

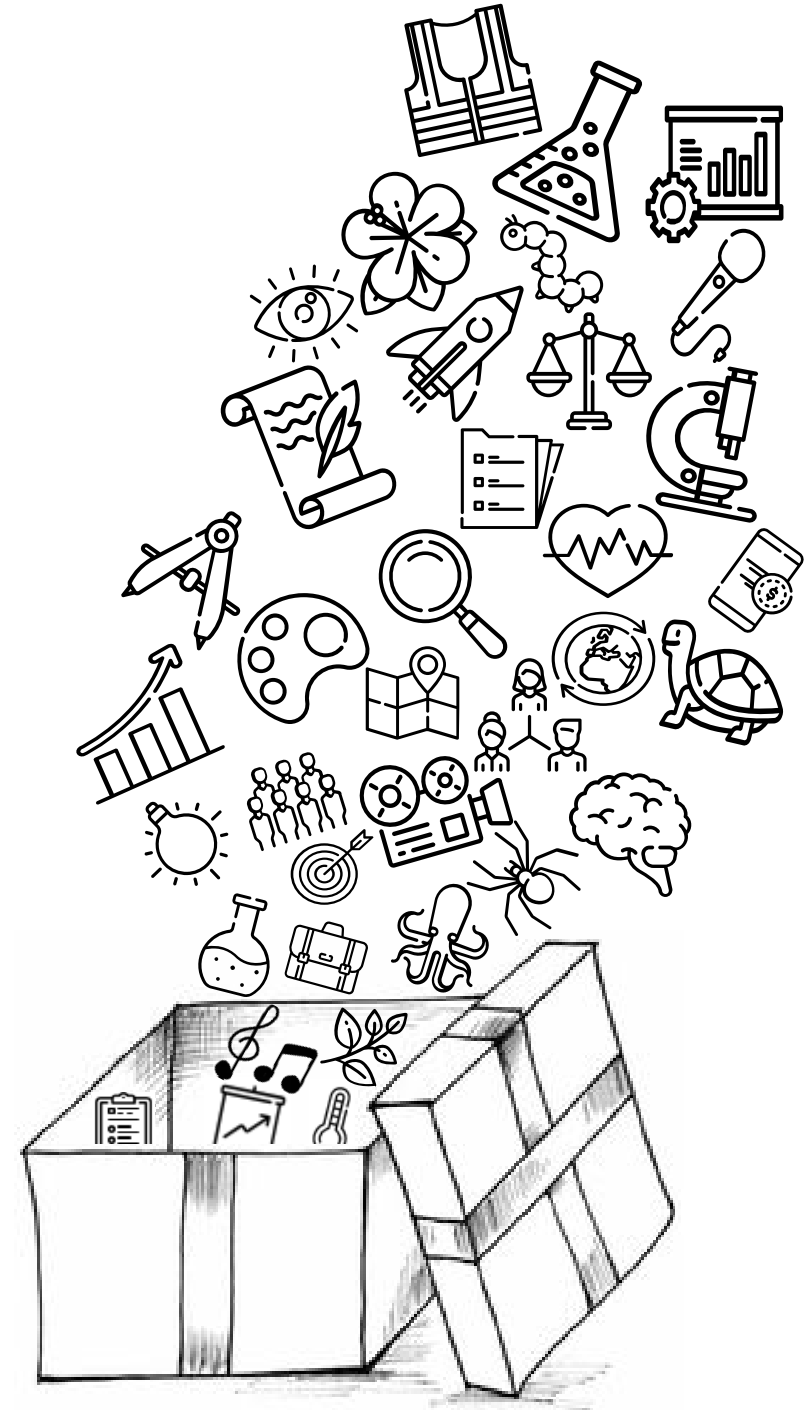
Traiter et analyser des données en sciences humaines et sociales

Échanges autour de l'ouverture des données de la recherche

démarrer, comprendre, se former

11 décembre 2019

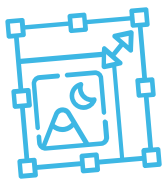
Caroline Sophie DONATI - SCD AMU



Qu'appelle-t-on données de la recherche ?

- Informations collectées, observées ou créées sous une forme numérique ou non, dans le cadre d'un projet de recherche et à partir desquelles sont bâties les hypothèses.
- Produits de la recherche, éléments factuels dont la communauté scientifique admet la nécessité pour documenter, valider ou invalider les résultats de recherche.
- Concernent les producteurs de données **ET** les métiers qui viennent en appui à la recherche.

