
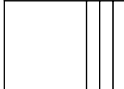



Ouvrir ses données

	URFIST de Strasbourg
	Université de Strasbourg
	

Noël Thiboud

Sommaire

1. Définition(s)
2. Les politiques publiques en matière de données
3. Principes et bonnes pratiques
4. Outils et ressources

Définition

OCDE (2007) :

« 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. »

OCDE, Organisation de Coopération et de Développement Economiques. *Principes et lignes directrices de l'OCDE pour l'accès aux données de la recherche financée sur fonds publics*. p. 18



Définition(s)

Les données de la recherche

De nombreuses définitions des données de la recherche existent. Dans notre guide, les données de la recherche concernent à la fois les échantillons, les données administratives et les données scientifiques tant manuscrites (cahiers de laboratoires...) que numériques...

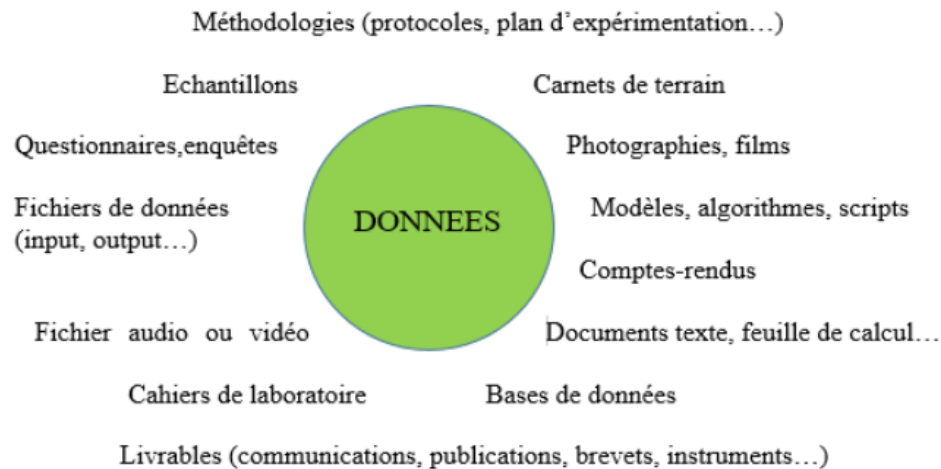
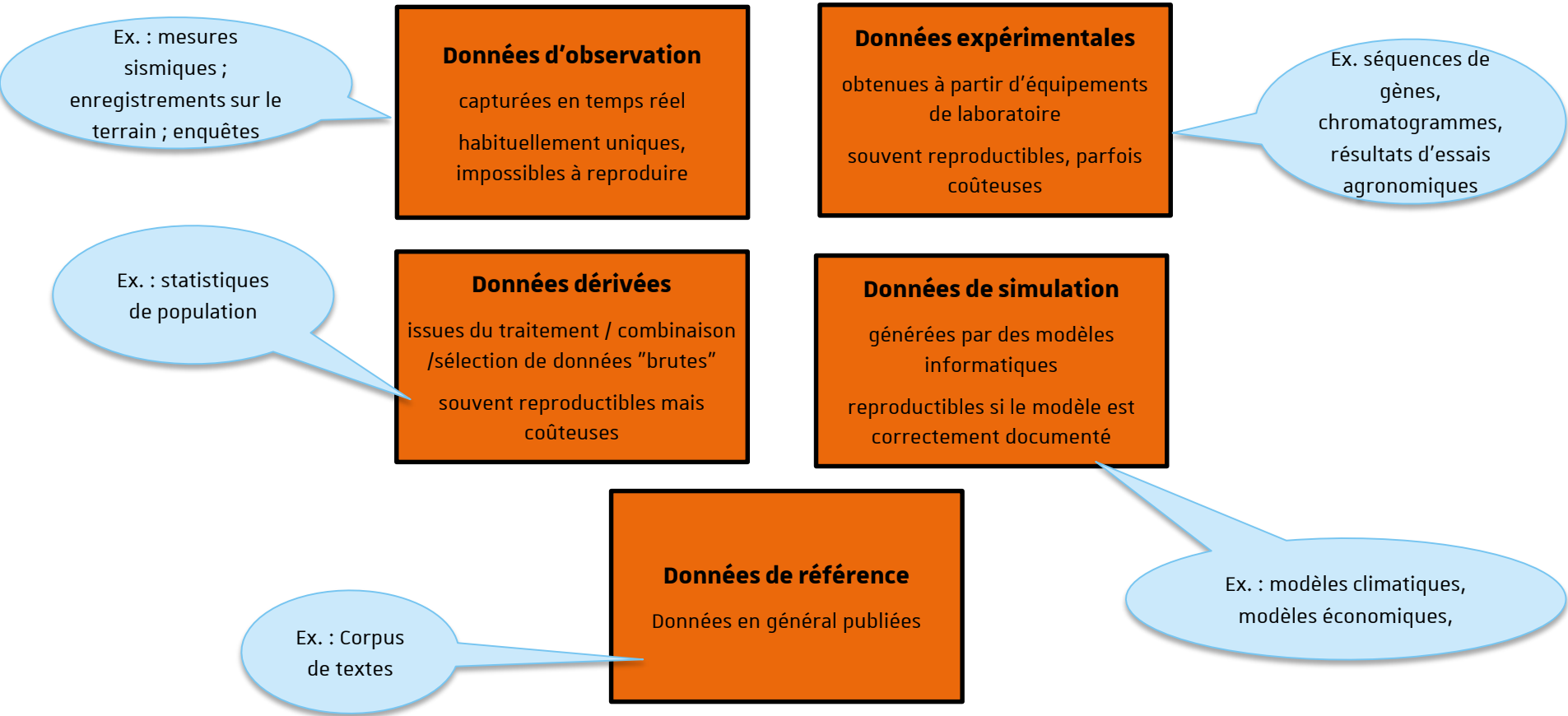


Figure 3 : La diversité des données de la recherche

Source : Alain Rivet, Marie-Laure Bachèlerie, Auriane Denis-Meyere et Delphine Tisserand - Traçabilité des activités de recherche et gestion des connaissances - Guide pratique de mise en place – 2018 - http://qualite-en-recherche.cnrs.fr/IMG/pdf/guide_tracabilite_activites_recherche_gestion_connaissances.pdf

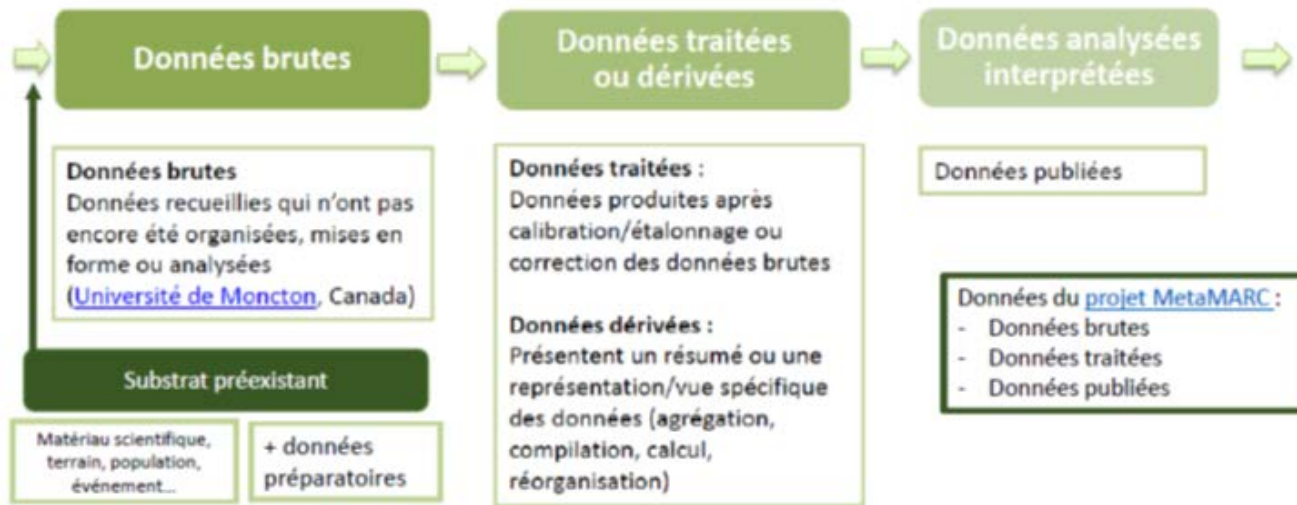
Typologie(s)



D'après Doranum, Typologie par origine <https://doranum.fr/plan-gestion-donnees-dmp/origine-description-donnees-recherche/> et Research Information Network). *Stewardship of digital research data: a framework of principles and guidelines*. London : Research Information Network, 2008.

Typologie(s)

Continuum données brutes ---> analysées



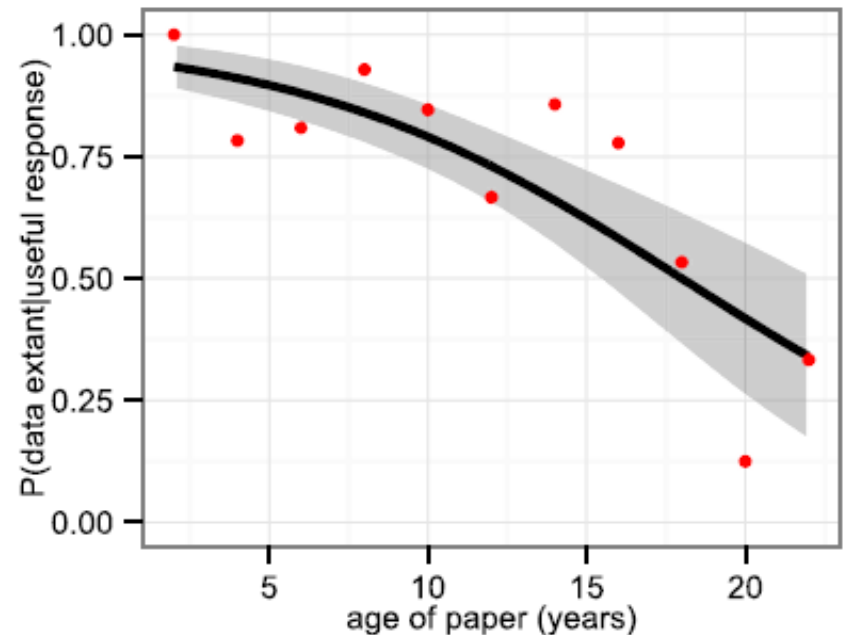
- **Données primaires** « données collectées par la personne qui effectue la recherche », **Données secondaires** « données collectées par une personne différente de l'utilisateur » ([Wikipédia](#), 2016)
- **Données sources** « données utilisées par les chercheurs pour leur recherche » **Données résultats** « données produites comme résultats de recherche » (Prost, 2015)

Source : A. Ciolek-Figiel et F. Cosserrat, « Gestion et diffusion des données de la recherche », 2016 ([en ligne](#))

Pourquoi ouvrir les données ?

