

Réaliser un plan de gestion de données FAIR

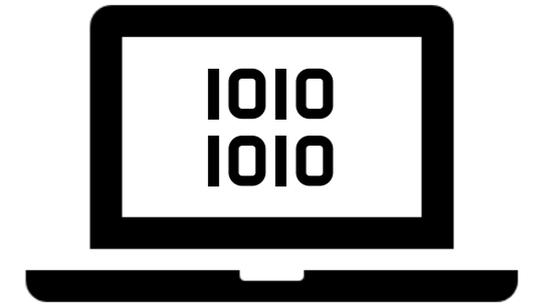
—
7 décembre 2021

 **Université
Paris Nanterre**



Nathalie Reymonet
Chargée de mission
science ouverte - UPL

Données de la recherche



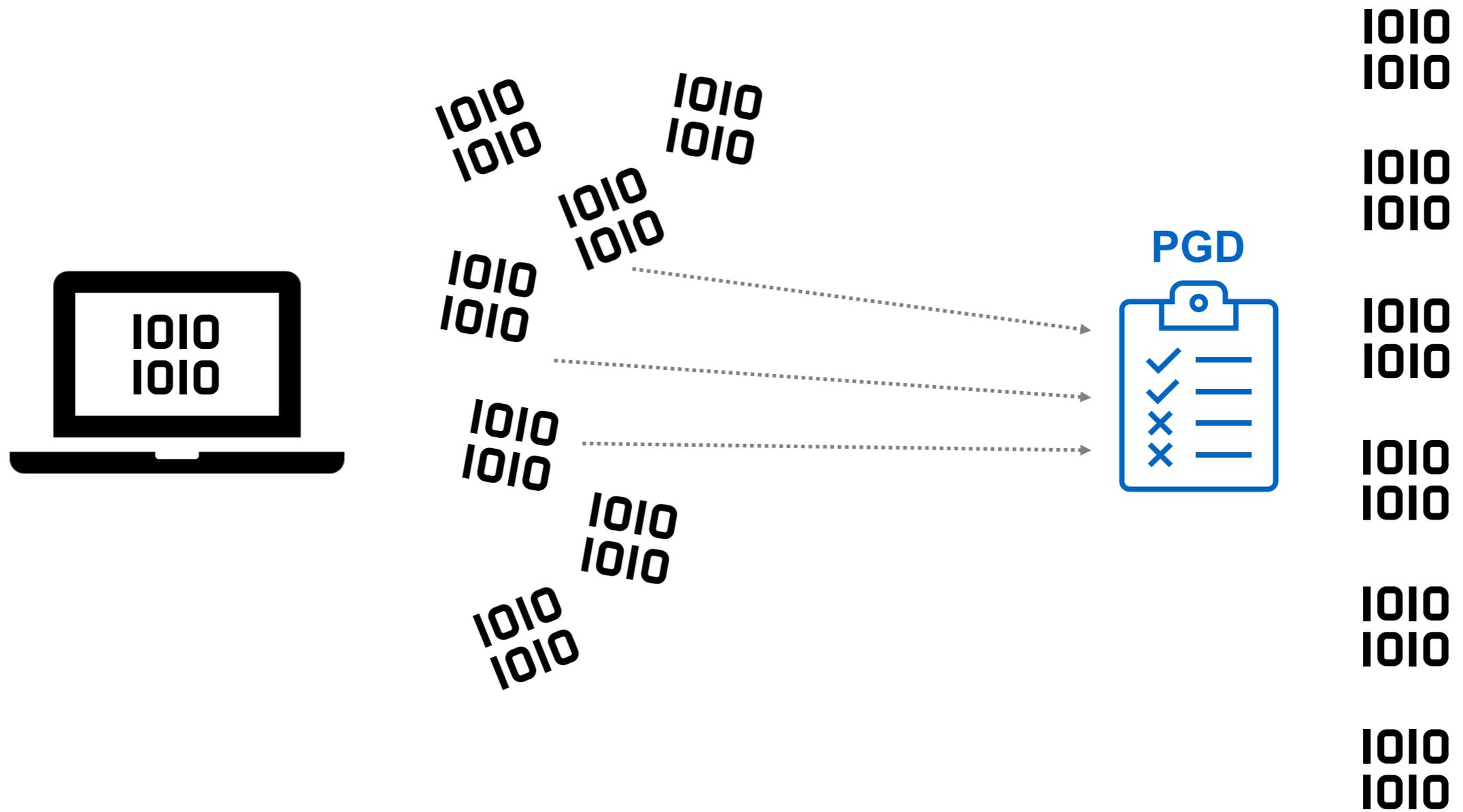
Les copies ou reproductions numériques réalisées à partir d'une source licite, en vue de l'exploration de textes et de données incluses ou associées aux écrits scientifiques pour les besoins de la recherche publique, à l'exclusion de toute finalité commerciale.