⇒ Données issues de l'observation, données expérimentales, données dérivées ou compilées, données de références, données de simulation



Que veut dire « ouvrir ses données » ?

- Mettre en libre accès les données issues d'un projet de recherche.

« mise à disposition gratuite sur l'Internet public de données, permettant à tout un chacun de lire, télécharger, copier, transmettre, [...] ou s'en servir à toute autre fin légale, sans barrière financière, légale ou technique autre que celles indissociables de l'accès et l'utilisation d'Internet »
(Initiative de Budapest pour l'Accès Ouvert, 2002)



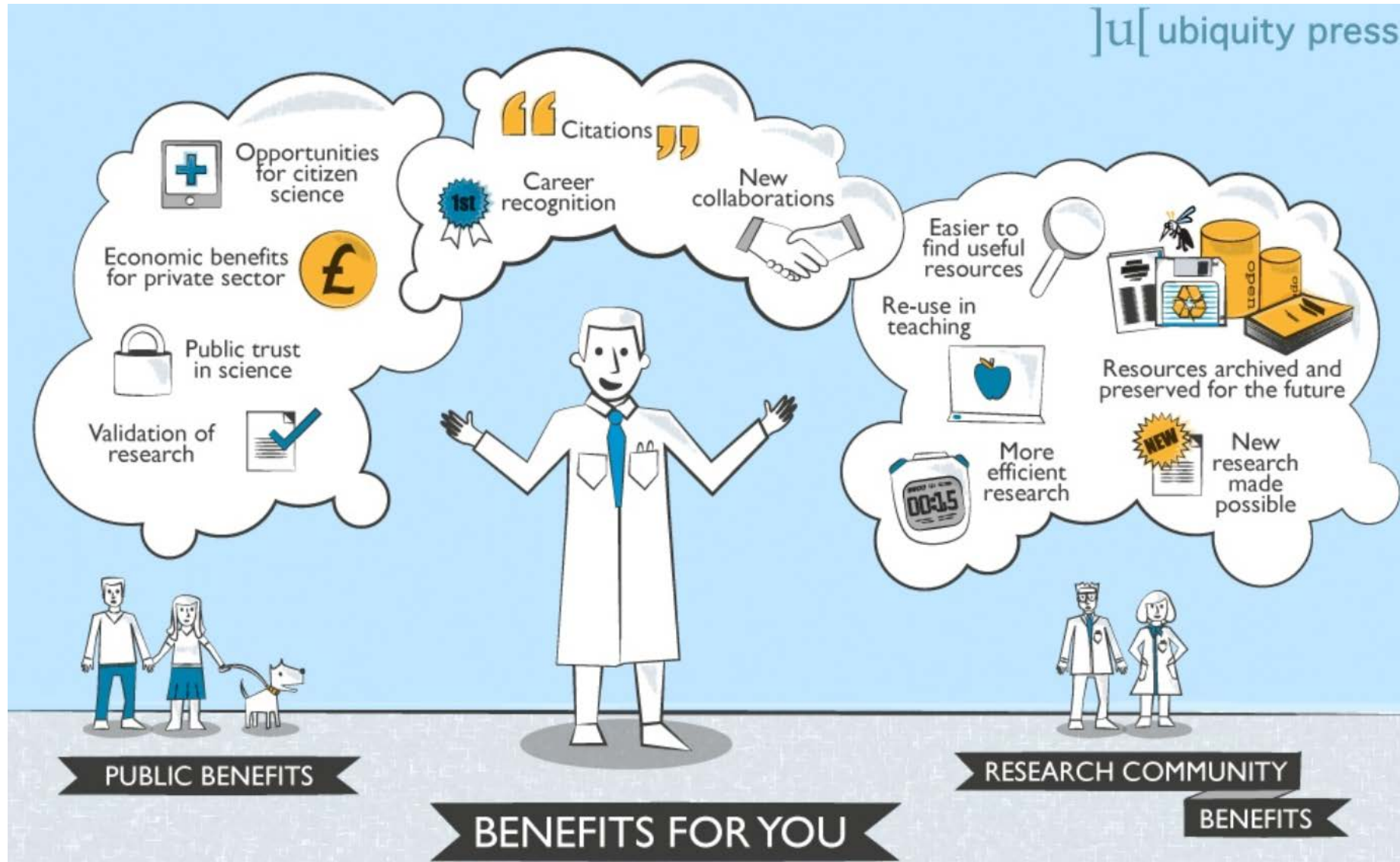
Attention

Ouvrir l'accès ne signifie pas nécessairement rendre accessible **toutes** les données de recherche.

