



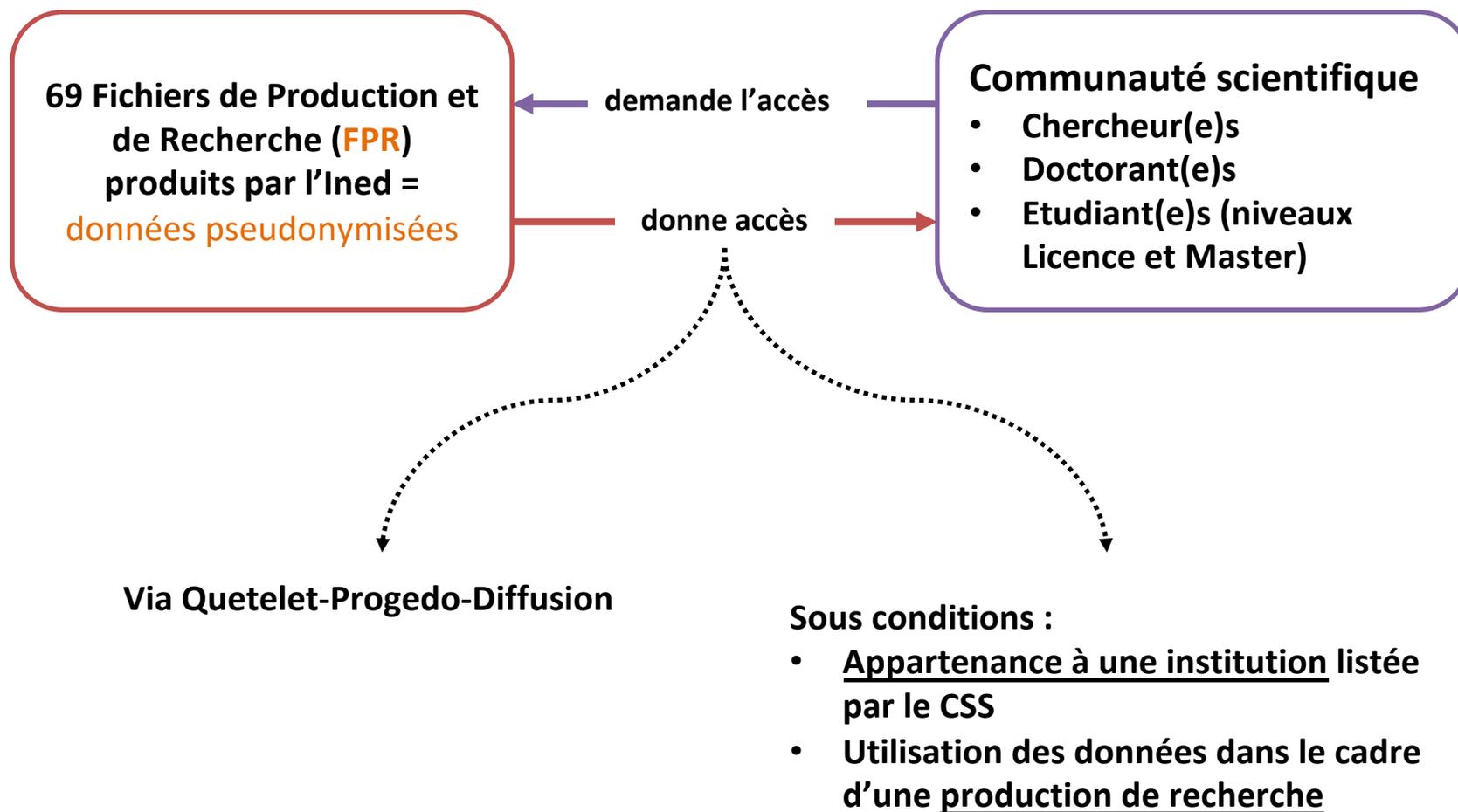
---

# Anonymisation des enquêtes

## Présentation à l'Assemblée Générale LifeObs

Julie LENOIR

# Contexte (1/2)



# Contexte (1/2)

## Besoins non remplis par les FPR

- Accès aux données pour des besoins pédagogiques :
  - Apprentissage de logiciels de statistiques
  - Apprentissage de méthodes quantitatives

## Demande des PUD

- Fichiers accessibles avec un coût d'entrée moindre
- Organisation de cours d'introduction aux logiciels / méthodes quantitatives

# Objectifs

---

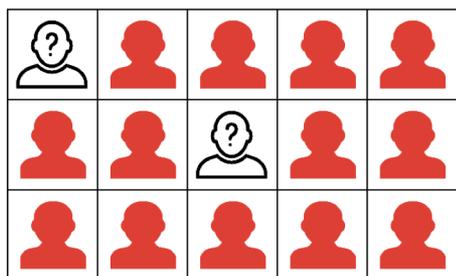
- Fichiers plus accessibles que les FPR
  - Accès simplifié (= fichiers anonymes)
  - Coût d'entrée faible
- A valeur pédagogique
  - Reproductibilité de résultats de recherche
  - Diversité de variables

# Anonymiser des données



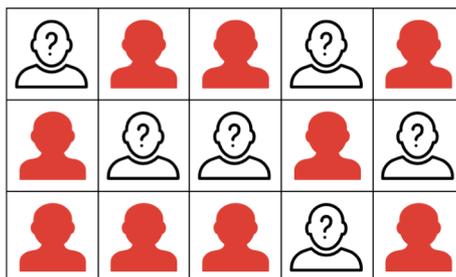
**Fichiers détaillés =**

données  
**désidentifiées** (nom,  
prénom, adresse)



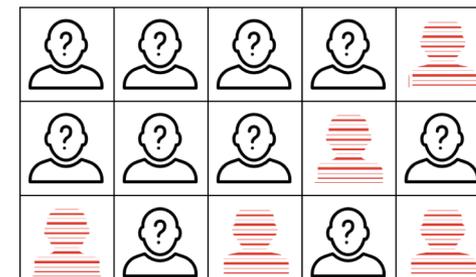
**FPR =**

données  
**pseudonymisées**  
(regroupement de  
modalités)



**FPA =**

données  
**anonymisées** (k-  
anonymat,  
brouillage, PRAM, ...)



# Fabriquer des Fichiers Pédagogiques Anonymes (FPA)

## Identification des résultats à reproduire

Population et sociétés

Tableaux / graphiques

Variables

## Anonymisation

Regroupement de modalités

Recodage de variables

Brouillage, perturbations, ...

## Vérification du fichier

Risque de réidentification

Perte d'information

Reproductibilité des résultats

## Trajectoires et origine (TeO1)

4 Population et Sociétés

73 variables

Brouillage des années de  
rencontre / de mariage

Risque de réidentification  
évalué à 1.80%

## ERFI 1 (vague 1)

1 Population et Sociétés

86 variables

Brouillage des âges

Risque de réidentification  
évalué à 0.19%

# Les prochaines étapes

---

- Fin des vérifications sur les FPA
- Mise à disposition
- Fichiers anonymes de plus grande ampleur : travaux à démarrer en collaboration avec l'INRIA (recrutement en cours)