... **fichiers produits au terme des activités de recherche** pour lesquelles elles ont été produites ; **ces fichiers constituent des données de la recherche**

(LPRN, Art. 38 = Code de la PI, Art 122-5)

Réaliser un plan de gestion des données... mais pourquoi ?

—



Plan de gestion de données



Document formel précisant la manière dont les données seront produites, traitées, décrites, partagées ou protégées et conservées au cours et à l'issue du projet.

Ses objectifs

- décrire le cycle de vie des données produites ou collectées au cours du projet de recherche.
- anticiper les questions de gestion qui surviennent au cours d'une recherche et les conditions d'une diffusion et d'une conservation futures des données.

Enjeux de l'ouverture des données

—



Valoriser la production scientifique de l'établissement



Conformité aux exigences réglementaires



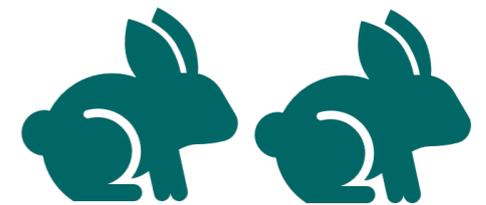
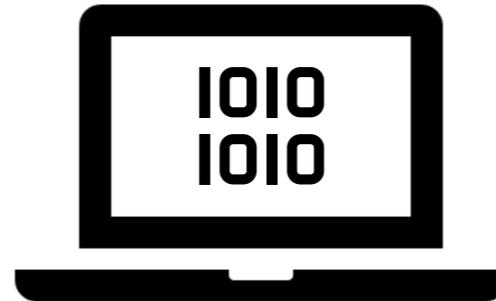
Nouvelles sources pour la recherche



Fouille de données / *data mining*



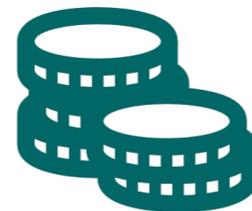
Confiance science/société



Préserver les données dont les conditions de production ne peuvent pas être reproduites

