



# Un entrepôt de données de Santé pour la Recherche

A. Moreau-Gaudry, P. Mossuz



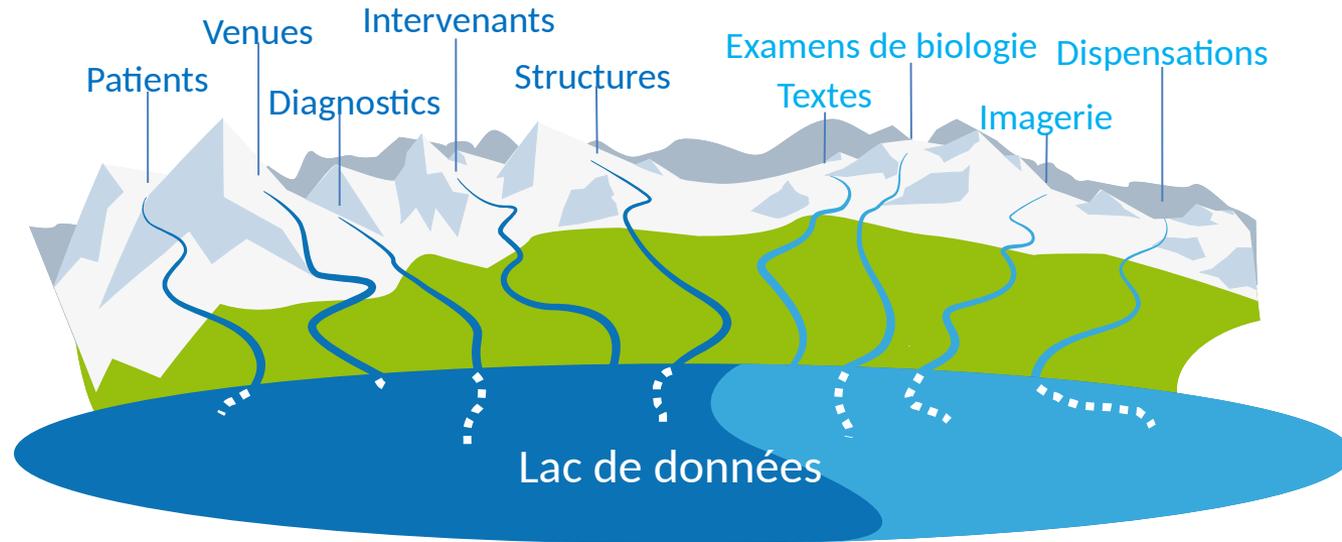
# Plan de la présentation

# Structuration du diaporama

- ✓ Usages et déclaration CNIL associée
- ✓ Principes directeurs et Architecture
- ✓ Outils développés
- ✓ Organisations & exemples de projets

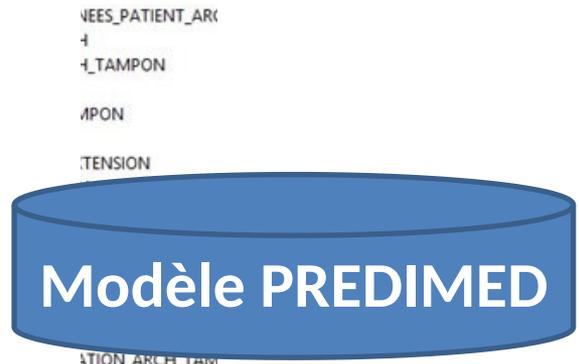
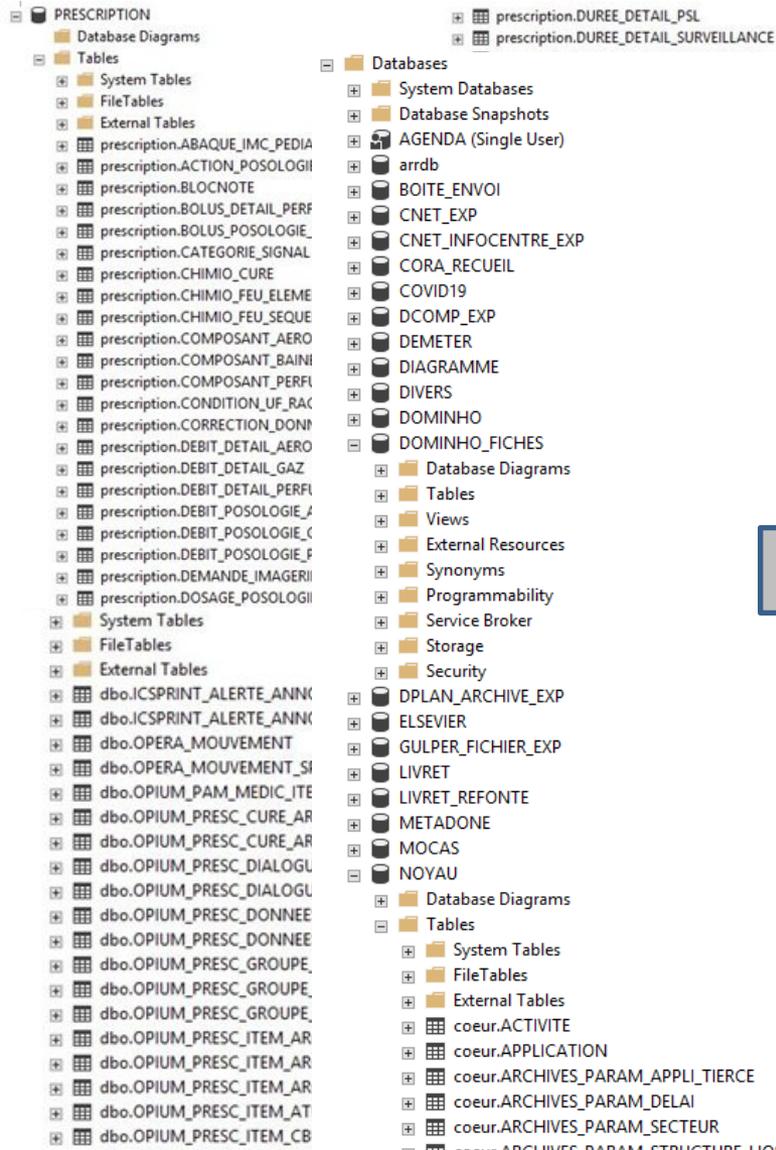
# Usages et déclaration CNIL associée