Enjeux scientifiques

- deux années après la publication d'un article, les chances d'accéder aux données scientifiques de l'étude chutent de près de 17 % par an
- réutiliser des données de la recherche :
 - capitaliser les savoirs et savoir-faire d'une structure de recherche
 - : témoignage sur le site [Datacc](#)
 - favoriser de nouvelles recherches



(D) Predicted probability that the data were extant (either "shared" or "exist but unwilling to share") given that we received a useful response. In all panels, the line indicates the predicted probability from the logistic regression, the gray area shows the 95% CI of this estimate, and the red dots indicate the actual proportions from the data.

T. Vines et al., « The availability of research data declines rapidly with article age », *Journal of Current Biology*, 2014.

Pourquoi ouvrir les données ?

Enjeux scientifiques

- garantir la qualité scientifique d'une recherche :
 - reproductibilité des résultats
 - traçabilité des données
 - Intégrité scientifique

Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis



Mandeep R Mehra, Sapan S Desai, Frank Ruschitzka, Amit N Patel

Summary

Background Hydroxychloroquine or chloroquine, often in combination with a second-generation macrolide, are being widely used for treatment of COVID-19, despite no conclusive evidence of their benefit. Although generally safe when used for approved indications such as autoimmune disease or malaria, the safety and benefit of these treatment regimens are poorly evaluated in COVID-19.

Methods We did a multinational registry analysis of the use of hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19. The registry comprised data from 671 hospitals in 36 countries. We included patients hospitalised between Dec 20, 2019, and April 14, 2020, with a positive laboratory finding for SARS-CoV-2. Patients who received one of the treatments of interest within 48 h of diagnosis were included in one of four treatment groups (chloroquine alone, chloroquine with a macrolide, hydroxychloroquine alone, or hydroxychloroquine with a macrolide), and patients who received none of these treatments formed the control group. Patients for whom one of the treatments of interest was initiated more than 48 h after diagnosis or while they were on mechanical ventilation, as well as patients who received remdesivir, were excluded. The main outcomes of interest were in-hospital mortality and the occurrence of de-novo ventricular arrhythmias (as defined on the basis of sustained ventricular tachycardia or ventricular fibrillation).

Findings 96 032 patients (mean age 53·8 years, 46·2% women) with COVID-19 were hospitalised during the study period and met the inclusion criteria. Of these, 67 032 patients were in the treatment groups (1868 received chloroquine, 3783 received chloroquine with a macrolide, 3016 received hydroxychloroquine, and 6221 received hydroxychloroquine with a macrolide) and 29 000 patients were in the control group. 10 698 (11·1%) patients died in hospital. After controlling for multiple confounding factors (age, sex, race or ethnicity, body-mass index, underlying cardiovascular disease and its risk factors, diabetes, underlying lung disease, smoking, immunosuppressed condition, and baseline disease severity), we compared with mortality in the control group (9·3%), hydroxychloroquine (18·0%; hazard ratio 1·335, 95% CI 1·235–1·447), hydroxychloroquine with a macrolide (23·8%; 1·447, 1·368–1·531), chloroquine (16·4%; 1·365, 1·235–1·531), and chloroquine with a macrolide (22·2%; 1·368, 1·273–1·469) were each independently associated with an increased risk of in-hospital mortality. Compared with the control group (0·3%), hydroxychloroquine (6·2%; 2·366, 1·935–2·900), hydroxychloroquine with a macrolide (8·1%; 5·106, 4·106–5·983), chloroquine (4·3%; 1·701, 1·404–4·596), and chloroquine with a macrolide (6·5%; 4·011, 3·344–4·812) were independently associated with an increased risk of de-novo ventricular arrhythmia during hospitalisation.

Interpretation We were unable to confirm a benefit of hydroxychloroquine or chloroquine, when used alone or with a macrolide, on in-hospital outcomes for COVID-19. Each of these drug regimens was associated with decreased in-hospital mortality, but also with an increased frequency of ventricular arrhythmias when used for treatment of COVID-19.

Funding William Moseley Distinguished Chair in Advanced Cardiovascular Medicine at Brigham and Women's Hospital.

Copyright © 2020 Elsevier Ltd. All rights reserved.

Introduction

drugs have been shown in laboratory conditions to have

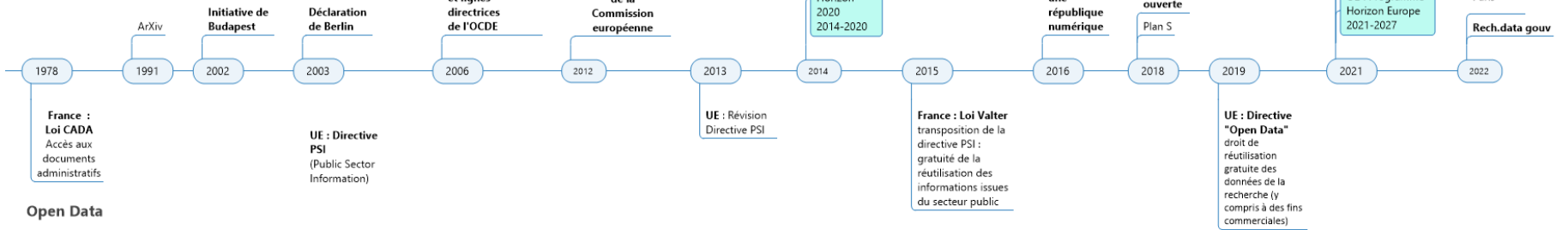
Preprint
DOI: [https://doi.org/10.1016/S0140-6736\(20\)31180-6](https://doi.org/10.1016/S0140-6736(20)31180-6)
This online publication has been corrected. The corrected version first appeared at thebmj.com on May 29, 2020.

See Online/Comment
[https://doi.org/10.1016/S0140-6736\(20\)31180-6](https://doi.org/10.1016/S0140-6736(20)31180-6)
Brigham and Women's Hospital
Heart and Vascular Center and
Harvard Medical School
Boston, MA, USA
(Prof M R Mehra MD);
Surgisphere Corporation,
Chicago, IL, USA (S S Desai MD);
University Heart Center,
University Hospital Zurich,
Zurich, Switzerland
(Prof F Ruschitzka MD);
Department of Biomedical
Engineering, University
of Utah, Salt Lake City, UT, USA
(A N Patel MD); and HCA
Research Institute, Nashville,
TN, USA (A N Patel)

Correspondence to:
Prof Mandeep R Mehra, Brigham
and Women's Hospital Heart and
Vascular Center and Harvard
Medical School, Boston,
MA 02115, USA
mmehra@bwh.harvard.edu

Les politiques publiques d'ouverture des données : de l'open data à l'open research data

Open Access /
Science ouverte



Open Data

Loi pour une République numérique (2016)

Dès lors que les données issues d'une activité de recherche financée au moins pour moitié par des dotations de l'État, 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.

Loi n°2016-1321 du 7 octobre 2016 - art. 30

Loi pour une République numérique (2016)

1. Les données sont des informations publiques
2. Principe d'ouverture par défaut
Des exceptions protègent néanmoins certains types de données
3. Principe de libre réutilisation (y compris à des fins commerciales) et de gratuité : la réutilisation n'est pas soumise au paiement d'une redevance (sauf pour liste limitative d'établissements)

2^e Plan national pour la science ouverte (2021)



OUVRIR LA SCIENCE

LA SCIENCE OUVERTE LE COMITÉ GROUPES BLOG AGENDA RESSOURCES

FR - EN



POLITIQUES PUBLIQUES

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 pour développer l'accès ouvert aux publications et – autant que possible – aux données de la recherche.

Sommaire

Introduction 6

Premier axe
Généraliser l'accès ouvert aux publications 8

Mesures 10

- 1 Généraliser l'obligation de publication en accès ouvert des articles et livres issus de recherches financées par appel à projets sur fonds publics
- 2 Soutenir les modèles économiques d'édition en accès ouvert sans frais de publication pour les auteurs (modèle « diamant »)
- 3 Favoriser le multilinguisme et la circulation des savoirs scientifiques par la traduction des publications des chercheurs français

Deuxième axe
Structurer, partager et ouvrir les données de la recherche 12

Mesures 14

- 4 Mettre en œuvre l'obligation de diffusion des données de recherche financées sur fonds publics
- 5 Créer Recherche Data Gov, la plateforme nationale fédérée des données de la recherche
- 6 Promouvoir l'adoption d'une politique de données sur l'ensemble du cycle des données de la recherche, pour les rendre faciles à trouver, accessibles, interopérables et réutilisables (FAIR)

4

Troisième axe
Ouvrir et promouvoir les codes sources produits par la recherche 16

Mesures 18

- 7 Valoriser et soutenir la diffusion sous licence libre des codes sources issus de recherches financées sur fonds publics
- 8 Mettre en valeur la production des codes sources de l'enseignement supérieur, de la recherche et de l'innovation
- 9 Définir et promouvoir une politique en matière de logiciels libres

Quatrième axe
Transformer les pratiques pour faire de la science ouverte le principe par défaut 20

Mesures 22

- 10 Développer et valoriser les compétences de la science ouverte tout au long du parcours des étudiants et des personnels de la recherche
- 11 Valoriser la science ouverte et la diversité des productions scientifiques dans l'évaluation des chercheurs et enseignants-chercheurs, des projets et des établissements de recherche
- 12 Tripler le budget de la science ouverte en s'appuyant sur le Fonds national pour la science ouverte et le Programme d'investissements d'avenir

Bilan du Premier Plan national pour la science ouverte 26

ANNEXES 28

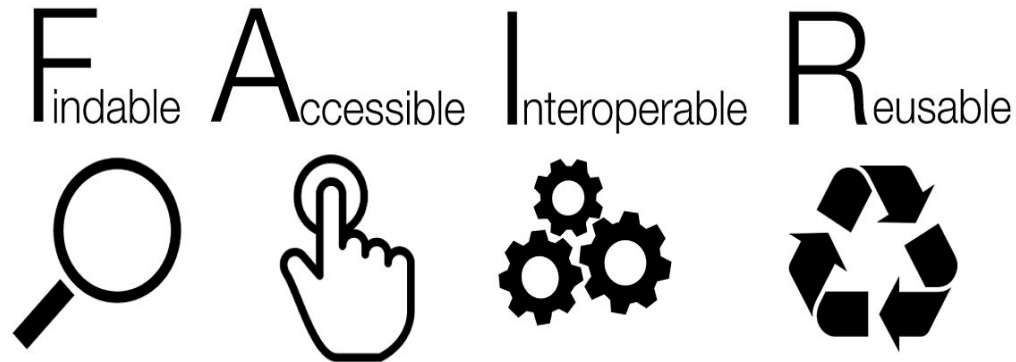
5

Gestion des données

“ As open as possible,
as closed as necessary ”

