

# Les données de la recherche en accès ouvert : évolutions/révolutions

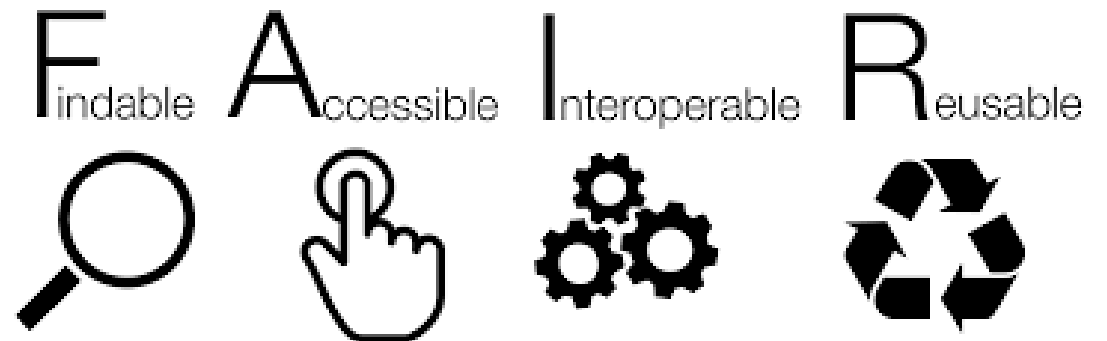
## Archives traditionnelles du chercheurs :

Des fonds (gros volume de doc papiers, enregistrements, photos, + des données numérisées) reversés à posteriori par le chercheur

- Des délais de communication des données parfois très long
- Des données perdus
- Un travail énorme de traitement, dépouillement, classement, indexation des données par les prof de l'IST pour rendre les données accessibles exploitables, partageables. Fonds des médiathèques, phonothèques, iconothèques.

Demain : Les archives de chercheur construites au fil de l'eau, nativement numériques et immédiatement FAIR

Un changement de pratiques (production, gestion des données). Pour le chercheur une culture de la donnée à acquérir.



# Les Données de la rech en Science Ouverte

Données : Ensemble d'info collectées, observées ou créées sous une forme numérique ou non, par les chercheurs dans le cadre d'un projet de recherche et à partir desquelles ils bâtissent leurs hypothèses. **Les données primaires ou brutes** (empiriques, observées, photographiées, mesurées, cartographiées, ) et **les métadonnées** qui décrivent structurent et facilitent l'accès à ces données primaires.

Exemples de données pour les Shs : Données d'enquêtes , de fouilles, d'archives privées, des données cartographiées...

# Les Données de la recherche en Science Ouverte

Depuis **juillet 2018** : Dans la continuité du mouvement qui a porté les publications vers l'Open access (~ 38 % des publi sc sont désormais en Open Access) , les données de la recherche sont sous le feu des projecteurs

Volonté forte du gvt français ainsi que 16 autres pays (dont les organismes de financement de la recherche (dont ANR, ERC) ont signé le Plan S que ces gisements énormes de connaissances, trop peu connus et exploités, est même souvent voué à l'oubli et à la destruction, faute de signalisation et de sauvegarde pérenne **soit en accès ouvert selon les principes du FAIR.**

**Projet ambitieux** , une montagne de données très diversifiées construite de façon « anarchique ». Mettre de l'ordre dans tt ça. **Définir de bonnes pratiques** = le **principe de données FAIR**, mettre en place des **infrastructures** pour accompagner les chercheurs dans le tournant numérique qui concerne aussi leurs données de recherche. En Shs infrastructures TGIR Huma-Num , PROGEDO

# Les textes fondateurs

France

- La loi pour la république numérique(2016)
- **Plan national pour la science ouverte (juillet 2018)**