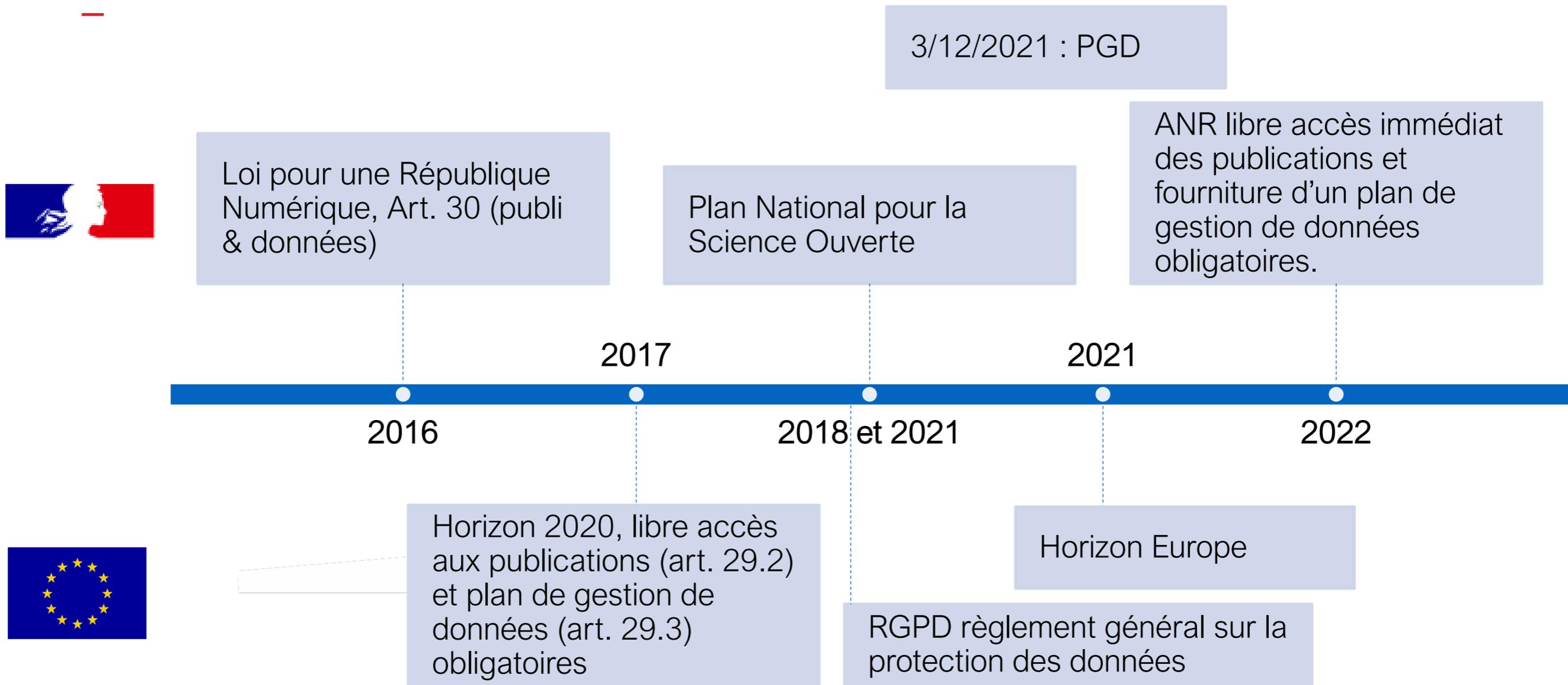


Fiabiliser les résultats de recherche



Economies dans les coûts de production

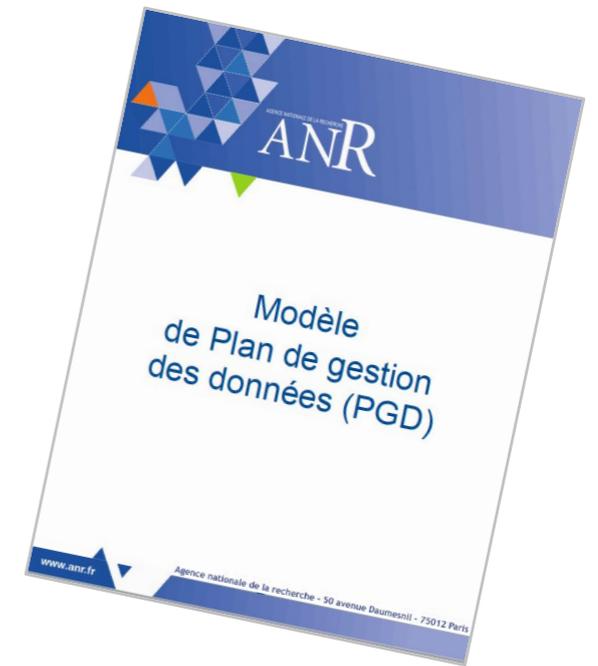
Contexte récent



Mieux gérer ses données : le plan de gestion de données



Exemple de structure : le PGD de l'ANR



1. Description des données



- a) Modalités de collecte
- b) Types de données

2. Documentation et qualité des données



- a) Métadonnées et documentation
- b) Contrôle de la qualité des données

3. Stockage, sauvegarde



- a) Stockage et sauvegarde
- b) Sécurité des données

4. Exigences légales et éthiques



5. Partage et conservation à long terme



- a) Partage
- b) Conservation
- c) Méthodes et outils pr lecture
- d) Identifiants pérennes