European Commission. *New Guidelines on FAIR Data Management in Horizon 2020*, 2016

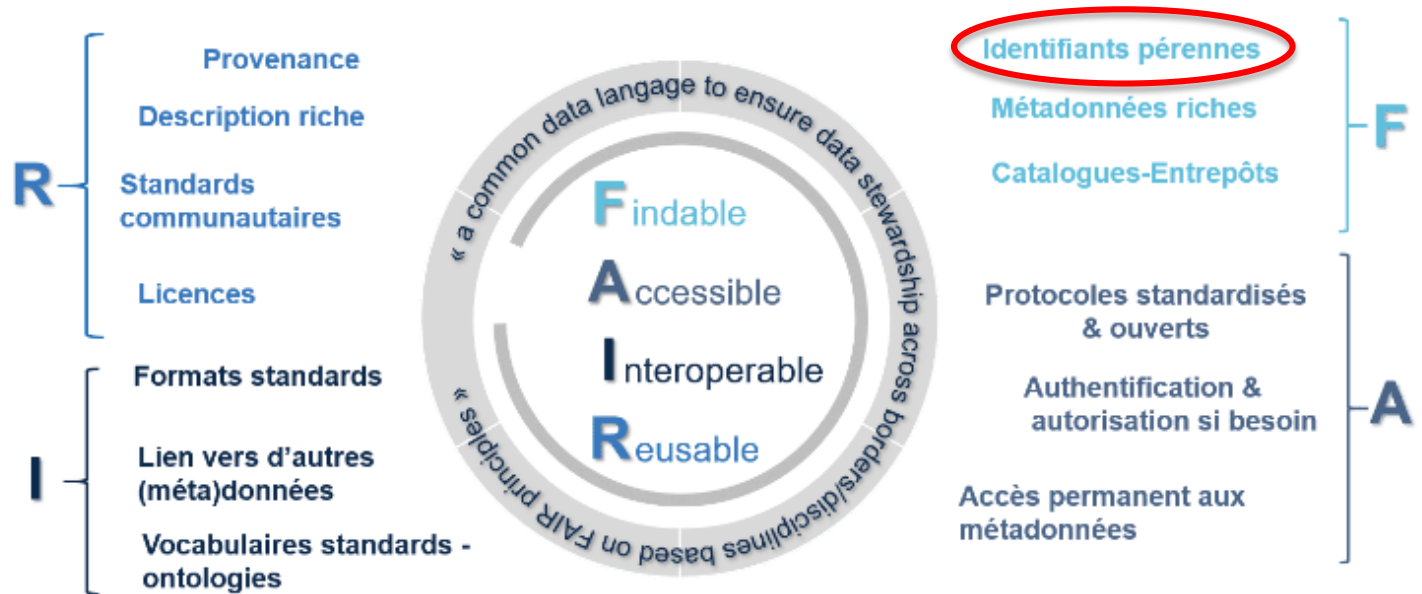
Les principes FAIR



Wilkinson Mark D., Dumontier Michel, Aalbersberg IJsbrand Jan, [et al.]. « The FAIR Guiding Principles for scientific data management and stewardship ». *Scientific Data*. 15 mars 2016, vol. 3, p. 160018. DOI < [10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18) >.

Logo FAIR : SangyaPundir

Les principes FAIR



Source:

Implementation Roadmap for the European Science Cloud (Staff Working Document SWD(2018) 83), 14 March 2018

Traduction Inra <https://www6.inra.fr/datapartage/Produire-des-donnees-FAIR>

DoRANum . [Parcours interactif sur la gestion des données de la recherche \(2020\)](#)

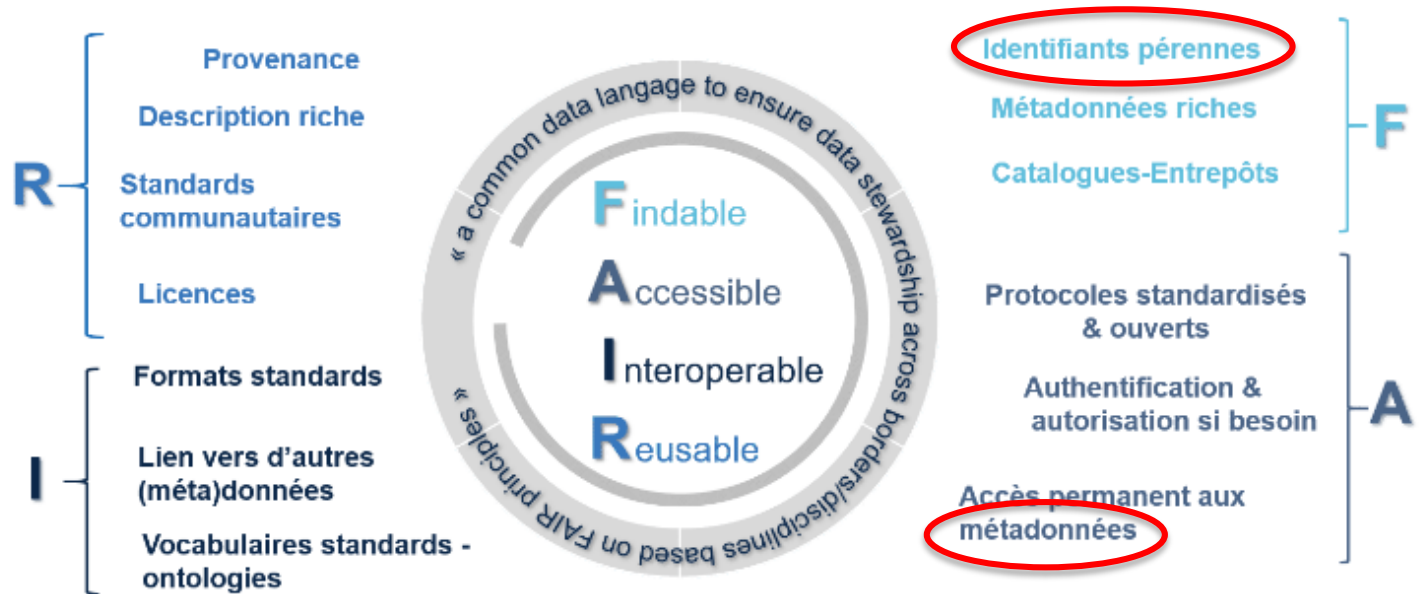
Un outil : les identifiants

- Définition : code numérique unique, invariable, associé de manière permanente à une personne physique ou morale (chercheur, institution) ou à un objet (> donnée, publication)
- pérenne quel que soit le chemin de la ressource ou de la personne
- ouverts et interopérables = indépendants des plateformes
- Rendent les données accessibles et citables, permettent de les lier aux publications, augmentent la visibilité
- Plusieurs systèmes d'identifiant (selon leur producteur et/ou leur objet)

Quelques Exemples

- DOI : Digital Object Identifier, identifiant numérique d'objet
 - Identifie aussi bien un article qu'un jeu de données
- ORCID (Open Research and Contributor ID) identifie de façon univoque un chercheur dans un environnement numérique (Base de données, site de dépôt de publications ou de données ...)
- En savoir plus, cf. page [Identifiants pérennes](#) sur DoRANum

Les principes FAIR



Source:

Implementation Roadmap for the European Science Cloud (Staff Working Document SWD(2018) 83), 14 March 2018

Traduction Inra <https://www6.inra.fr/datapartage/Produire-des-donnees-FAIR>

DoRANum . [Parcours interactif sur la gestion des données de la recherche \(2020\)](#)



Définition et cycle de vie

Métadonnées embarquées et enrichies

Différence entre standard et schéma

Exemples de standards

L'enrichissement des métadonnées

Utilité des métadonnées

EXEMPLES DE STANDARDS DE MÉTADONNÉES

Cliquer pour voir le détail

Dublin Core

DataCite
Metadata Schema

Le Dublin Core



Les 15 éléments de base

DDI
(Data Documentation Initiative)

MIDAS-
Heritage

DwC
(Darwin Core)

EML
(Ecological Metadata Language)

Guide OTELo



Définition et cycle de vie

Métadonnées embarquées et enrichies

Différence entre standard et schéma

Exemples de standards

L'enrichissement des métadonnées

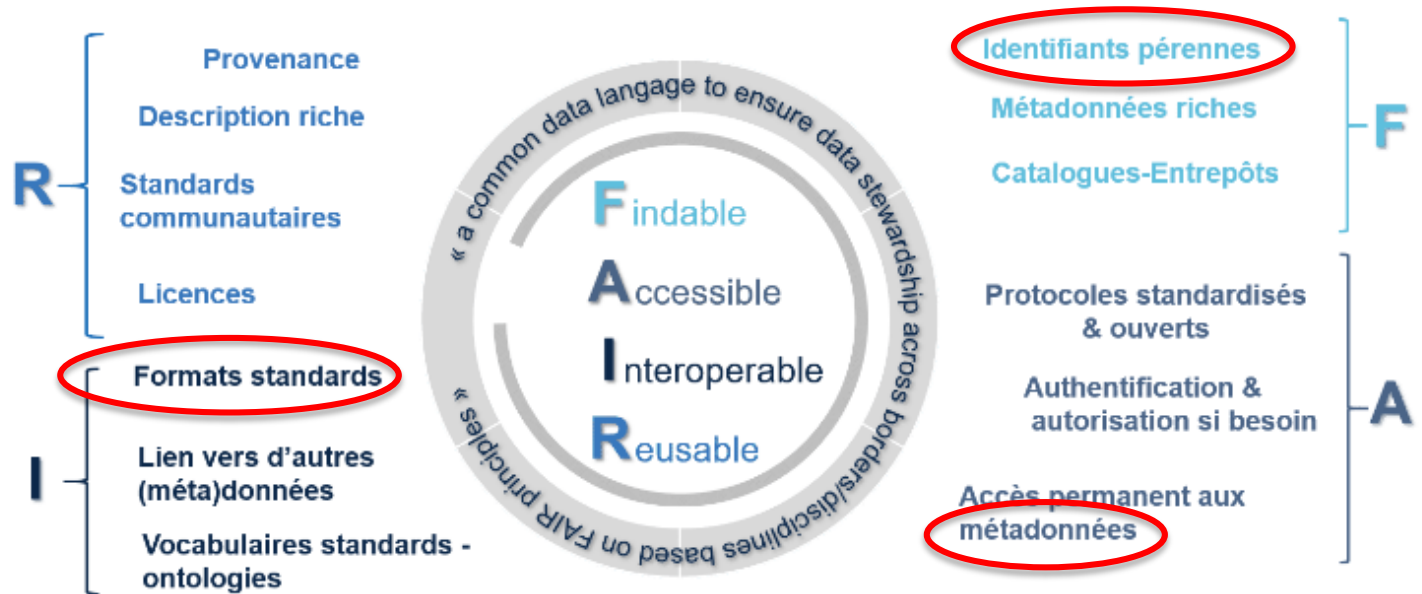
Utilité des métadonnées

■ UTILITÉ DES METADONNÉES

Pour résumer, voici à quoi servent les métadonnées en 5 points :

- ➔ Comprendre l'**origine** des données et leur **contexte** de création ou de collecte
- ➔ Améliorer le **moissonnage par les machines** (moteur de recherche)
- ➔ Garantir l'**interopérabilité**
- ➔ Connaître les **conditions de réutilisation et de partage** des données
- ➔ Fournir des **informations** très utiles lorsque les données ne peuvent pas être partagées (embargo, accès restreint) ou lors du retrait des données (données obsolètes, etc.).

