



Formation doctorale



**Comment bien gérer ses données
pour faciliter son travail de doctorat ?**

Cellule Data Grenoble Alpes – 27 avril 2023



- **Module principal (B10) sur 2 jours :**
 - organisation et description : jeudi 27 avril
 - stockage, diffusion et conservation : jeudi 04 mai
- **3 Modules supplémentaires :**
 - diffusion des données (B13) : jeudi 11 mai
 - diffusion des codes (B16) : mardi 23 mai
 - gérer des données à caractère personnel (B14): mardi 30 mai
- **Objectifs :**
 - sensibiliser à la gestion FAIR des données de la recherche
 - apporter des recommandations pratiques
 - accompagner à la rédaction du plan de gestion de données
- **Formation interactive :**
 - des présentations classiques, des sondages, des activités
 - **surtout : posez toutes les questions que vous souhaitez !**

Contact Cellule Data : sos-data@univ-grenoble-alpes.fr



Qu'est-ce que la gestion des données de la recherche ?





- Données d'observation
- Données expérimentales
- Données de simulation numérique
- Données dérivées ou compilées
- Données de référence
- Données d'enquêtes
- Code informatique

- **Ensemble des informations collectées, produites et utilisées par la recherche scientifique**
- **Des données très variées dans leur forme**
- **Très dépendantes des disciplines scientifiques**



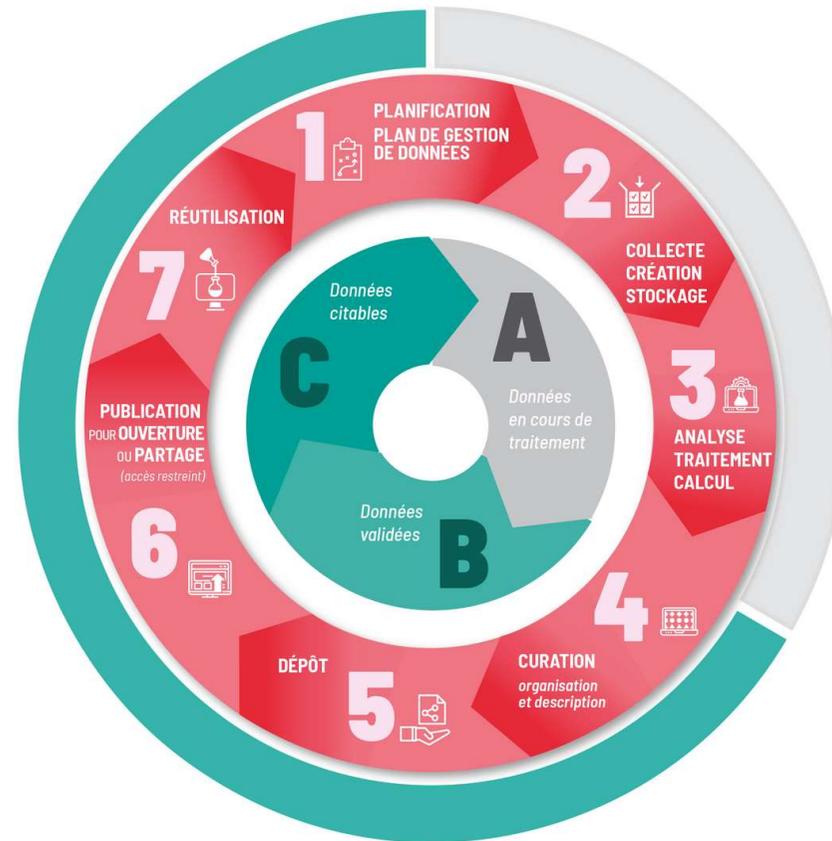
Dans le cadre du *Règlement Général sur la Protection des Données* (RGPD), il faut faire attention à distinguer parmi les données de recherche :

- les données personnelles
- les données personnelles sensibles (données pouvant induire des discriminations)

Cycle de vie des données de la recherche



Cycle de vie des données, Recherche Data Gov



On oublie bien souvent l'étape de **destruction des données**, ainsi que le fait que les données ont un **impact environnemental** !



Par défaut, **l'ouverture des données de la recherche est obligatoire selon le principe "aussi ouvert que possible, aussi fermé que nécessaire"**.

Exceptions :

- Données **personnelles** non anonymisées ou pseudonymisées
- Données **non achevées** (sauf données géographiques qui peuvent être inachevées, directive INSPIRE), certaines données environnementales (cf protection des espèces)
- **Photos** quand il y a une personne reconnaissable
- Données **scrapées** (car souvent interdit par les conditions d'utilisation) sauf :
 - si le préjudice du producteur est nul (risque encouru nul également)
 - le scraping d'une partie non substantielle des données reste possible



- En premier lieu **faciliter le travail de recherche** :
 - Assurer l'**exploitation des données** tout au long du projet
 - Assurer la **conservation** des données à moyen terme
 - Pouvoir **reproduire** les résultats scientifiques, y compris par l'équipe qui les a obtenus !
 - Pouvoir facilement les **réutiliser** pour produire de nouvelles recherches
 - **Valoriser** les résultats scientifiques, mais aussi les données et les codes qui ont permis de les obtenir et en augmenter la visibilité
 - Favoriser de **nouvelles collaborations**, de nouvelles approches
- Enjeux d'**ordre sociétaux** :
 - Assurer une **souveraineté** sur les données produites
 - Assurer l'**intégrité scientifique**
 - Garantir la **transparence**, et assurer la confiance des citoyens en la recherche



Tour de table





Présentez vous en *quelques mots* :

- votre laboratoire et vos thèmes de recherche
- les données manipulées dans votre thèse :
 - type
 - produite ou réutilisée
 - volume
- avez-vous une attente spécifique de la formation ?
- est-ce que vous avez déjà une question préalable ?