6. Responsabilités en matière de gestion des données



Les acteurs du PGD



Chercheur et ingénieur



Informaticien

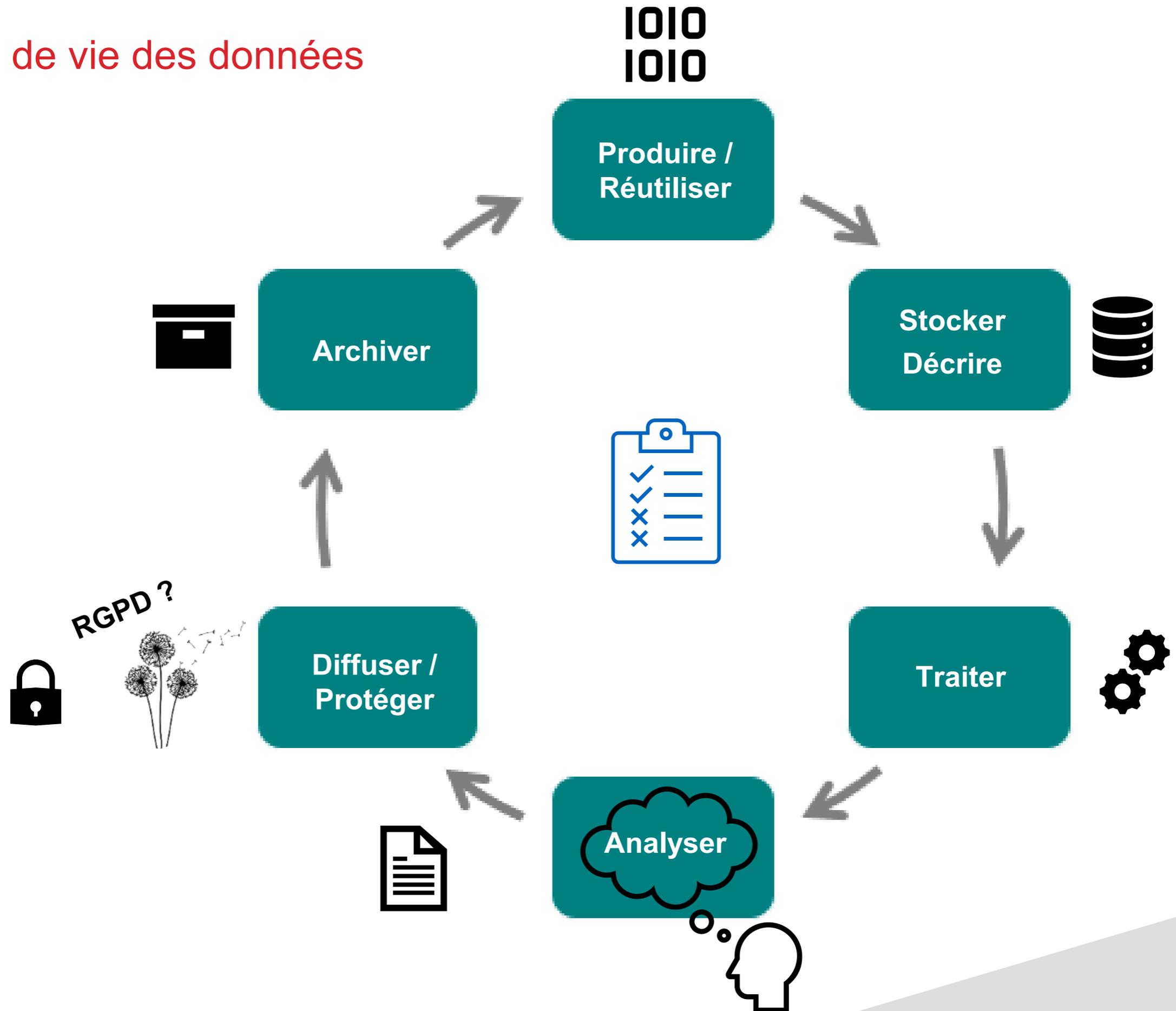


Professionnel de l'IST



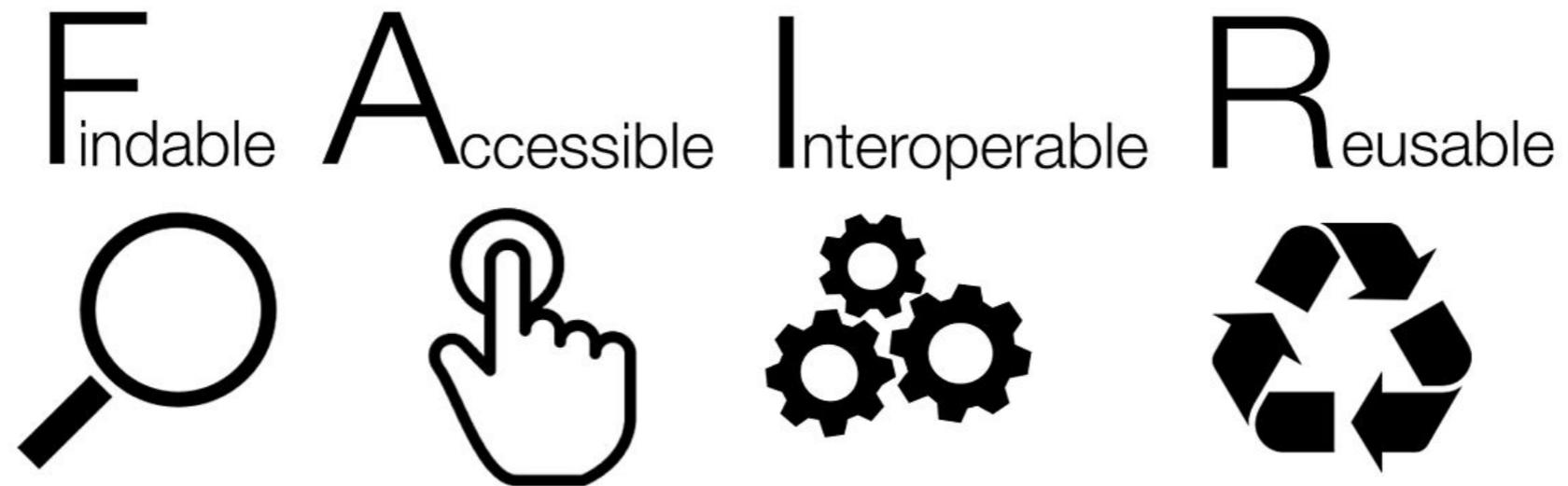
DPO, juriste, ingénieur-projet

Cycle de vie des données



Données FAIR

—



Sangya Pundir (2016) under the [Creative Commons Attribution-Share Alike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/) license.





À quoi ça sert ?

zenodo Search Upload Communities Log in Sign up

All versions Found 1267329 results. Sort by: Most recent asc. View

Access Right

- Open (1241927)
- Closed (23533)
- Restricted (1176)
- Embargoed (693)

File Type

- Pdf (721178)
- Jpg (294933)
- Png (152984)
- Zip (46076)
- Hdf5 (15048)
- Xml (9227)
- Docx (6878)
- Txt (4196)
- Json (4048)
- Gz (3572)

May 22, 2019 (v1) Photo Open Access

Sisymbrium chrysanthum Jord. (BR0000010525616)
Meise Botanic Garden;
Belgium Herbarium image of Meise Botanic Garden.
Uploaded on May 22, 2019

May 22, 2019 (v1) Photo Open Access

Sisymbrium chrysanthum Jord. (BR0000010525760)
Meise Botanic Garden;
Belgium Herbarium image of Meise Botanic Garden.
Uploaded on May 22, 2019

May 22, 2019 (v1) Photo Open Access

Sisymbrium chrysanthum Jord. (BR0000010525623)
Meise Botanic Garden;
Belgium Herbarium image of Meise Botanic Garden.
Uploaded on May 22, 2019

Les métadonnées servent à faciliter la recherche d'une donnée : lorsque vous tapez des mots clés dans une barre de recherche ou lorsque vous sélectionnez des filtres.



Pourquoi une authentification ?