Les principes FAIR



Source:

Implementation Roadmap for the European Science Cloud (Staff Working Document SWD(2018) 83), 14 March 2018

Traduction Inra <https://www6.inra.fr/datapartage/Produire-des-donnees-FAIR>

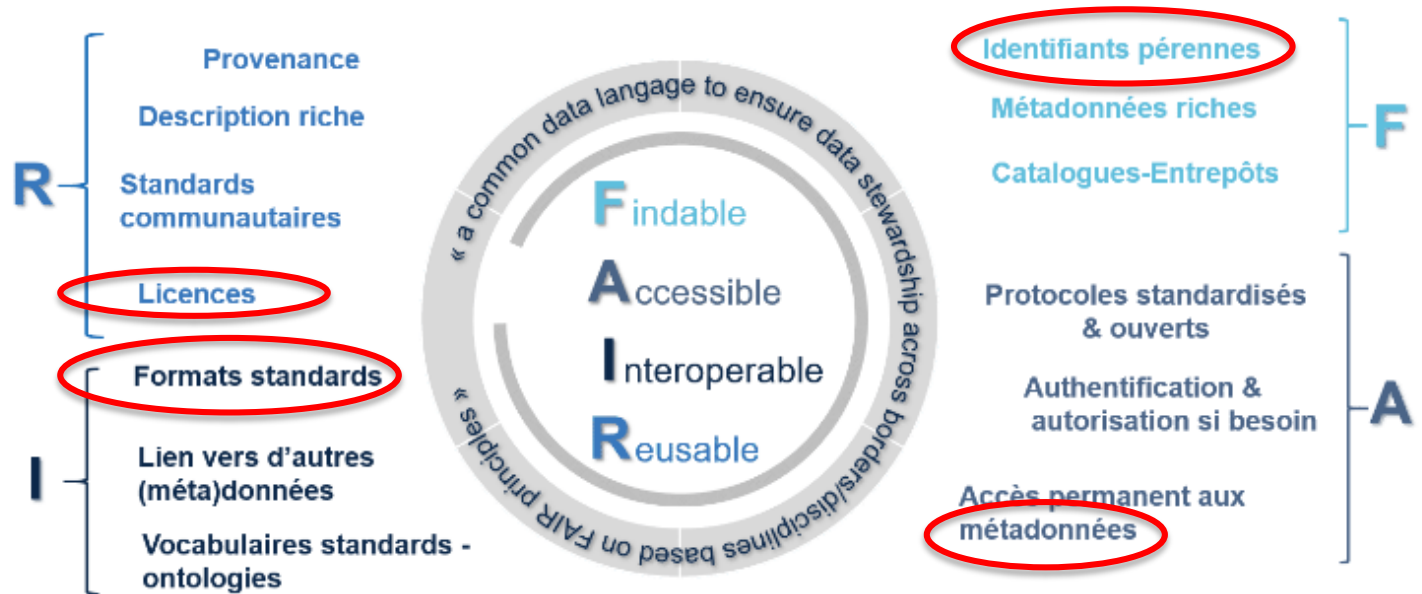
DoRANum . [Parcours interactif sur la gestion des données de la recherche \(2020\)](#)

Format de fichier déconseillé	Format de fichier à privilégier
Excel (.xls, .xlsx) Word (.doc, .docx) Powerpoint (.ppt, .pptx) Photoshop (.psd) Quicktime (.mov)	Comma Separated Values (.csv) Texte ascii (.txt) ou PDF/A si formatage PDF/A (.pdf) TIFF (.tif, .tiff) MPEG-4 (.mp4)

Le Centre Informatique National de l'Enseignement Supérieur (CINES) met à disposition un service de validation de formats destiné à vérifier l'éligibilité de documents sur sa plateforme d'archivage (<https://facile.cines.fr/>).

Source : Alain Rivet, Marie-Laure Bachèlerie, Auriane Denis-Meyere et Delphine Tisserand - Traçabilité des activités de recherche et gestion des connaissances - Guide pratique de mise en place – 2018 - http://qualite-en-recherche.cnrs.fr/IMG/pdf/guide_tracabilite_activites_recherche_gestion_connaissances.pdf

Les principes FAIR



Source:

Implementation Roadmap for the European Science Cloud (Staff Working Document SWD(2018) 83), 14 March 2018

Traduction Inra <https://www6.inra.fr/datapartage/Produire-des-donnees-FAIR>

DoRANum . [Parcours interactif sur la gestion des données de la recherche \(2020\)](#)

Les licences de réutilisation

- La licence d'utilisation précise les conditions dans lesquelles vos données peuvent être réutilisées.
- Suite à la loi pour une République numérique, un décret précise [la liste des licences](#) qui peuvent être utilisées par les administrations pour la réutilisation à titre gratuit de leurs informations publiques
- Exemples :



- Licence ouverte : pour publier des données mais aussi des textes et des œuvres originales numérisées
- ODBL pour les Bases de données (ex [OpenStreetMap](#))



- Creative Commons : pas dans la liste officielle mais très utilisées

Les licences Creative commons

- Principe : Passer d'une approche « Tous droits réservés » (nécessitant une autorisation préalable) à « certains droits réservés » (permettant certains usages a priori)



- Objectifs :
 - autoriser gratuitement la reproduction et la diffusion (sous certaines conditions)
 - accorder plus de droits aux utilisateurs **en complétant le droit d'auteur** qui s'applique par défaut
- ensemble de licences avec 4 options combinables pour ouvrir plus ou moins de droits à l'utilisateur
- En savoir plus : [Article Creative commons \(Wikipedia\)](#)

Documentation et partage des données



« On refait la même chose parce qu'on ne sait plus ce qu'on a fait. Et on ne sait plus ce qu'on a fait parce que c'est mal archivé, mal valorisé, mal documenté.»

« Nous, le problème (...) c'est de trouver des moyens de partager nos données à l'intérieur d'une équipe qui mène un même projet. On a déjà du mal actuellement à avoir des outils qui nous permettent de mettre en commun nos données, même à petite échelle, à échelle intersite.»

Conservation et archivage des données

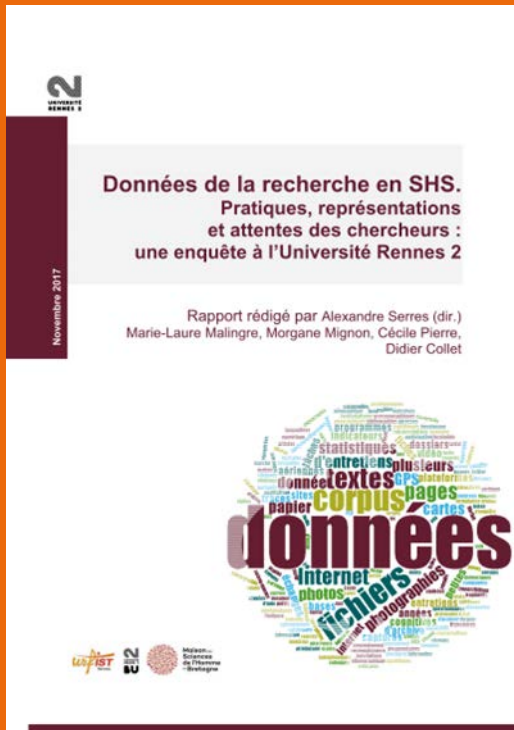
- Organiser un espace collaboratif dédié au projet
 - serveurs locaux (machines virtuelles), cloud institutionnel...
 - contrôle des accès et modalités de partage des données
- Estimer, indiquer la volumétrie prévisionnelle des données

Archivage / Partage des données

Bonnes pratiques

- règles de nommage des fichiers
- gestion des versions
- Règles de classement des données pour un partage intelligible entre partenaires
- privilégier des formats de fichiers et des logiciels ouverts, standardisés et pérennes.
- Comment seront préservées les données à long terme ? (obsolescence des formats)

Questions juridiques



« sur des données communes, qui a la propriété des données ? Et là, je ne parle pas des résultats scientifiques, je parle des données brutes ». [...] « on s'est retrouvé bloqués pour les utiliser parce qu'on n'avait pas de réponse sur ce qu'on avait le droit de faire avec ces données »

« on a été confronté au problème il y a quelque temps. En l'occurrence j'ai écrit un article que je n'ai jamais pu soumettre parce que je risquais de m'exposer à des problèmes, j'ai préféré laisser mes résultats de côté, c'est un peu dommage parce que c'était un projet chronophage, intéressant et scientifiquement porteur »

Questions juridiques autour des données

- Propriété
- Traitement
- Diffusion
- Réutilisation

réutilisation de données issues d'une activité de recherche

La loi pour une République numérique du 7 octobre 2016 prévoit qu'elle est libre si :

