but non lucratif financée par les cotisations des établissements membres et subventionnée par le Ministère de l'Enseignement Supérieur, de la Recherche. Il repose sur l'engagement actif et volontaire des agents des établissements qui le composent et d'une équipe de permanents dédiés. Spécialistes de l'information scientifique et technique, les personnels des bibliothèques et des centres de documentation des établissements y jouent un rôle prépondérant. Au sein du consortium, un groupe de travail [Science ouverte](#) s'est organisé, à l'intérieur duquel se trouve un [groupe de travail](#) se consacrant spécifiquement aux données de la recherche. Ce dernier se donne notamment pour objectif de « *produire des outils à destination des chercheurs et services supports* ».



Chaque année, un [programme de webinaire](#) est proposé à l'ensemble de la communauté universitaire. Les séances sont enregistrées et disponibles sur [Youtube](#).

## Webinaires



## Études de fond

Le GT produit également des études de fond à destination des chercheurs. Signalons notamment [Faire entrer la science ouverte dans son projet ANR : un guide pratique](#)

# Quelques ressources bibliographiques

## Généralités sur les données de la recherche

- BORGMAN Christine, *Qu'est-ce que le travail scientifique des données ? Big data, little data, no data*, traduit par Charlotte Matoussowsky, Marseille, OpenEdition Press (coll. « Encyclopédie numérique »), 2020, 420 p. [[Disponible en ligne](#)]
- DENIS Jérôme, *Le travail invisible des données. Éléments pour une sociologie des infrastructures scripturales*, Paris, Presses des Mines (coll. « Sciences sociales »), 2018, 206 p. [Disponible dans les BU de Caen ; du Havre et de Rouen]
- THEVIOT Anaïs (ed.), *Gouverner par les données ? Pour une sociologie politique du numérique*, Lyon, ENS Éditions (coll. « Gouvernement en question(s) »), 2023, 288 p. [[Disponible en ligne](#)]
- « Les sciences humaines et social au travail (II) : Que faire des données de la recherche », *Tracés*, n°19, 2019 [[Disponible en ligne](#)]

# Quelques ressources bibliographiques

## Aspects juridiques

- ROBIN Agnès, « L'ouverture des données publiques scientifiques : de l'examen de la règle “open as possible, closed as necessary” », *Communication Commerce Électronique*, septembre 2020, vol. 9, n° 15, p. 5-10. [[Disponible en ligne](#)]
- ROBIN Agnès, *Droit des données de la recherche. Science ouverte, innovation, données publiques*, Louvain-la-Neuve, Larcier, 2022, 660 p. [Disponible dans les BU de Caen ; du Havre et de Rouen]
- Rossi Julien, *La recherche en SHS au croisement de multiples contraintes. Rapport sur le cadre juridique applicable à l'usage de données issues de réseaux socio-numériques à des fins de recherche scientifique*, Paris, Centre d'expérimentation en méthodes numériques pour les recherches en SHS. Faculté des Lettres de Sorbonne Université, 2023. [[Disponible en ligne](#)]
- *Les sciences humaines et sociales et la protection des données à caractère personnel dans le contexte de la science ouverte. Guide pour la recherche*, Paris, CNRS, 2021, 35 p. [[Disponible en ligne](#)]

# Quelques ressources bibliographiques

## Ouverture et fairisation des données

- WILKINSON Mark D., DUMONTIER Michel, AALBERSBERG IJsbrand Jan, *et al.*, « The FAIR Guiding Principles for scientific data management and stewardship », *Scientific Data*, 2016, vol. 3, n° 1. [[Disponible en ligne](#)]
- *La standardisation des données ouvertes : favoriser l'interopérabilité, accroître l'impact de l'open data*, Paris, Observatoire Data Publica (coll. « Les cahiers de l'observatoire »), 2022, 79 p. [[Disponible en ligne](#)]

## Les données et la question écologique

- MALLARINO Didier, BRAS Sylvie Le et BONAMY Cyrille, « Les impacts environnementaux et sociaux des données : un défi pour l'avenir », Marseille, 2022. [[Disponible en ligne](#)]
- MOALIC Anthony, LEHOUX Élise, PION Christophe et LASNE Christophe, « La science ouverte à l'épreuve de la sobriété », *Arabesques*, 2023, n° 109, p. 12-14. [[Disponible en ligne](#)]
- THE SHIFT PROJECT, *Déployer la sobriété numérique*, Paris, 2020. [[Disponible en ligne](#)]

# Quelques ressources bibliographiques

## L'archivage des données de recherche

- HADROSSEK Christine, « Synthèse des journées d'étude RIP Data : Quelle sélection, conservation et suppression des données de recherche ? », 2024. [[Disponible en ligne](#)]
- HADROSSEK Christine, BÉZARD Laure et Boissat Romain, VALENCIA Océane (Sorbonne Université), et BACHÈLERIE Marie-Laure, « Clarifions la notion d'archivage pour les données de recherche », 2024. [[Disponible en ligne](#)]