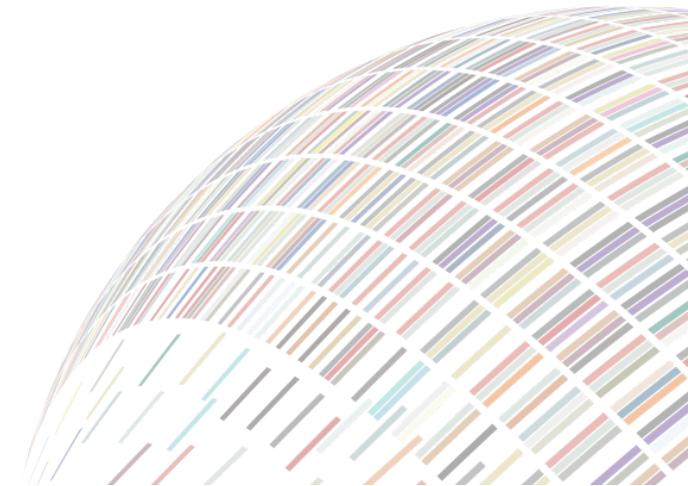


Quetelet Progedo Diffusion à l'heure de la science ouverte

23 novembre 2021

Lorraine Adam (Ingénieure d'étude à l'ADISP - Progedo)

Erik Zolotoukhine (Ingénieur d'étude à l'ADISP - Progedo)



Préambule : les origines de la diffusion de données statistiques

Aux origines... Demandes individuelles des chercheurs auprès de l'INSEE

Années 80 : Chercheurs du **Lasmas** : enquêtes FQP puis enquêtes Emploi
→ Le Lasmas deviendra le **Centre Maurice Halbwachs** (2004)

Années 90 : Diffusion des données (INSEE) étendue au sein du CNRS

Années 2000 : Diffusion "**générale**" via le Réseau Quetelet
→ Champ étendu aux SSM (services stat. ministériels) et autres producteurs "publics"
→ Accès pour l'ensemble de la communauté scientifique (universités, EPST, ...)

2012 : Création de **Progedo**
→ Objectif : **développer la culture des données**
→ Coordination du Réseau Quetelet

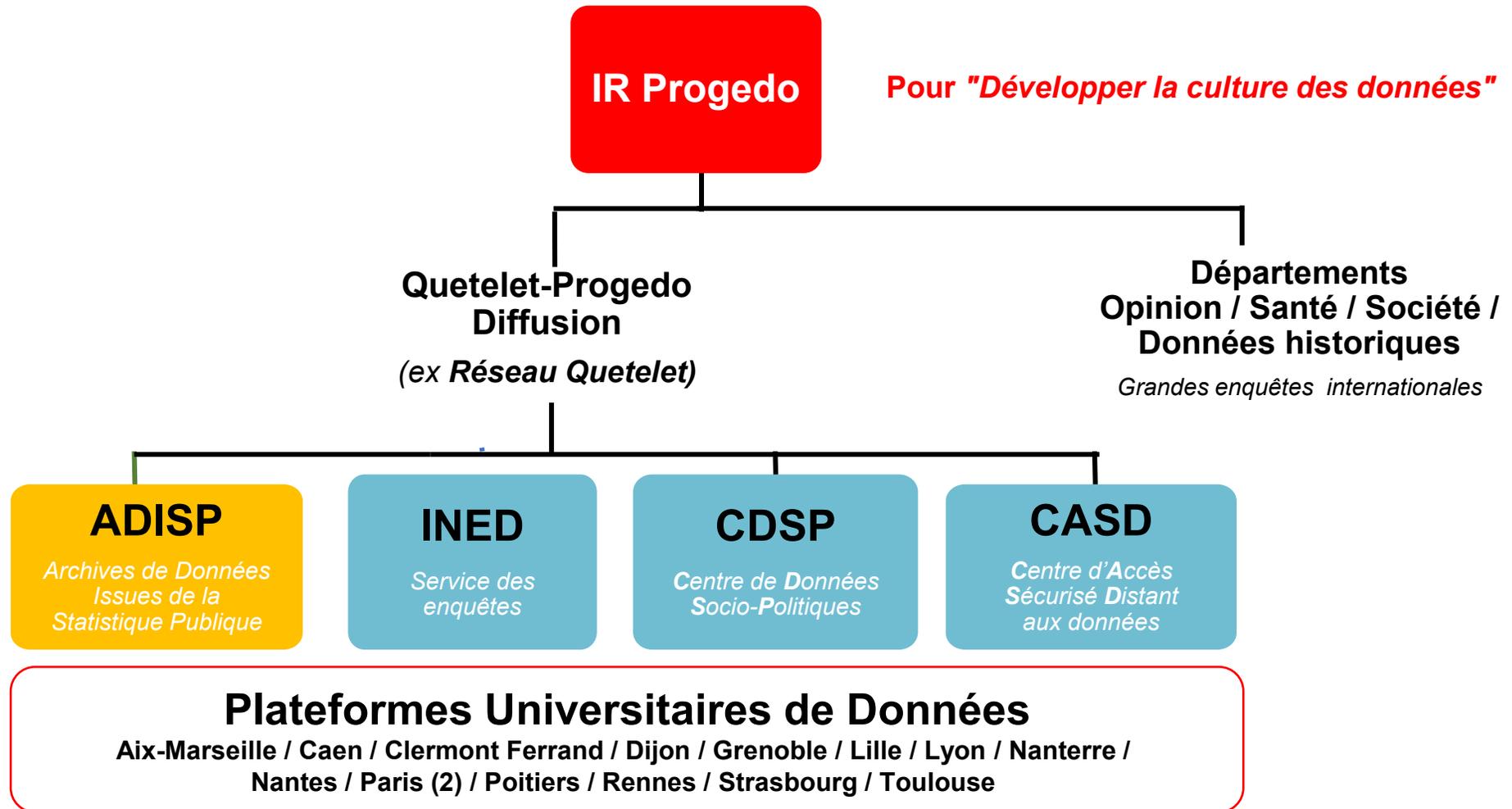
2021 : ouverture aux données de la recherche académique