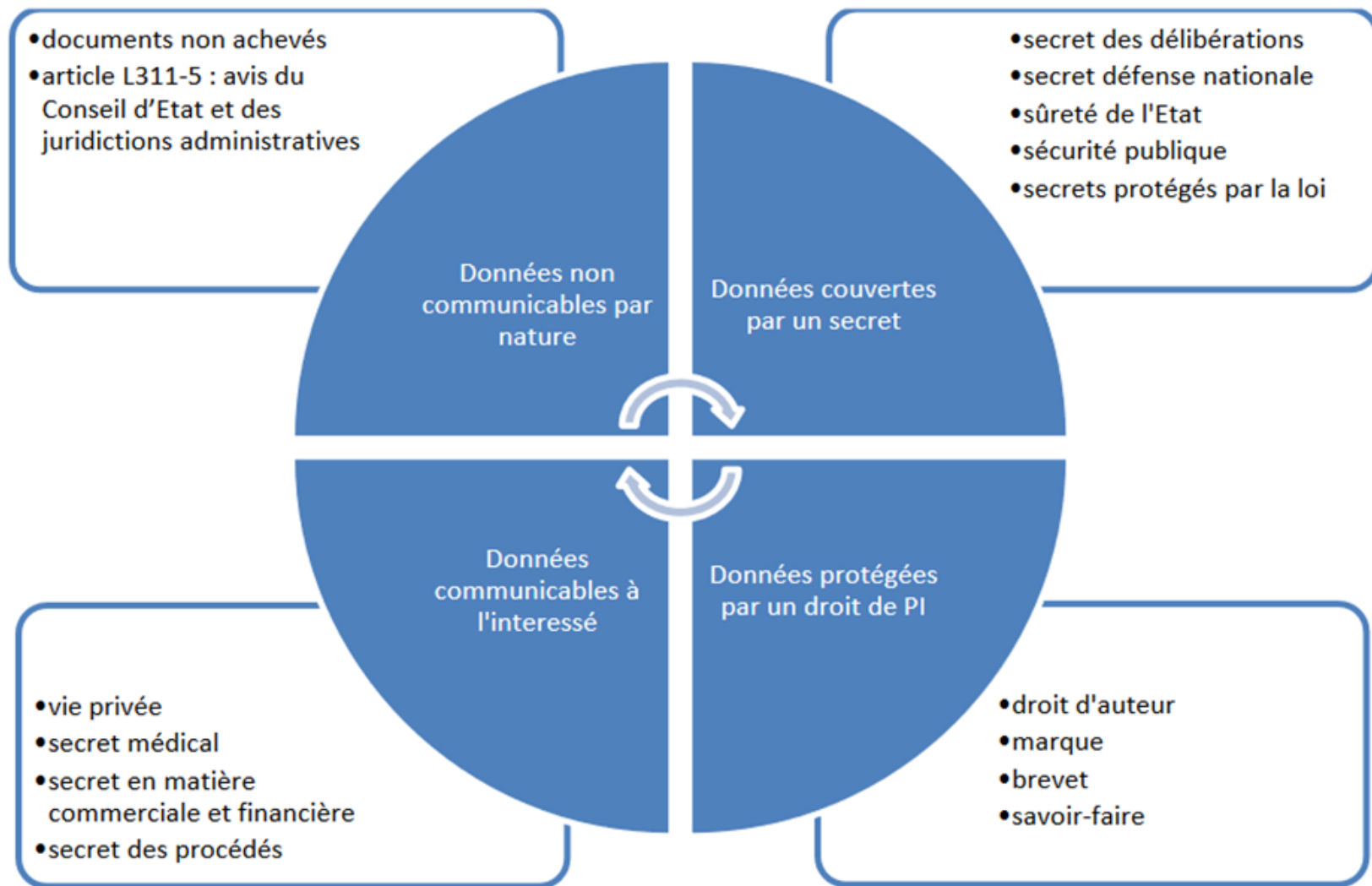
- ces données sont issues d'une activité de recherche financée au moins pour moitié par des fonds publics ;
- ces données ne sont pas protégées par un droit spécifique ;
- ces données ont été rendues publiques par le chercheur ou l'établissement.
- **Licence de diffusion** : attribuer une licence de diffusion lorsque les données sont communicables permet d'afficher clairement les modalités de réutilisation

Questions juridiques autour des données

- Propriété
- Traitement
- Diffusion

- Les données produites par les établissements de recherche et d'enseignement dans le cadre de leur mission de service public sont considérées comme des **documents administratifs** et sont donc communicables à toute personne qui en fait la demande, sauf exceptions légales.
 - « documents administratifs » : également les données produites par ces établissements = données brutes, données élaborées et métadonnées.
- ne peuvent être accessibles au public que les documents « achevés ». Par conséquent, tous les documents préparatoires ne sont pas communicables.

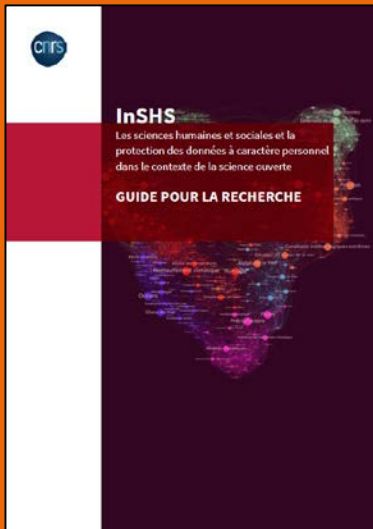
Exceptions légales à la diffusion des données



Source : Y, Lafosse. *Initiation aux données de la recherche* – [Formation Urfist Paris – 27 et 28 mars 2019](#)

Questions juridiques autour des données

- Propriété
- Traitement



Recueil des données : identifier les différents droits mis en jeu (à l'occasion d'entretiens, de prises de vues, de collecte d'informations, de consultation d'archives..., etc.).

- **Règlement Général sur la Protection des Données (RGPD) (2016)** : si collecte de données personnelles (données permettant l'identification directe ou indirecte d'une personne), respecter des principes essentiels sur
 - la durée de conservation des données
 - le droit à l'information (consentement éclairé)
 - obligation de sécuriser les données
 - Consulter le Délégué à la protection des données (DPD/DPO) de votre établissement
-
- CNRS-InSHS, Isabelle André-Poyaud, et Sandrine Astor. « Les sciences humaines et sociales et la protection des données à caractère personnel dans le contexte de la science ouverte. Guide pour la recherche ». InSHS, 2021, 36 p.
 - [webinaire](#) du 11/01/2021 sur DoRANum

Sur les questions juridiques



ACCUEIL RESSOURCES ACTUALITES A PROPOS CONTACT



— Où en sommes-nous par rapport à la loi sur les données ? —



Doranum. *Données de la recherche – contexte juridique : qui a les droits, quelles obligations ?* Logigramme [DOI : 10.13143/8dh5-d615](https://doi.org/10.13143/8dh5-d615)

Ouverture des données de la recherche. Guide d'analyse du cadre juridique en France, 2017, 45 p.



Un outil : le plan de gestion des données (PGD)

en anglais : Data management plan ou DMP

Le plan de gestion des données est un outil d'aide à la gestion des données d'un projet de recherche. Il a pour objectif d'expliquer la façon dont les données de recherche seront gérées durant le projet et après son terme

Certains financeurs de la recherche (ANR, Horizon Europe) l'intègrent dans les livrables du projet financé.



FRANCE. MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE, *Passeport pour la science ouverte : Guide pratique à l'usage des doctorants*, 2021

Structure

Le PGD /DMP se présente sous la forme d'un document structuré en rubriques.

Il permet

- d'anticiper les questions qui peuvent se poser en matière de données au cours d'un projet de recherche : propriété, partage, réutilisation, pérennisation des données...
- de faciliter la coordination de la gestion des données au sein d'une équipe de recherche



modèle [DMP H2020](#) établi par la Commission européenne

Source : :Doranum

Autres modèles : <https://dmp.opidor.fr>

Rechercher des entrepôts



Re3data est un répertoire mondial de dépôts de données de recherche pluridisciplinaire. Il a été mis en ligne à l'automne 2012 et est financé par la Fondation allemande pour la recherche (DFG).

Recherche Data Gov

Entrepôt pluridisciplinaire pour les données de la recherche

Caractéristiques d'un entrepôt

Quel type d'entrepôt
privilégieriez-vous ?
Disciplinaire ? Pluridisciplinaire ?

L'entrepôt est-il certifié, reconnu
au niveau international ?

Quels types de données sont
acceptés par l'entrepôt ?

Quels formats sont acceptés ?

L'entrepôt impose-t-il un
standard de métadonnées ?

Impose-t-il une licence ?
Vous convient-elle ?

Attribue-t-il un identifiant
pérenne aux données ?

La plateforme est-elle simple
d'utilisation ?
Des aides sont-elles proposées ?

Pourrez-vous modifier les
données déposées et accéder aux
anciennes versions ?

L'entrepôt est-il localisé en
France ?

L'entrepôt permet-il de
restreindre l'accès aux données ?
De mettre un embargo ?

L'entrepôt s'engage-t-il à
préserver l'accès aux données sur
le long terme ?

Le dépôt est-il payant ou gratuit ?

L'entrepôt propose-t-il un lien
entre les données et les
publications associées ?

Des statistiques d'utilisation, de
consultation, de téléchargement
sont-elles disponibles ?

Doranum. *Dépôt et Entrepôts : fiche synthétique*
[DOI : 10.13143/a3d4-7553](https://doi.org/10.13143/a3d4-7553)