# Cadre de l'EDS



- Regroupement des données produites lors de la prise de charge des patients par le CHUGA a des fins de:
  - **Recherche** dans le domaine de la santé effectuée par les professionnels du CHUGA et ses partenaires extérieurs ;
  - **Amélioration de la prise en charge des patients** dans l'établissement ;
  - **Pilotage de l'établissement.**
- EDS encadrée par autorisation CNIL du 19/10/2019

# Alimentation de l'EDS (Lac structuré)



# Bases de données récupérées

## Quotidiennement:

- Cristal Net dans son intégralité + les archives / Easily dans son intégralité
- Biologie, prescription médicamenteuse, actes, administratif, mouvements, origine géographique du patient, ...
- CORA (PMSI)
- Diane (anesthésie)
- DMU (dossier médical des urgences)

## Mensuellement:

- Base de données des décès de l'INSEE

## En cours:

- Imagerie, données DICOM
- Textes via Easily
- Comptes rendus (hospitalisation, consultations, opératoire, radiologie, biologie type anapath, urgences), notes, fiches j0, relèves infirmiers, ...

## Envisagé:

- À court terme : Reasic (réanimation)
- À moyen terme : base de données nationales : SNIIRAM, CEPIDC, autres

# Objectifs visés

- Parcours de l'EDS de façon interactive pour :
  - Explorer, Visualiser,
  - Analyser qualitativement,
  - les données cliniques complexes et massives
- De façon interactive (temps réel)
- De manière « conviviale »
- Pour, à terme, créer un « **datamart projet** »
  - Mis à disposition du **projet**
  - Dans le respect de la **législation** en vigueur



# Principes directeurs et Architecture



## Concernant la gestion des données

- Un modèle de lac structuré ouvert et évolutif sans remise en cause des briques développées
- Des espaces spécifiques pour les données volumineuses : textes, images, omiques, autres ...
- Des métadonnées sur les données volumineuses reportées dans le lac structuré
- Un lac brut servant d'alimentation du lac structuré

## Concernant les composants logiciels

- Open source
- « Meilleur » composant du moment pour une fonction donnée
- Architecture « LEGO » construite par intégration de composants logiciels
- Données du lac accessible via des API REST accessibles en R, Python et Java Script

## Concernant l'architecture physique

- Mise en œuvre majoritaire sur des OS LINUX
- OS Windows pour quelques briques logicielles
- Des espaces de stockages dont certains doivent être partageables (NFS par exemple, Samba...)



# Données non structurées (imagerie, texte, données génomiques)

## Des « univers » dédiés

### Texte

- Stockage de l'ensemble des textes dans un système de fichiers
- Création d'une base NoSQL pour index
- Certaines métadonnées reportées dans le lac structuré

### Imagerie

- Injection des données d'imagerie dans un système de fichiers
- Métadonnées dans BD associée à l'outil (Orthanc)
- Certaines métadonnées reportées dans le lac structuré

### Données génomiques

- Chantier en cours

# Outils développés



# Requêtes d'interrogation

Elementary hop expression in the datalake, from a concept to another one :

“source objects class[filters]—named relation[filters]—> target objects class[filters]...”