Le principe Accessible n'est pas synonyme de libre et ouvert

"As open as possible, as closed as necessary"

Bien que produites sur fonds publics, certaines données ne peuvent pas être accessibles publiquement pour des raisons légitimes. Ce sont par exemple :

- des données à caractère personnel (données permettant d'identifier une personne, directement ou indirectement) ;
- des données relevant de la sécurité nationale ;
- des données sujettes à un dépôt de brevet.

Si vos données doivent rester privées, spécifiez les conditions exactes dans lesquelles elles peuvent être accessibles : qui a le droit d'y accéder et comment.



Qu'est-ce qu'un vocabulaire FAIR ?

Il s'agit de vocabulaires contrôlés dont on peut retrouver la documentation grâce à un PID (identifiant pérenne et unique).

Exemple : Animal Diseases Ontology (ANDO)

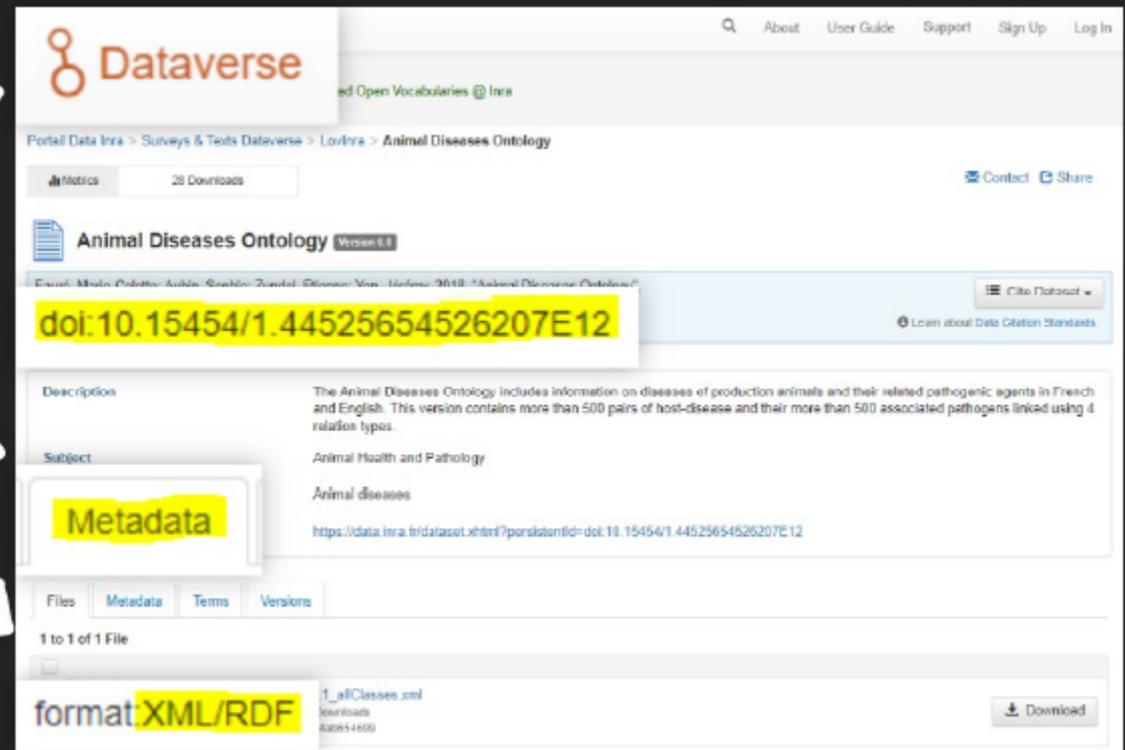
Ce vocabulaire respecte les principes FAIR :

Accessible sur l'entrepôt Dataverse

A son propre identifiant (DOI)

Est décrit par des métadonnées

Est publié dans un format interopérable



<https://data.inra.fr/dataset.xhtml?persistentId=doi:10.15454/1.44525654526207E12>



Pourquoi attribuer une licence ?

En France, les données issues de recherche financée sur **fonds publics** doivent être **ouvertement partagées et librement réutilisables**, sauf exceptions légales (données sensibles par exemple).