Quelques exemples d'entrepôts



Entrepôt multidisciplinaire

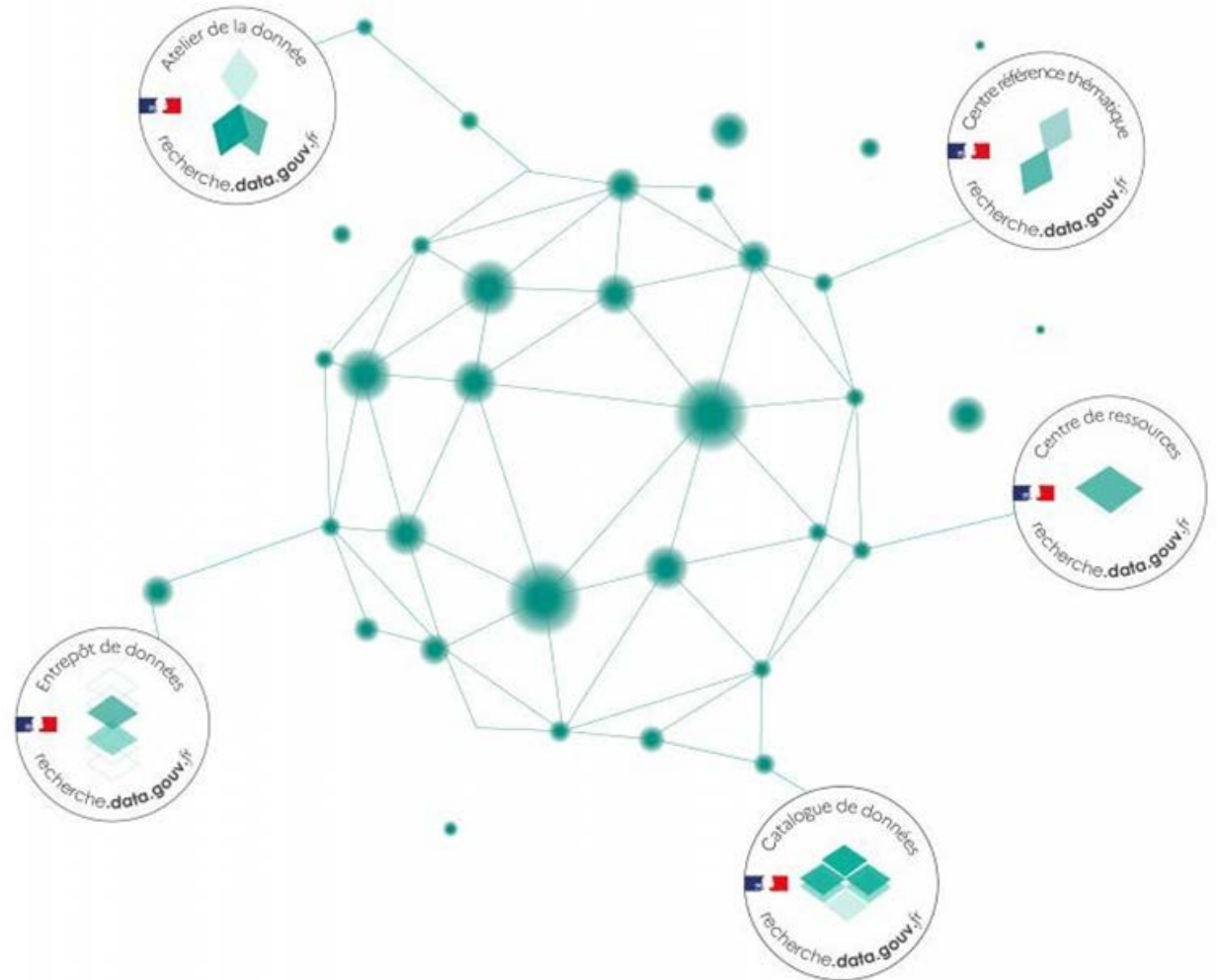


Entrepôt multidisciplinaire

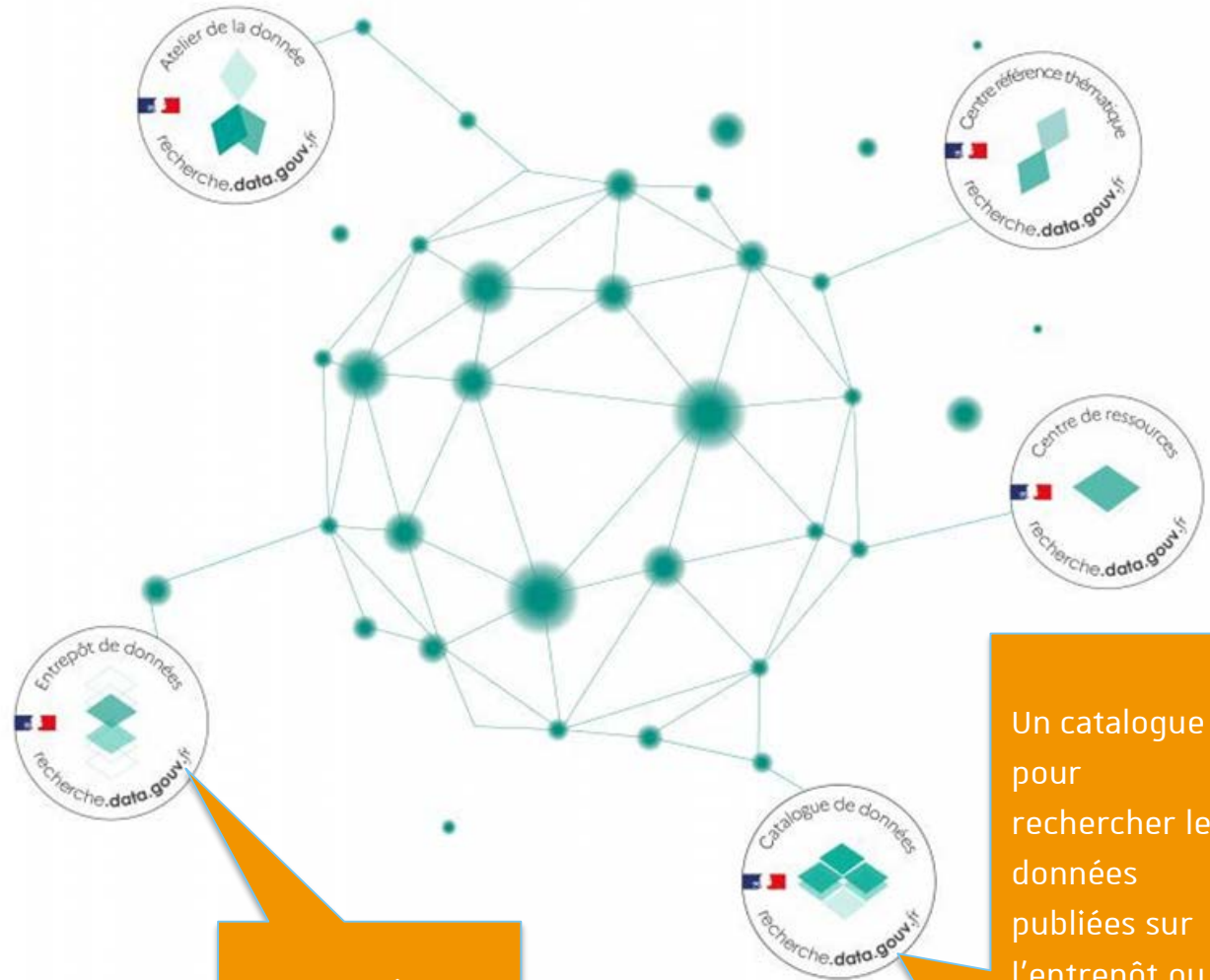


Entrepôt - Sciences de la vie

L'écosystème Recherche Data Gouv



L'écosystème Recherche Data Gouv

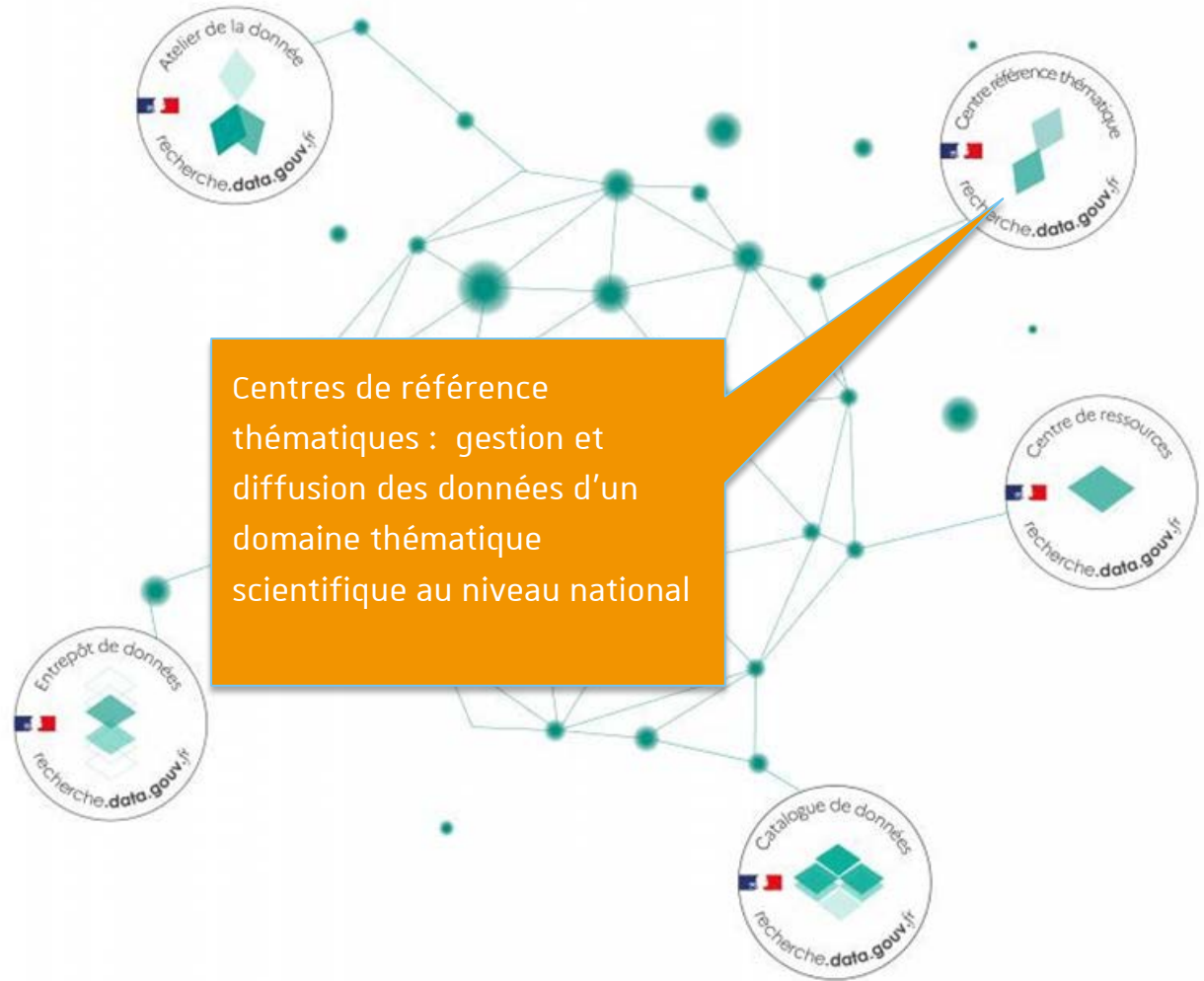


Un entrepôt pour déposer et utiliser des données

Un catalogue pour rechercher les données publiées sur l'entrepôt ou sur des entrepôts externes

Centres de référence thématiques

Recherche Data Gouv



Centres de référence thématiques



Astronomie et astrophysique



Système Terre et Environnement



Sciences Humaines et Sociales



Biologie-Santé



Réutilisation et citation des données

- Vérifier que la licence de diffusion appliquée par le(s) auteur(s) aux données que l'on souhaite utiliser est adaptée à ses besoins
- Respecter la propriété intellectuelle des auteurs
- Citer les données si la licence l'exige (il est recommandé de toujours citer ses sources)
- Lier les données aux publications
- Partager toute nouvelle donnée qui pourrait être générée en utilisant le travail des auteurs originaux conformément à la licence de diffusion

Publier ses données

- Les déposer dans un entrepôt de données
- dans des fichiers associés à l'article dans une revue scientifique (***supplementary materials***).
- dans un *data journal* : revue spécialisée dans les ***data papers***

Publier ses données : les data papers

Exemples

- **En STM :** <https://www.nature.com/articles/sdata201828>
- **En SHS:** [Research Data Journal for the Humanities and Social Sciences](#)(Brill)

Le *Data paper* est un article scientifique qui permet d'informer la communauté scientifique de l'existence d'un jeu de données et de le documenter pour faciliter sa réutilisation.

