

# Projet BNDMR: Récolte de Données et Interopérabilité

7 000 maladies rares  
300 000 patients estimés  
131 centres de référence  
504 centres de compétence  
23 groupes de maladies rares  
54 laboratoires de diagnostic

Meriem Maaroufi<sup>1,2</sup>, Rémy Choquet<sup>1,2</sup>, Marie-Christine Jaulent<sup>1</sup>, Paul Landais<sup>3</sup>

<sup>1</sup>UMRS872 Eq20, <sup>2</sup>Assistance publique des hôpitaux de Paris, <sup>3</sup>Université Montpellier 1, EA 2415, BESPIM CHU de Nîmes

## La BNDMR : Banque Nationale de Données Maladies Rares

Projet de mise en place d'une base de donnée nationale pour la collecte de données de patients maladies rares dans le cadre du PNMR2\* 2011/2014.

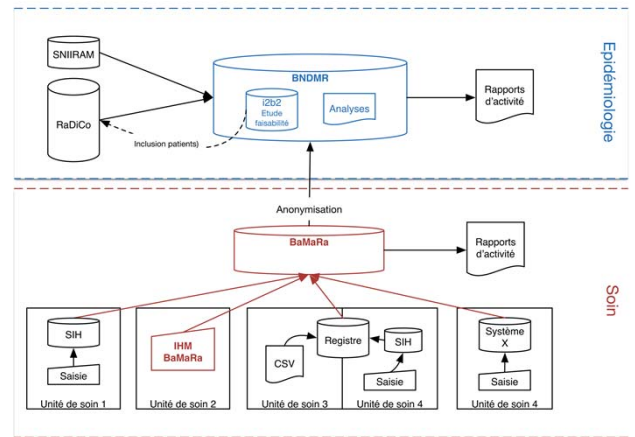
### BaMaRa - Soins (niveau nominatif)

- Recueil de données de patients maladies rares issues des différents Centres de Référence et de Compétence.
- Suivi de la file active des patients maladies rares.
- Evaluation de l'activité et reporting pour la DGOS\*\*.

### BNDMR - Epidémiologie (niveau anonyme)

- Mesure de l'adéquation entre l'offre et la demande de soins dans le domaine des maladies rares.
- Etude de faisabilité pour l'inclusion de patients dans des cohortes.

## Architecture Générale



## 1<sup>er</sup> niveau d'interopérabilité : Identification des patients

### Objectifs :

- Méthode d'identification fiable qui minimise les **doublons** et les **collisions**.
- Chainage intra et inter site de soin des données d'un même patient : un identifiant patient **unique** et **permanent**.
- Chainage SNIIRAM et DMP.

### Types d'identifiants

#### Données nominatives :

- Noms (naissance, marital...)
- Prénoms...

#### Identifiants significatifs :

- NIR

#### Identifiants anonymes :

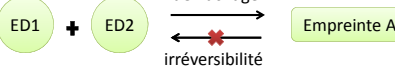
- Identifiants incrémentés
- NIP, IPP...

#### Identifiants dé-identifiés :

- INS (Calculé ou Aléatoire)
- Identifiant SNIIRAM

- Rapprochement d'identités (gérer les doublons).
- Contraintes réglementaires.

### Dé-identification<sup>1</sup>

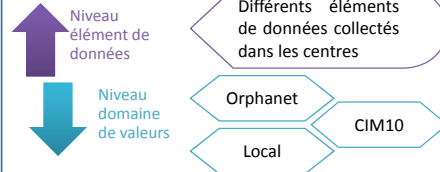


## 2<sup>e</sup> niveau d'interopérabilité : Les données de santé

### Objectifs :

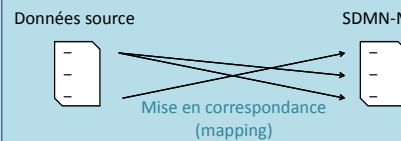
- Inciter les unités de soins à recueillir le même jeu de données : **Standardisation**.
- Essayer d'intégrer les données déjà existantes chez les sources : **Alignement**.

### Standardisation



=> Proposer un **set de données minimal national maladies rares** validé par la DGOS comme un standard pour la collecte des données maladies rares

### Alignement



**Problématique** : +500 alignements à faire (connecteurs à développer) pour connecter toutes les unités de soin.  
=> Comment factoriser?

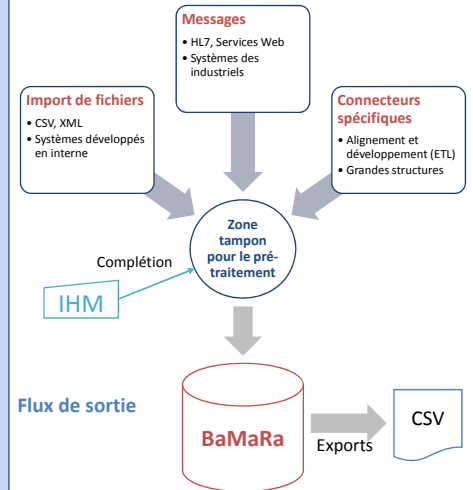
## 3<sup>e</sup> niveau d'interopérabilité : Gestion des flux de données

### Objectifs :

- Rendre les différents systèmes **communicants** au niveau technique.
- Organiser les flux de données pour mieux les gérer.

### Flux d'entrée

=> Ouverture grâce aux standards technologiques.  
=> 3 modes d'intégration de données : *ad hoc*, périodique ou en temps réel.

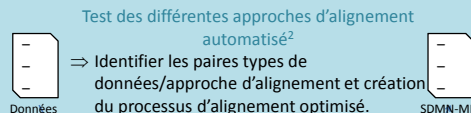


## Méthodologie



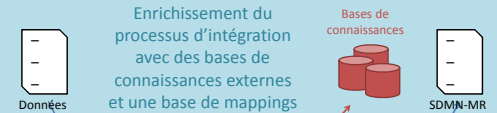
- Créer une typologie des correspondances pour la génération des règles de transformation et de mappings.
- Garder les résultats de l'intégration de données comme une base de validation pour les autres phases.

### Phase I - caractériser



=> Identifier les paires types de données/approche d'alignement et création du processus d'alignement optimisé.

### Phase II - définir les stratégies



### Phase III - factoriser

\* PNMR : Plan National Maladies Rares  
\*\* DGOS : la Direction Générale de l'Offre de Soins  
\*\*\* ED : Elément de Donnée

### Bibliographie

- R. Fraser and D. Willison. Tools for de-identification of personal health information. <https://www.ehealthinformation.ca/presentations/webinars/infoway/deid.pdf>
- J. Euzenat, P. Shvaiko, Classification of ontology matching techniques, *Ontology Matching 2007* Springer, pp 61-72