Alors pourquoi ouvrir – réfléchir à l'ouverture des données ?

- Répondre aux exigences des organismes subventionnaires et respecter la loi ;
- Permettre la reproduction et la validation des résultats de la recherche ;
- Aller vers la transparence et la vérifiabilité des données ;
- Favoriser la réutilisation pour le développement de nouveaux savoirs ;
- Éviter la répétition des travaux de recherche ;
- Lutter contre la fraude scientifique et le plagiat ;
- Respecter les règles éthiques.

Quels bénéfices ?



From Open Access to Open Data, <https://fr.slideshare.net/brianhole/from-open-access-to-open-data>, Publié dans *Technologie*. Licence : CC Attribution License.

Qu'est-ce que la gestion des données ?

- Ensemble des activités mises en œuvre **pour faciliter l'exploitation et la sécurisation** des données pendant un projet de recherche, et leur **préservation** après la fin du projet.
- De bonnes pratiques de gestion sont un préalable **indispensable** à l'ouverture des données et à leur réutilisation, pour :
 - accompagner vos données tout au long du projet de recherche ;
 - les organiser dès la collecte ;
 - sauvegarder régulièrement les données en cours d'utilisation ;
 - préserver/archiver les données **achevées** et rendre leur lecture accessible, leur exploitation future possible ;
 - respecter les règles éthiques pour leur partage.

Comment faire ?

1. Penser gestion des données dès la soumission du projet de recherche

Quelles données ?

Quels besoins d'analyse/traitement ?

Avec qui ? Comment ?

Quel coût ?

...

- ↳ Fiabiliser les données et faciliter leur gestion ;
- ↳ Anticiper et favoriser leur éventuelle diffusion ;
- ↳ Décrire la façon dont les données scientifiques d'un projet de recherche seront **produites, traitées, diffusées, protégées...**

2. Structurer et conserver les données selon les principes FAIR

- ↳ Les données FAIR ne sont pas nécessairement ouvertes
- ↳ Les principes FAIR permettent, **si les données sont communicables**, de poser les bonnes bases pour leur ouverture, selon le principe de la Commission Européenne

« aussi ouvert que possible, aussi fermé que nécessaire ».

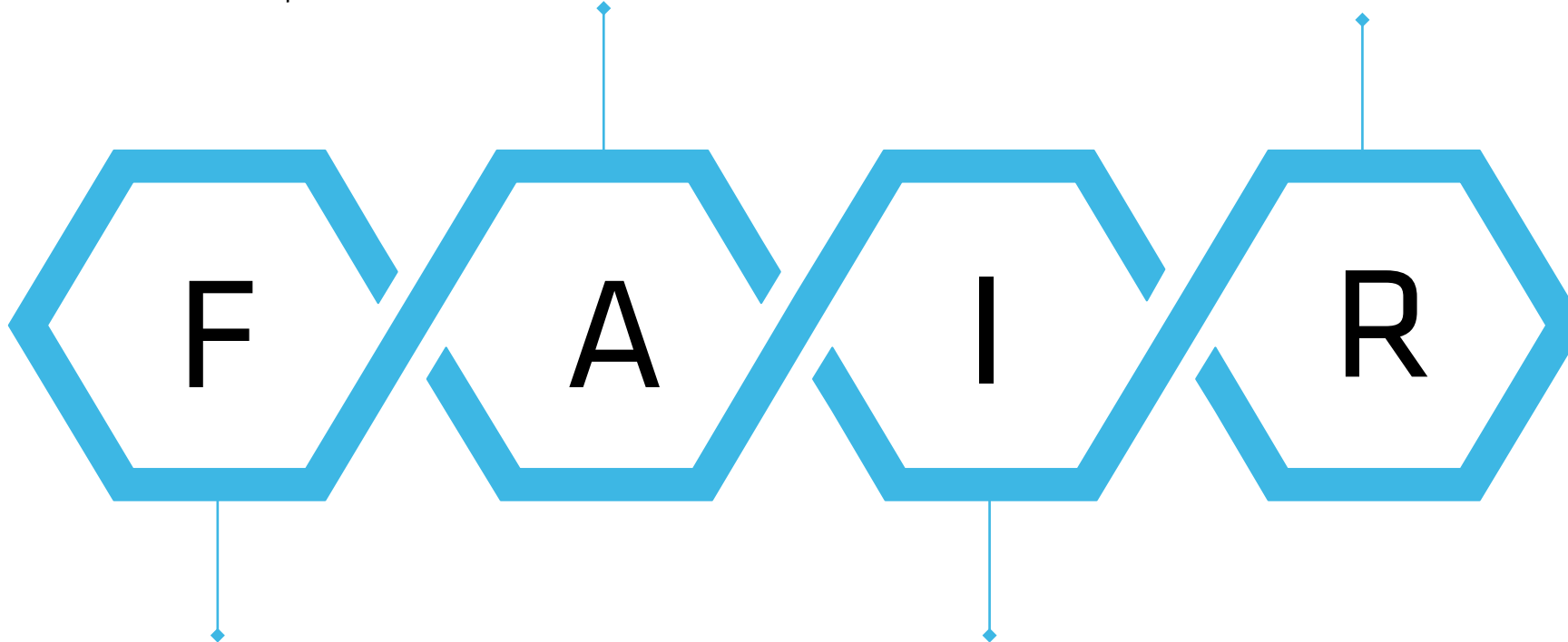
3. Ensuite, prendre le temps de rendre la science plus ouverte, plus transparente et d'accès plus universel.

