# Anonymisation et FHIR

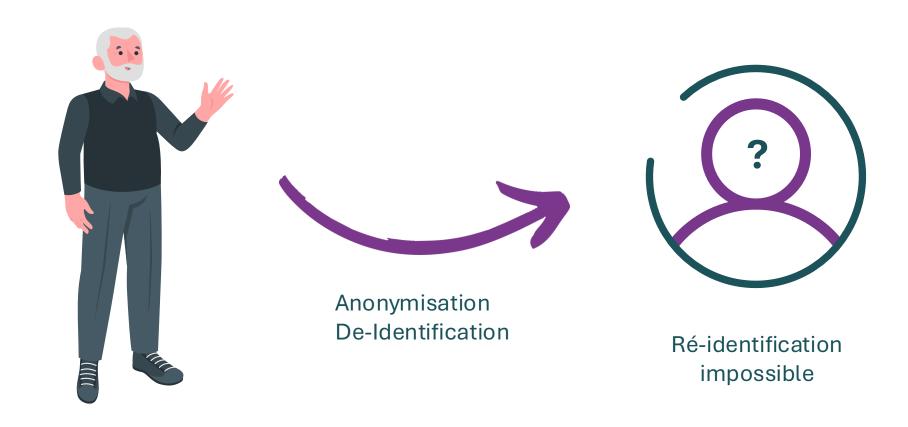


# Anonymisation

Sans FHIR



# **Anonymiser**





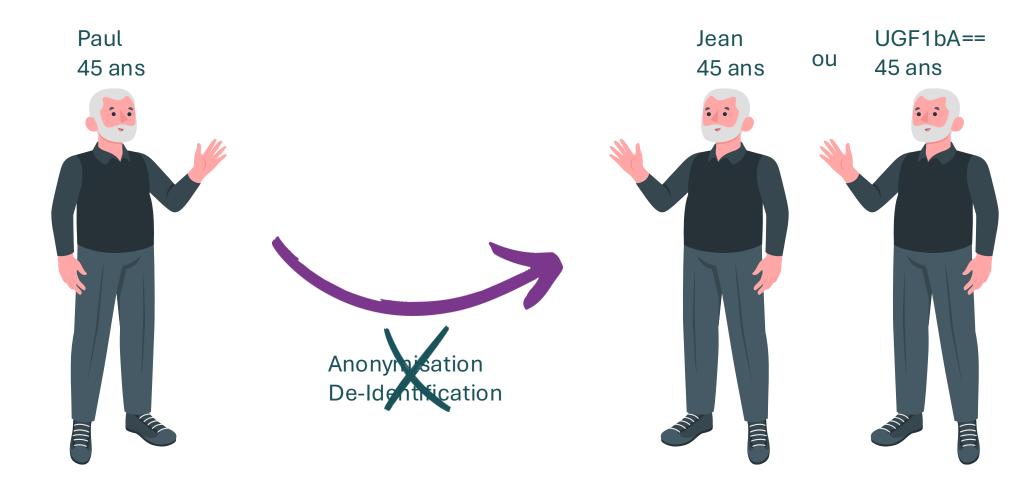
### Pourquoi?

- Exploiter des données personnelles dans le respect des droits et libertés des personnes
- Publier en ligne des informations publiques sans données personnelles (open data)
- Conserver des données au-delà de leur durée de conservation





# Pseudonymisation: piège







### Identification directe: identifiants

- Noms, prénoms
- Identifiants
- Adresses

•

• Valeurs rares : *âge* : seulement 2 personnes ont 116 ans dans le monde.





### Organe de l'UE prévu par le RGPD

- **CEPD Comité européen de la protection des données edpb – European Data Protection Board**
- Reglement sur la protection des données
- Directives protections des données
- Instruments juridiques...

Critères pour déterminer l'impossibilité de la ré-identification





### Individualisation

Isoler une partie ou la totalité des enregistrements identifiant un individu dans l'ensemble des données.

Identifiant un individu même sans ses identifiants directs.

Problèmes de la pseudonymisation.





### Corrélation

Relier au moins 2 enregistrements se rapportant à la même personne ou groupe de personnes. (avec une base extérieure ou

non)





### Inférence

La possiblité de déduire, avec un degré de probabilité élevé, la valeur d'un attribut à partir des valeurs d'un ensemble d'autres attributs.





# Techniques d'anonymisation

Aucune n'est infaillible seule



### Ajout de bruit

Modifier les valeurs originales pour les rendre mois précises. Permet de garder la distribution de l'attribut.

Rend la ré-identification moins fiable.

Choisir le bruit en fonction de la distribution initiale.



#### **Permutation**

Echanger les valeurs entre les individus Permet de garder la distribution de l'attribut.

Ne permet pas de garder la corrélation entre les attributs.

Attentions aux permutations d'attributs corrélés

Si 2 attributs sont très fortement corrélés et l'un des 2 est permuté, il sera possible de détecter cette permutation et de l'inverser.

(Situation professionnelle & salaire)





# Agrégation & k-anonymat (k-anonymity)

Regrouper une personne avec au moins k autres individus

*Individualisation :* plus possible d'individualiser une personne au sein d'un groupe de k utilisateurs.

Perte en finesse et parfois en cohérence Inférence toujours possible







## Diversité l-diversité (l-diversity)

Extension du k-anonymat : au sein d'un groupe, chaque valeur de l'attribut sera représentée au moins | fois.

On garde un degré d'incertitude

#### t-proximité (t-closeness):

Affinement de la l-diversité

Des groupes qui ont la même distribution que l'ensemble de données original.







### **Conclusion anonymat**

- N'est pas une simple pseudonymisation
- Processus unique au contexte : données initiales et analyse voulue
  - Anonymisation au cas par cas
  - Supervision humaine
- Anonymisation n'est pas toujours possible
- Respecter les **trois exigences** réduis considérablement les risques de ré-identification





# Anonymisation et FHIR

Comment appliquer ce processus d'anonymisation?



### **Implementation Guide**

- IG publié sur le build FHIR
- Github IG

- Publier son propre IG
- Bonnes pratiques IG



#### Ressources

- Résumé CNIL
- Avis comission européenne (CEPD)
- Malentendus récurrents (CEPD)
- Résumé Anonymisation (Octopize)
- Résumé des techniques (Octopize)





4 Rue Paul Vatine 17180 Périgny, France +33 5 46 43 46 19 <a href="mailto:contact@xtremsante.fr">contact@xtremsante.fr</a> <a href="mailto:www.xtremsante.fr">www.xtremsante.fr</a>