**PREMIER AXE : GÉNÉRALISER L'ACCÈS OUVERT AUX PUBLICATIONS**

## **MESURES**

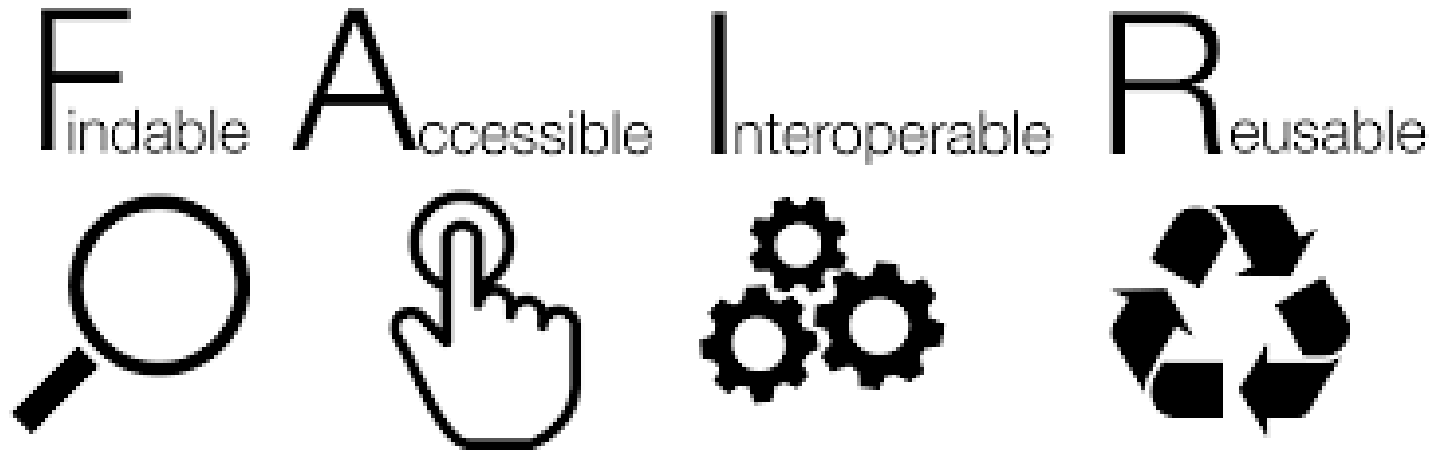
1. Rendre obligatoire la publication en accès ouvert des articles et livres issus de recherches financées par appel d'offres sur fonds publics.
2. Créer un fonds pour la science ouverte.
3. Soutenir l'archive ouverte nationale HAL et simplifier le dépôt par les chercheurs qui publient en accès ouvert sur d'autres plateformes dans le monde.
4. Reconnaître la science ouverte dans les évaluations des chercheurs et des établissements

## **DEUXIÈME AXE : STRUCTURER ET OUVRIR LES DONNÉES DE LA RECHERCHE**

L'ambition est de faire en sorte que les données produites par la recherche publique française soient progressivement structurées en conformité avec les principes FAIR (Facile à trouver, Accessible, Interopérable, Réutilisable)