I. Le périmètre de Quetelet-Progedo Diffusion

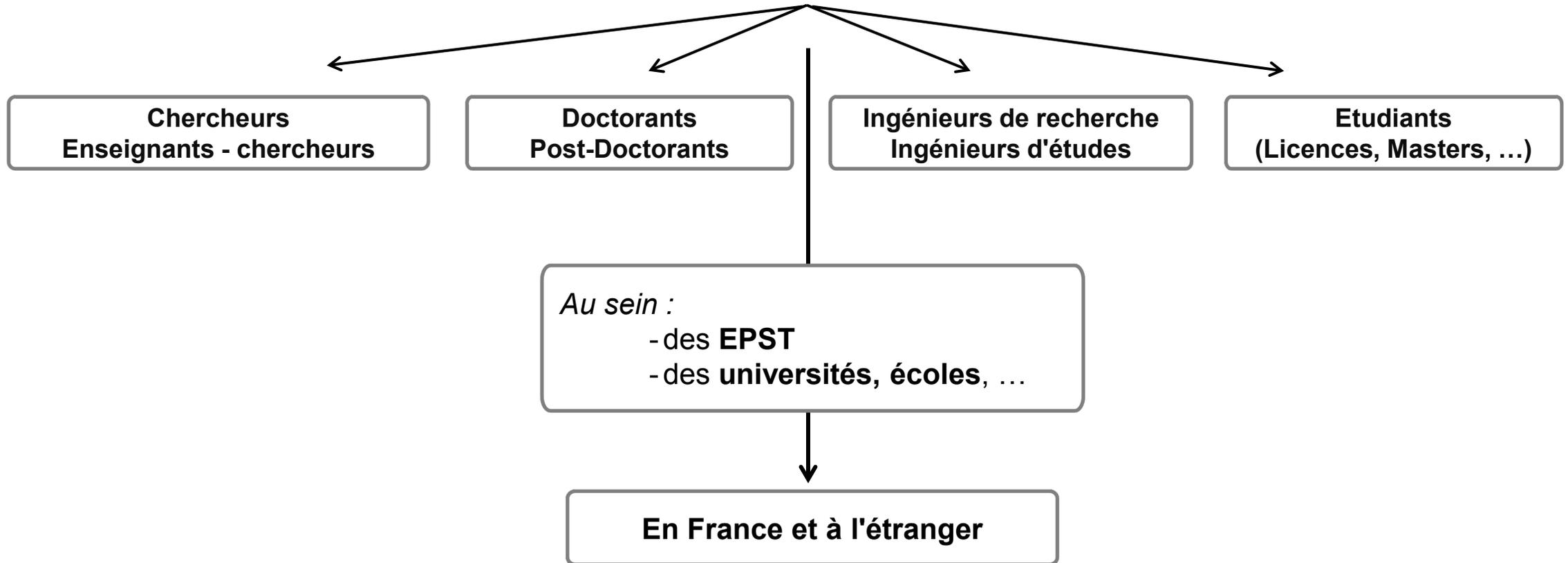


Les 4 partenaires du réseau de diffusion

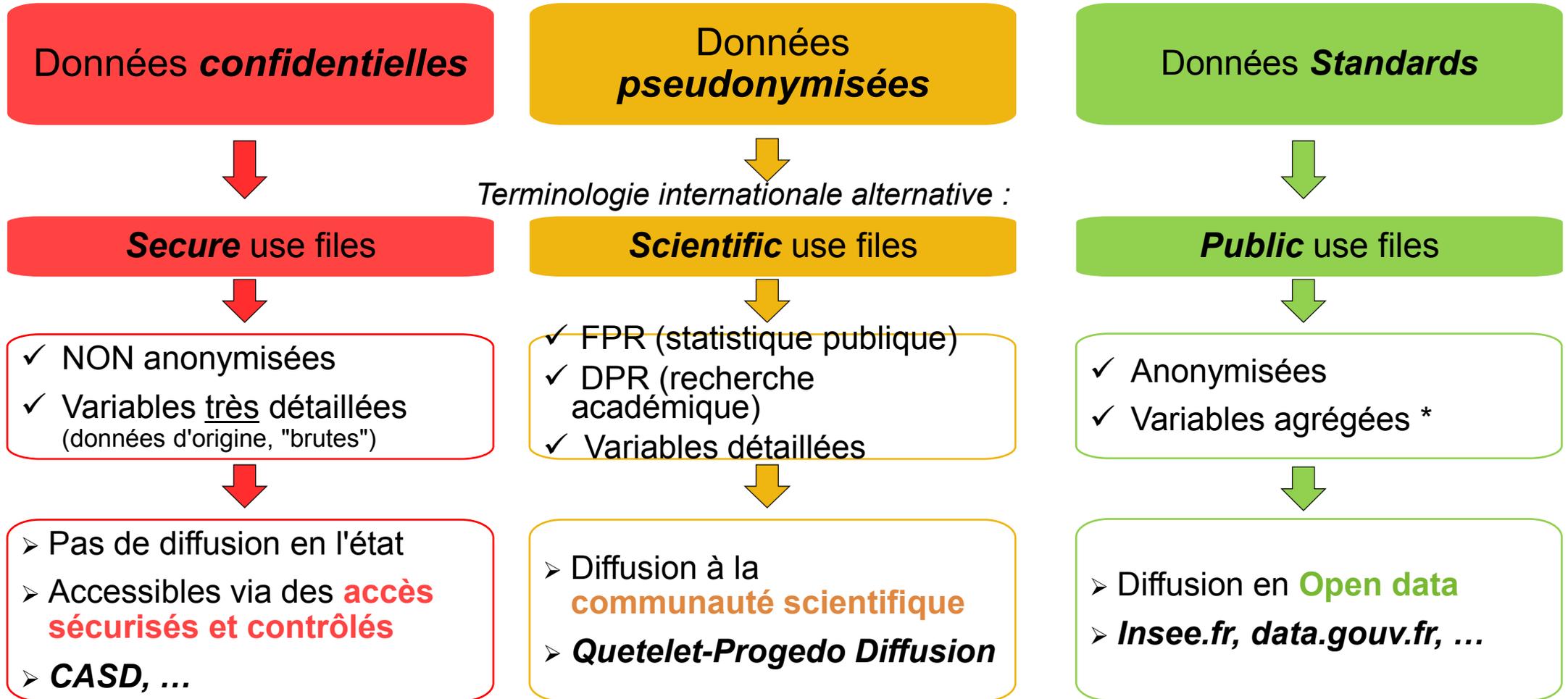


Un service pour la recherche

Accès à la communauté scientifique



Précisions sur les types de fichiers



(*) Agrégations le plus souvent opérées sur les variables géographiques, de professions et autres nomenclatures ; ainsi que sur des variables numériques synthétisées en tranches



II. Elargissement du périmètre aux données produites par la recherche



Rappel du contexte

La loi pour une **République numérique**
Le Plan National pour la **Science Ouverte**
Les **plans de gestion des données**
La **reproductibilité** des résultats
Etc.

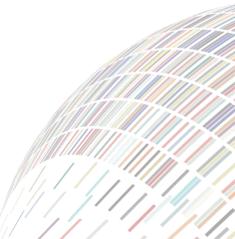


La **question posée** est donc celle du **partage des données** à l'issue d'un projet de recherche.