Structure

- **Partie descriptive** : Éléments communs aux articles classiques : titre, résumé, mots-clés... + Éléments spécifiques aux données : types de données, formats, processus et méthodes de production, métadonnées, réutilisation...
- **Accès aux données** : intégrées dans l'article ou déposées dans un entrepôt
- **L'identifiant des données** (exemple DOI) permet d'établir le lien du data paper vers les données

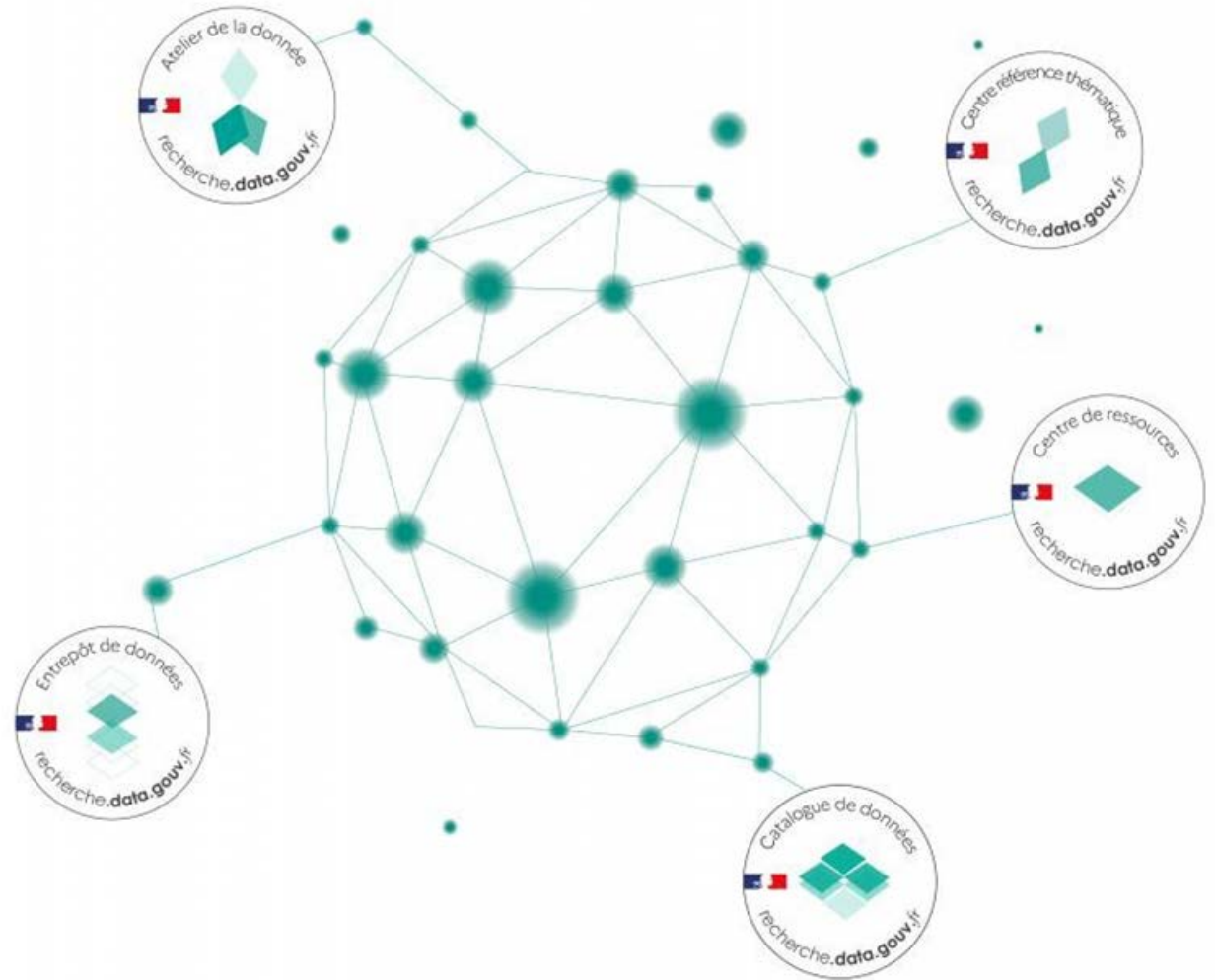


Personnels d'appui à la recherche

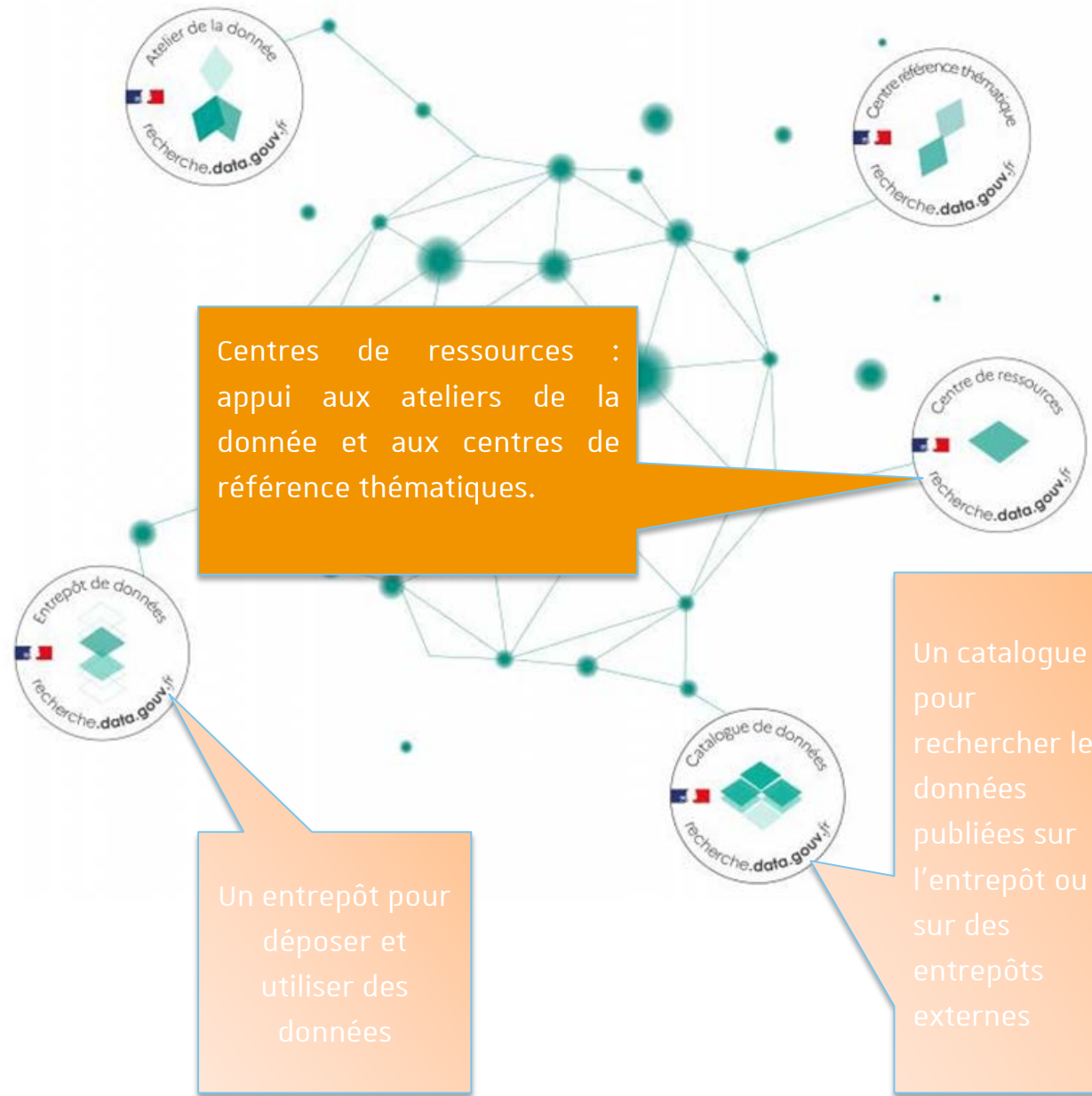
Dans les établissements, de nombreux métiers apportent leur expertise tout au long du cycle des données de recherche :

- **bibliothécaires ,documentalistes, archivistes**
: indexation, gestion des métadonnées, participation à l'archivage, gestion des archives de la recherche...
- **Ingénieurs** : collecte, traitement et analyse des données : [plateformes universitaires de données](#)
- **informaticiens** : interopérabilité, stockage, diffusion des données, gestion des entrepôts de données,
- **Délégué(e) à la protection des données (DPO – Data Protection Officer)** : conseils sur les aspects juridiques, analyse juridique des projets de traitements des données, données à caractère personnel etc.
- **Les formateurs IST** : Urfist, formateurs des établissements : sensibilisation, formation.