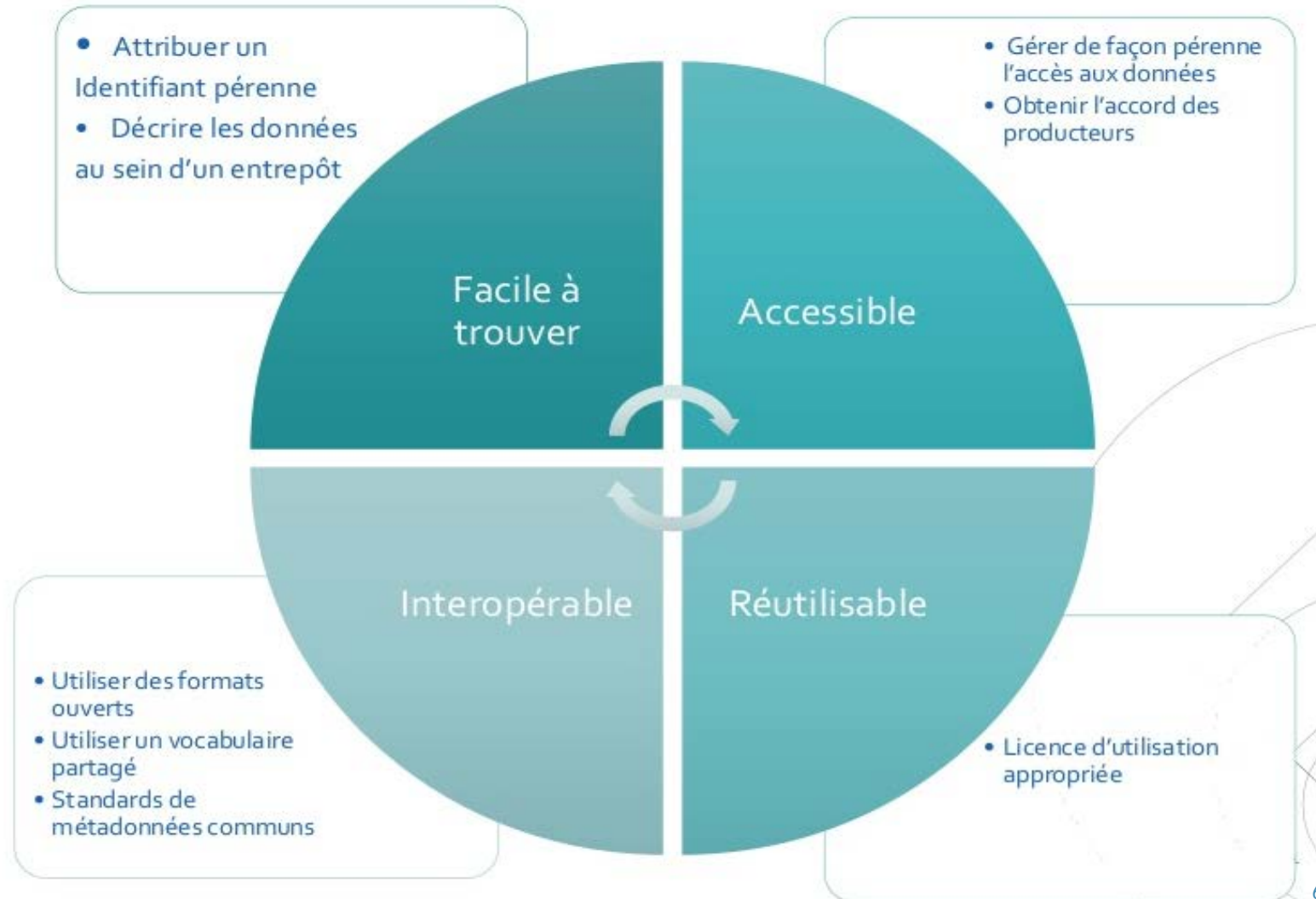
Les chercheurs seront invités à déposer les données dans des entrepôts de données certifiés, dont la gouvernance et les règles de propriété intellectuelle seront conformes aux bonnes pratiques et principes FAIR

# FAIR Principles



Écho aux principes fondamentaux à l'origine de la SO pour les publications : Accessibilité, Pérennité, Intéropérabilité (des formats/protocoles de données intéropérables), Gratuité de l'accès, Utilisation et réutilisation des contenus dans le respect des droits d'auteur

# Principes FAIR : Augmenter le potentiel des données



# Quelques mesures concrètes autour de l'axe 2 du plan national pour la science ouverte

1. Rendre obligatoire la diffusion ouverte des données de recherche issues de programmes financés par appels à projets sur fonds publics. **Généraliser dès 2019 la mise en place de plans de gestion des données dans les appels à projets de recherche.**

# Obligation qui est aussi une exigence des financeurs membre de la coalition S

- ▶ DMP ou PGD plan de gestion de données = **un livrable obligatoire** des appels à projet. Un livrable à fournir dans les 6 mois qui suivent le démarrage du projet.
- ▶ Exemples de l'ANR, de l'ERC
  - ❑ [La politique science ouverte de l'ANR dans les appels à projets](#)
  - ❑ [Le modèle ANR de Plan de Gestion des Données](#)





# Modèle de Plan de gestion des données (PGD)

# Quelques mesures concrètes autour de l'axe 2 du plan national pour la science ouverte

2. Créer la fonction d'administrateur des données et le réseau associé au sein des établissements. Construire autour de l'administrateur des données un réseau de correspondants dans les établissements, pour répondre aux questions que se posent les chercheurs sur les données de la recherche.