Accessible

- Encourager à stocker durablement les données et les métadonnées
- Faciliter leur accès et/ou leur téléchargement
- Spécifier les conditions d'accès et d'utilisation

Réutilisable

- Mettre en avant les caractéristiques qui rendent les données réutilisables pour de futures recherches ou d'autres finalités



Facile à trouver

- Faciliter la découverte des données par les humains et les systèmes informatiques
- Décrire et indexer des données et des métadonnées

Interopérable

- Téléchargeable
- Utilisable
- Intelligible
- Combinable
 - avec d'autres données, par des humains et des machines

FAIR ≠ OPEN

- Tout ce qui est FAIR n'est pas forcément Open
 - **Notions communes** : l'interopérabilité et l'accessibilité
 - **Différences** : considérations juridiques, éthiques ou économiques / restriction de l'accès à des données qui respectent par ailleurs les principes techniques FAIR

FAIR

- la prise en compte de questions juridiques et éthiques (quelles données communiquer et comment) ;
- l'adoption de bonnes pratiques de gestion des données (formats ouverts, vocabulaires et standards de métadonnées communs) ;
- l'utilisation d'outils et de technologies adéquats

OPEN

- contenu librement accessible à l'utilisation, la modification et le partage par quiconque pour quelque fin que ce soit.
- contenu mis à disposition sous une licence ouverte et sans restrictions aux réutilisateurs au-delà de celles qui préservent la provenance ou l'ouverture.

Dans la pratique

L'obligation : le plan de gestion de données

↳ détaille la façon dont vous obtenez, documentez, analysez et utilisez les données à la fois au cours de votre recherche et une fois le projet terminé.

↳ décrit les méthodes et les processus de création, fourniture, maintenance, conservation et de protection des données.

- Mettre de l'ordre dans ses données via une démarche qualité, une politique d'archivage, de tri...
- FAIR = prendre le temps / ne pas s'inquiéter / aller selon l'agenda du projet, du chercheur, du service d'appui

Pour vous accompagner

- AMU s'engage dans la science ouverte
 - HAL AMU
 - Données de la recherche
- Service commun de la documentation > cellule Open Access
 - archiveouverte@univ-amu.fr
 - caroline-sophie.donati@univ-amu.fr
- Le cycle de la donnée doit être présenté au plus tôt : dans les projets / dans les carrières.
 - ↳ Accompagnement à la rédaction de plan de gestion des données / data management plan
 - ↳ Sensibilisation à la gestion des données de la recherche (rencontres dans les unités)
 - ↳ Formation doctorale (en cours)
 - ↳ Formations thématiques EC/IST (en cours)

Ressources

- Espace institutionnel AMU sur [DMP OPIDoR](https://dmp.opidor.fr/) : <https://dmp.opidor.fr/>
 - Initiative portée par l'Inist et le CNRS
 - Outil d'aide à la rédaction de plans de gestion des données (PGD) / data management plan (DMP)
 - Accessible à la communauté scientifique de l'ESR
 - Des modèles répondant aux exigences d'organismes de financement et d'institutions
 - Collaboratif et personnalisable : il facilite les échanges entre les partenaires d'un même projet et les services d'accompagnement
- Un [guide](https://bu.univ-amu.libguides.com/donneesrecherche) dédié aux données de la recherche : <https://bu.univ-amu.libguides.com/donneesrecherche>
- Une [liste de diffusion, d'échanges](mailto:scd-donnees-recherche@univ-amu.fr) : scd-donnees-recherche@univ-amu.fr

Traiter et analyser des données en sciences humaines et sociales

Merci

Des questions ?