A complex query can be solved by traversing the data graph with a sequence of elementary hops, forming a path.

“**diagnostic**[icd\_cod~E11]—**concerns**[\*]—>**patient\_did**[age>45]—**concerned\_by**[\*]—>  
**diagnostic**[\*]...”

The events of the datalake navigation interface are converted into elementary cypher-like query strings, successively concatenated to describe the entire filtered path traversed in the graph.

# Recherche dans les textes

Charger config. Choix index ES : txt\_metadone

- Paramètres de la requête (Tous )
- Compte-rendu
  - CR consultation
  - CR urgences
  - Ordonnance
  - CR imagerie
  - CR hospit.
  - Rés. biologie
  - Ordo. médic.
  - CR opér.
  - CR accouch.
  - ECG
  - CR exam.
  - Doc hospit.
  - Dossier social
  - Transm./Consignes
  - Mot suivi
  - Obs./Exam clin.
  - CR RCP
  - CR anapath
  - Antécédents
  - Mot synthèse
  - Expl. fonction
  - Autres
  - Autres doc.

Date début : 01 / 01 / 2012

Date fin : 31 / 12 / 2022

Expression recherchée :  AND  Nbr fautes acceptées

AND  Expression avec distance de mots :  Distance maximale en mots

Send query Documents loaded : 900 on gte 10000 Next data Displayed 1 to 900 Export selected data

Load query Save context

<input type="checkbox"/>	100044	5000031726500		pdf	2020-05-12	SUD MPR ORTHO TRAUMATO SUD	
<input type="checkbox"/>	100044		Bilan rééducation	pdf	2020-05-28	MPR ORTHO TRAUMATO SUD	
<input type="checkbox"/>	100081	1800036479441	Compte-rendu urgences	pdf	2018-04-13	CHIR ARTHROS ET SPORT SUD	
<input type="checkbox"/>	100109	5000023301152	Compte-rendu urgences	pdf	2018-12-28	SAU ADULTES NORD	
<input type="checkbox"/>	100109	5000023301152	Compte-rendu urgences	pdf	2018-12-28	SAU ADULTES NORD	
<input type="checkbox"/>	100144	1500035683001	Compte-rendu urgences	pdf	2015-06-09	CHIR ARTHROS ET SPORT SUD	
<input type="checkbox"/>	100243	1500035698508	Compte-rendu urgences	pdf	2015-06-16	CHIR ARTHROS ET SPORT SUD	
<input type="checkbox"/>	100269	5000011597540	Compte-rendu	pdf	2018-06-07	MPR VASCULAIRE SUD	
<input type="checkbox"/>	10028	1500017017609	Compte-rendu	pdf	2015-06-11	NEURO GEN	
<input type="checkbox"/>	100384	1500011452365	Compte-rendu urgences	pdf	2015-08-14	CENTRE URG CHIR	
<input type="checkbox"/>	10045	5000002293205	Compte-rendu imagerie	pdf	2016-10-24	MED NUC	
<input type="checkbox"/>	100457	5000040622618		pdf	2021-07-26	ORL CHIR CERVICO-FACIALE	
<input type="checkbox"/>	100461			pdf	2021-02-12	EFCR	
<input type="checkbox"/>	100523	1500035804172	Compte-rendu	pdf	2015-09-02	REED CARDIAQUE	
<input type="checkbox"/>	100523	1500035804172	Autre	pdf	2015-08-17	REED CARDIAQUE	