# Pourquoi rendre mes données FAIR et les partager ?

Quelques avantages/bénéfices pour le chercheur

- **La garantie d'un archivage à long terme et d'un accès pérenne à mes données vs archivage privé/individuelle = Données en danger ! Perte avec les années ~ 17 % par an (Pierre Corvol Collège de France)**
- **Améliorer la visibilité de mes travaux et de mon profil chercheur.**  
L'interopérabilité permet des passerelles entre données et publications. Mes données pointent sur mes articles, mes articles sont plus lus et plus cités
- **Avoir des statistiques d'usage de mes données**
- **Une occasion/opportunité de me former à « la culture de la gestion de ses données », me former à de nouvelles pratiques ,de nouvelles connaissances, aux évolution techniques de traitement et de gestion des données (cf accompagnement formation des chercheurs à la gestion des données à AMU).**
- **Renforcer mes possibilités de recevoir des financements.** Des données FAIR sont compatibles avec les exigences des financeurs (les membres de la coalition <sup>11</sup> dont ANR,ERC)

# Comment faire pour produire des données FAIR ?

## Comment faire pour concevoir un PGD/MDP ?

Questions techniques : Comment faire pour produire des données bien décrites, bien documentés, dans un format réutilisable/interopérable ?

Questions juridiques : Quels droits s'appliquent à mes données ?

L'accompagnement individuelle , les actions d'info et de formation collectives à venir très prochainement à AMU.

- Des interlocuteurs, personnes ressources à AMU
- Évènements d'info et de formation à la gestion des données à venir très prochainement à AMU Mmsh

- Quelques personnes ressources interlocuteurs à AMU pour m'aider ?

Caroline-Sophie DONATI , chargée de mission Données de la recherche,  
Service commun de la documentation - Cellule Open Access (basée à St Charles).

[caroline-sophie.donati@univ-amu.fr](mailto:caroline-sophie.donati@univ-amu.fr)

<https://bu.univ-amu.libguides.com/donneesrecherche>

À la Mmsh :

Clément De Belsunce (ingénieur d'études) : **Plateforme universitaire de données d'Aix-Marseille Université**

La PUD-AMU, implantée à la Mmsh. Soutenue par le programme LabexMed, la PUD-AMU, développée pour les études en sciences humaines et sociales dans l'aire méditerranéenne. **Spécialisation en données quantitatives.** Elle accompagne la communauté scientifique dans la recherche et l'utilisation des données quantitatives, leur traitement logiciel et leur analyse.

Elle est animée par Clément De Belsunce (ingénieur d'études, statisticien) et pilotée par [Isabelle Blöss-Widmer](#) (référente scientifique).

Contact : [pud-amu@univ-amu.fr](mailto:pud-amu@univ-amu.fr)

# Actions d'information et de formation d'AMU à la gestion des données de la recherche

<http://corist-shs.cnrs.fr/node/945>

Journée 9:30 - 16:30 La matinée sera consacrée à l'intervention de la plateforme universitaire de données de l'université d'Aix-Marseille (PUD-AMU) et l'après-midi verra l'intervention de l'équipe PROGEDO-ADISP sur l'accès à la documentation des données en sciences sociales ([programme détaillé et inscription pour le buffet](#)). Occasion de rencontrer les interlocuteurs clefs d'AMU



## Tour National des Données de la Recherche

**Mardi 12 novembre 2019**

Amphithéâtre, Maison méditerranéenne des sciences de l'homme,  
5 rue du Château de l'horloge, Aix en Provence,  
Entrée libre

# Actions d'information et de formation à la gestion des données de la rech.

Très Grande Infrastructure de Recherche TGIR **PROGEDO** et le réseau national des Msh RnMsh organisent une semaine de formation et d'information autour de l'utilisation des données en sciences humaines et sociales

Du 9 au 13 déc

Au sein des Msh qui mobilisent leurs PUD Plateformes Univ de Données

(TGIR), PROGEDO acronyme de PROduction et GESTion des DONnées a pour but de développer la culture des données, d'impulser et structurer une politique des données d'enquêtes pour la recherche en sciences sociales.



## En conclusion

Pour le chercheur les données de la rech en Science ouvertes = de nouvelles **exigences** mais aussi de **nouvelles opportunités**.

Acquérir une culture de la donnée, se former à de nouvelles méthodologies, aux nouvelles techniques de production et gestion des données, aux questions juridiques (la RGDP).