L'écosystème Recherche Data Gouv

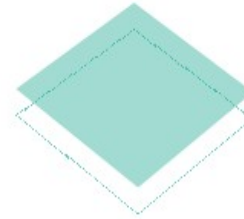


L'écosystème Recherche Data Gouv



Centres de ressources

 Recherche Data Gov



Centre de ressources
entrepôt-catalogue



Centre de ressources
supports pédagogiques, e-formations



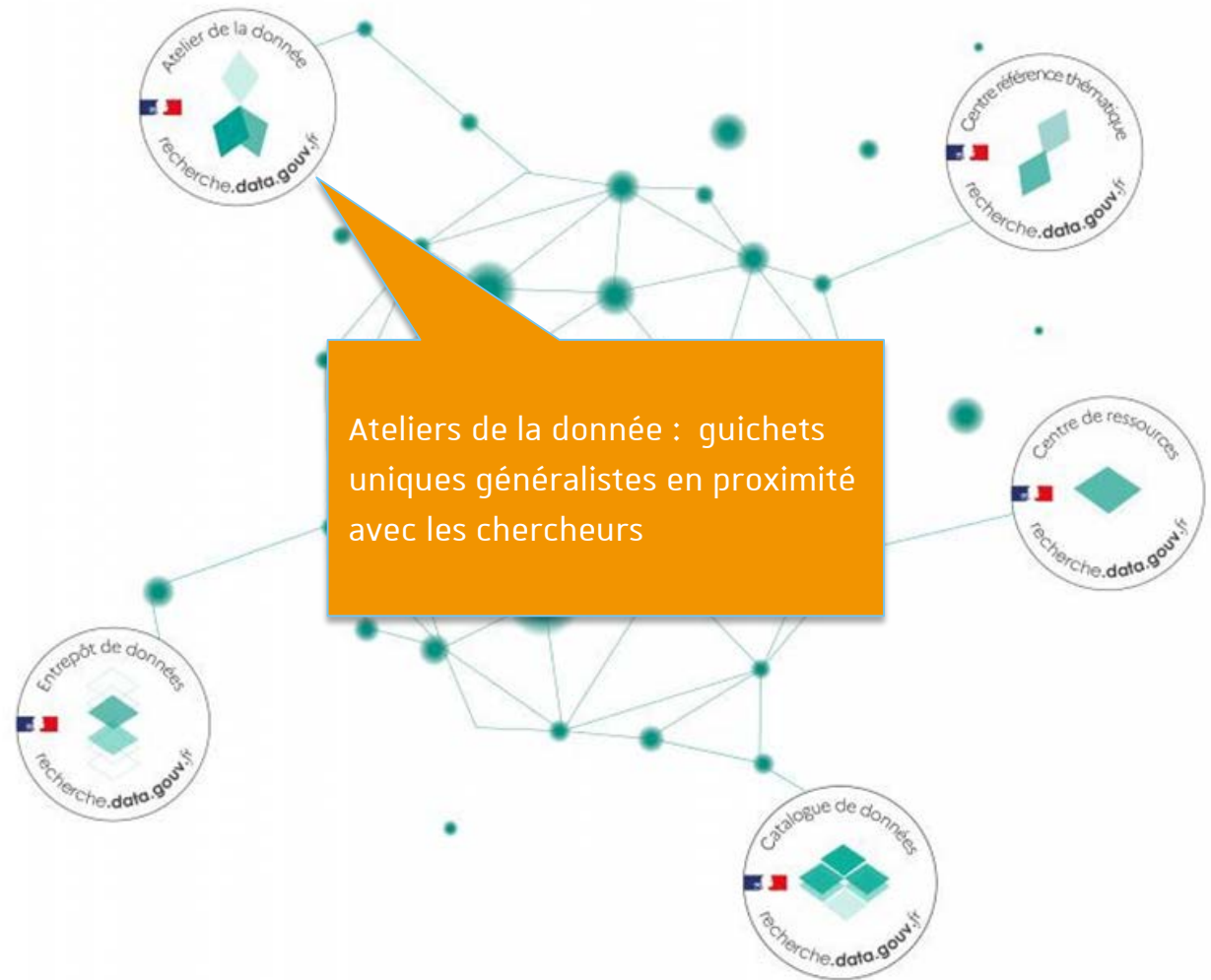
Centre de ressources
compétences



Centre de ressources
outils communs aux ateliers

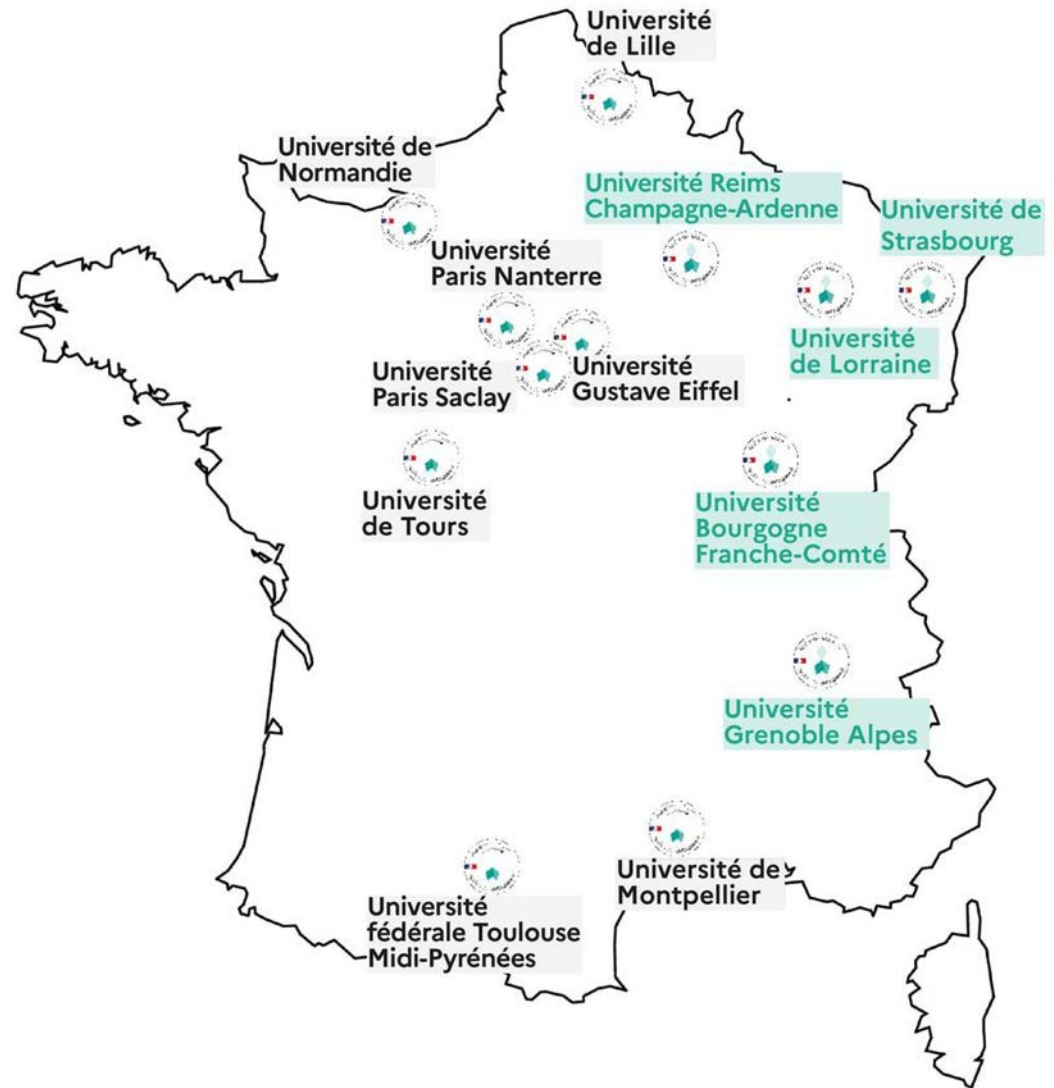
Ateliers de la donnée

 Recherche Data Gov



Ateliers de la donnée

 Recherche Data Gouv



L'atelier de la donnée Alsace

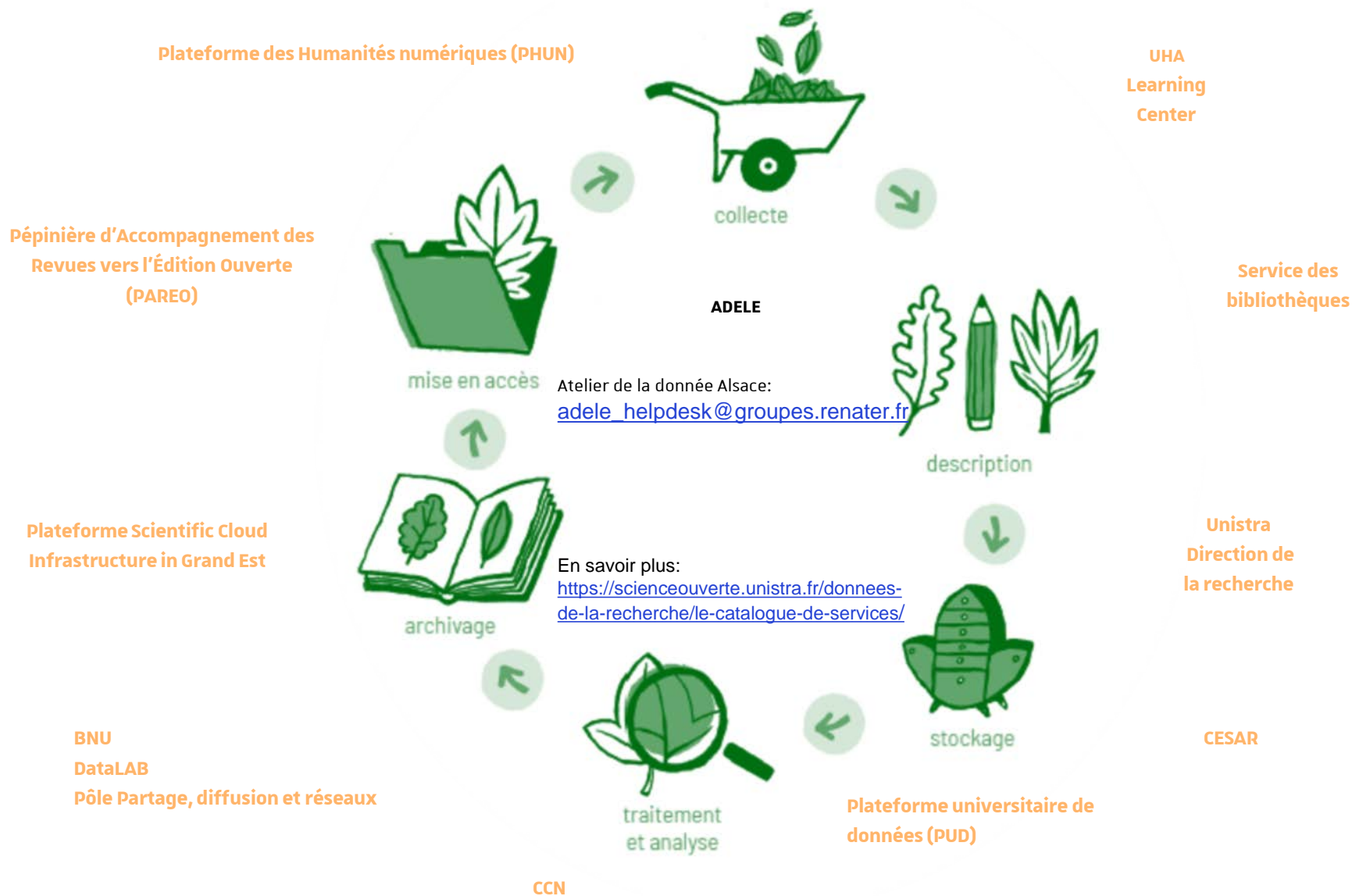
ADELE Helpdesk



5 établissements alsaciens



Atelier de la donnée : un guichet unique



Formations des Urfist

Réseau URFIST

Programme à venir de votre URFIST | Formations du réseau | Formateurs | Partenaires | 0 | FAQ

Bienvenue sur le site SYGEFOR du Réseau des URFIST !

Les URFIST (Unité Régionale de Formation à l'Information Scientifique et Technique) ont pour missions la recherche et la formation des usagers universitaires sur les outils, les pratiques et les évolutions de l'information scientifique.

[Accéder aux formations](#) | [Accéder à mon espace](#)

URFIST DE STRASBOURG

Université de Strasbourg
47 avenue de la Forêt-Noire
67082 STRASBOURG CEDEX
☎ 03 68 85 08 00
✉ urfist@unistra.fr
🌐 <http://urfist.unistra.fr>

[Voir le programme de l'URFIST de Strasbourg](#)

ARU ASSOCIATION RESEAU URFIST

Map of France with a red dot in the Alsace region.

- Collection [*Passeport pour la science ouverte*](#)



Pour aller plus loin

- [*Partager les données liées aux publications scientifiques – Guide pour les chercheurs*](#)



- *Autres documents sur* [*Ouvrir la science*](#)

Sur le web

- DoRANum : <https://doranum.fr/>



- Parcours interactif sur la gestion des données de la recherche

DOI : [10.13143/3xnz-as06](https://doi.org/10.13143/3xnz-as06)

- DEBOIN, Marie-Claude. S'initier en ligne aux données de la recherche. Dans : *CoopIST - Coopérer en information scientifique et technique* [en ligne]. 23 janvier 2020. <https://coop-ist.cirad.fr/gerer-des-donnees/s-initier-en-ligne-aux-donnees-de-la-recherche/1-familiarisez-vous-avec-le-concept-de-donnees-de-la-recherche>