L'essentiel est de se demander comment partager et quelles sont les règles à suivre pour rendre le partage compatible avec le respect de la vie privée (RGPD).

Cela implique une documentation tout au long du projet et une documentation spécifique des données.

(cf seconde partie de la présentation)



Le rôle de Progedo

**Centralisation des données statistiques en SHS sur Quetelet-Progedo
Diffusion**



**Accompagnement des chercheurs / équipes de
recherche dans la constitution des données / Métadonnées**



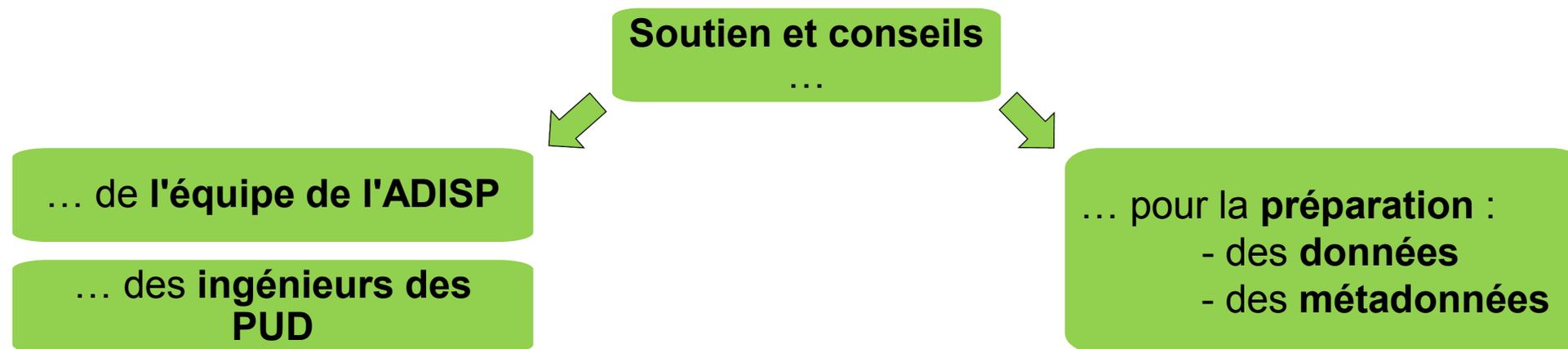
Mise en place des **cadres techniques et juridiques**
de la diffusion



Diffusion des données auprès de la communauté scientifique



L'accompagnement des chercheurs



- ?** Qui peut déposer ? → **Tout chercheur, doctorant...**
constituant ou ayant constitué par le passé un corpus de données
- ?** Quelles données ? → **Données quantitatives** auprès de l'ADISP
Possibilité de dépôt de **données qualitatives** auprès de BeQuali...
- ?** Quelle diffusion ? → Accès restreint à la **communauté scientifique**,
- avec **engagement** des utilisateurs ;
- après **habilitation** pour les données pseudonymisées



Les cadres techniques et juridiques

Les cadres techniques :

- Attribution d'un **identifiant unique** (DOI) au jeu de données / métadonnées
- Intégration dans le **catalogue de QPD** (visibilité / moissonnage par Isidore, Cessda...)
- Intégration de la référence dans l'**application de commande** :
 - **accès restreint** (contrôle par l'Adisp, licence - engagement des utilisateurs (avec habilitation si données pseudonymes))
 - de **suivi des projets** de recherche sur les données diffusées
- Possibilité de mise à jour des données / métadonnées

Les cadres juridiques :

- **Convention** entre Progedo (Cnrs) et l'institution de recherche de l'équipe de recherche
- Signature de l'**engagement des utilisateurs**
 - utiliser les données dans une finalité scientifique
 - respecter les règles de l'art et du secret statistique
 - citer les sources de données
 -



Les étapes

1. **Préparation** du projet de dépôt des données sur Quetelet-Progedo Diffusion (pendant le PGD si possible)

2. Soumission au **conseil scientifique** de Progedo

- 3a. **Convention** entre l'institution de l'équipe de recherche et le CNRS.
- 3b. Fourniture des **données et métadonnées**

4. **Diffusion** des données et de la documentation aux personnes qui en font la demande sur le portail Quetelet-Progedo Diffusion.



Merci pour votre attention

Contact : diffusion.adisp@cnrs.fr