Mais même si vos données peuvent être librement réutilisables, il est préférable d'attribuer une licence pour **explicitier vos conditions**. Par exemple, une licence CC-BY exigera au réutilisateur de reconnaître votre paternité.

Caractéristiques de la licence
 Vos choix sur ce panneau vont mettre à jour les autres panneaux sur cette page

Vous souhaitez autoriser le partage des adaptations de votre Oeuvre ?
 ?

Oui Non Oui, sous condition de partage dans les mêmes conditions.

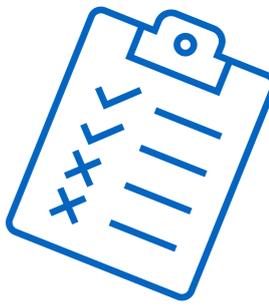
Autorisez-vous les utilisations commerciales de votre œuvre ?
 ?

Oui Non

?
Licence sélectionnée
 Attribution 4.0 International

  ← Reconnaissance de la paternité

C'est une Licence "Free Culture" (libre au sens donné par ce groupe) 



5. Partage des données et conservation à long terme

Données collectées : Données collectées par les membres du programme Transfunéraire

Les données seront partagées à la fin du programme de recherche sur plateforme Didomena, l'entrepôt de données de recherche de l'EHESS. (<https://didomena.ehess.fr>). Cette plateforme est dédiée aux sciences sociales et permet de partager et valoriser les données de recherche.

Il n'existe pas de raison motivant un embargo, les données partagées seront immédiatement accessibles. L'accès aux données sera libre.

Cependant, toutes les données issues des enquêtes qualitatives ne seront pas partagées. La somme de données recueillie implique de réaliser un tri pour identifier les données dont le partage présente le plus d'intérêt scientifique. Le partage d'une partie des données est restreint par leur caractère confidentiel et/ou par les autorisations spécifiques négociées avec les enquêtés.

Données produites : Données produites par les membres, dans le cadre du programme Transfunéraire

Les données seront partagées à la fin du programme de recherche sur plateforme Didomena, l'entrepôt de données de recherche de l'EHESS. (<https://didomena.ehess.fr>). Cette plateforme est dédiée aux sciences sociales et permet de partager et valoriser les données de recherche.

Il n'existe pas de raison motivant un embargo, les données partagées seront immédiatement accessibles. L'accès aux données sera libre.

L'ensemble des données générées seront uniques et devront donc être conservées sans mention de délai par leurs auteurs respectifs.

Toutes les données produites lors des enquêtes qualitatives de terrain ne seront pas partagées. Les données partagées feront l'objet d'une sélection préalable. Le partage des données est restreint par le respect du caractère confidentiel de certaines données et par les autorisations spécifiques négociées avec les enquêtés.

Droits d'auteur

Le(s) créateur(s) de ce plan accepte(nt) que tout ou partie de texte de ce plan soit réutilisé et personnalisé si nécessaire pour un autre plan. Vous n'avez pas besoin de citer le(s) créateur(s) en tant que source. L'utilisation de toute partie de texte de ce plan n'implique pas que le(s) créateur(s) soutien(nen)t ou aient une quelconque relation avec votre projet ou votre soumission.

Rédiger son plan de gestion de données

—

Pour de bonnes pratiques de gestion des données

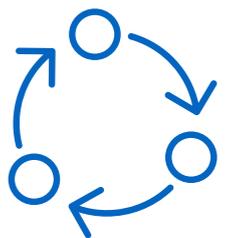
IOIO IOIO
IOIO IOIO
IOIO IOIO
IOIO IOIO

Rassembler dans un document unique, prospectif, évolutif



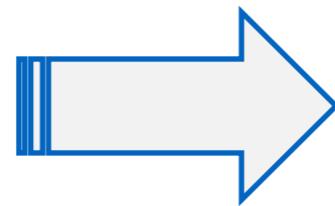
Les contributions des différents intervenants sur les données du projet

Suivant la temporalité du cycle de vie des données



Rédiger son plan de gestion de données

—



Merci de votre attention !

Nathalie Reymonet – nathalie.reymonet@u-plum.fr
nreymonet@parisnanterre.fr