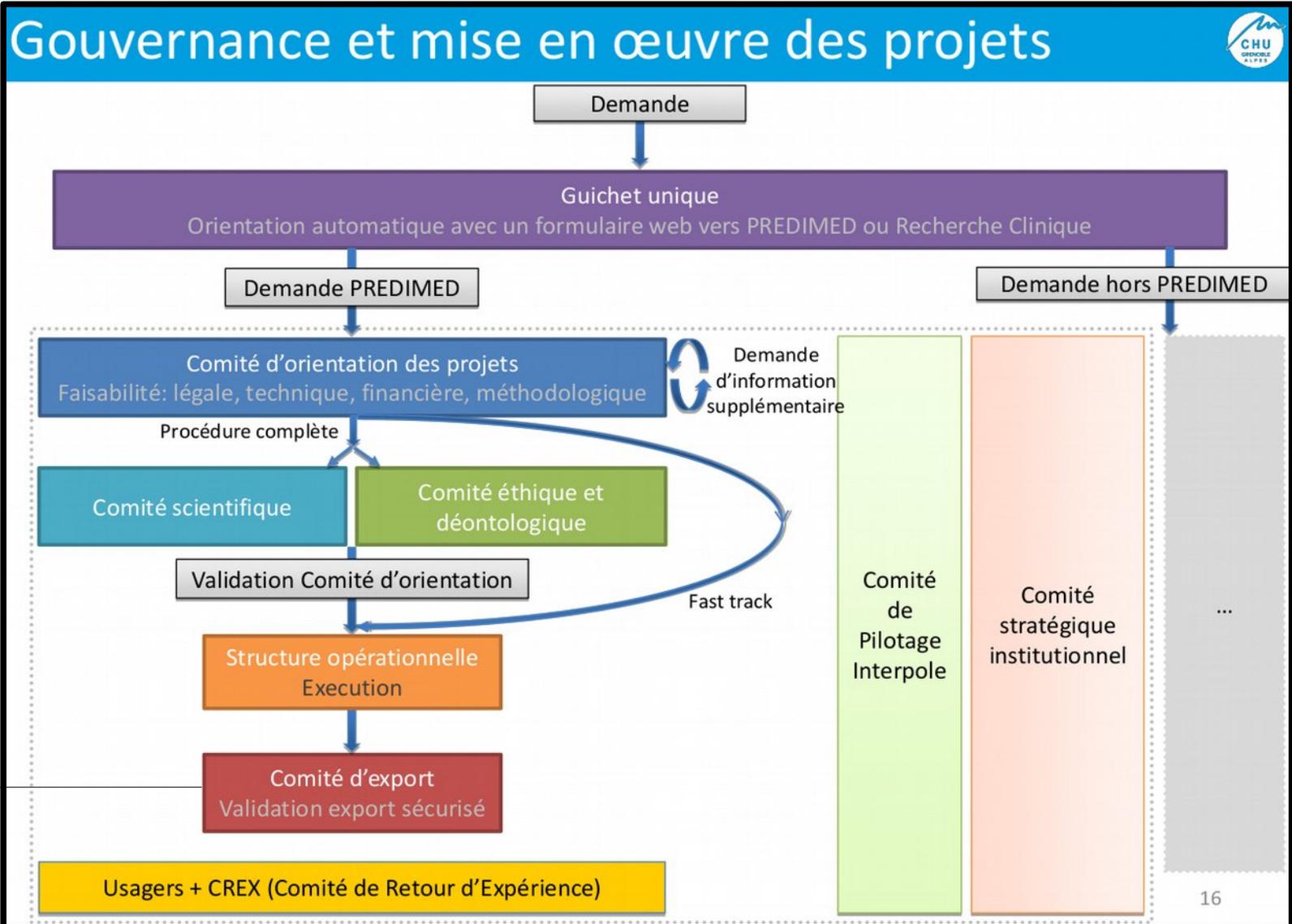
Fragment de texte contenant les expressions recherchées

patient de 13 ans vient pour douleur coude droit ATCD Ttt et allergies : 0 Choc samedi au judo sur chute, dimanche chute à vélo, mal à la palpation du coude un peu gonflé mais aucune limitation de mouvement, pronation supination et flexion

IPP Patient : 100243 Venue : 1500035698508

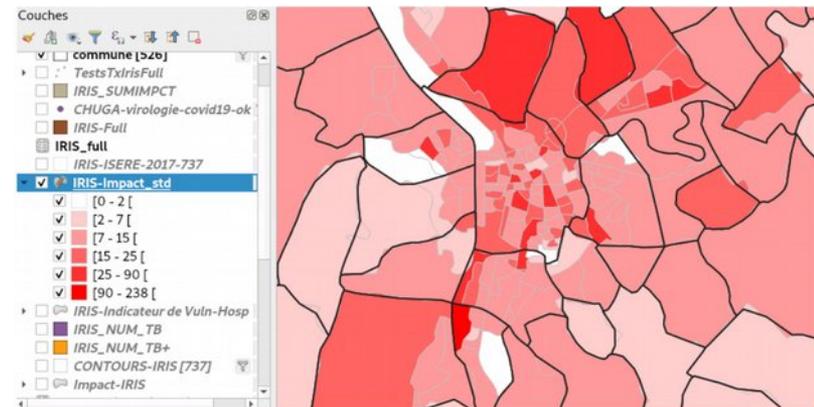
# Organisation projet

# Les différents comités



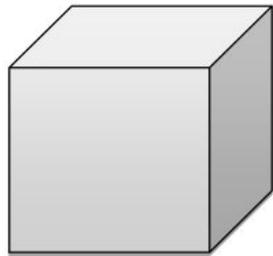
# Exemples de projets

- Un Projet d'imagerie : ODIASP
- DEMETER : Développement de modèles de Machine Learning prédictifs d'événements indésirables liés aux soins
- IA & PMSI MCO : un enjeu de performance
- Cartographies COVID
- DAMON (projet en réanimation ...)
- NewCare (organisation aux urgences)
- Etc...

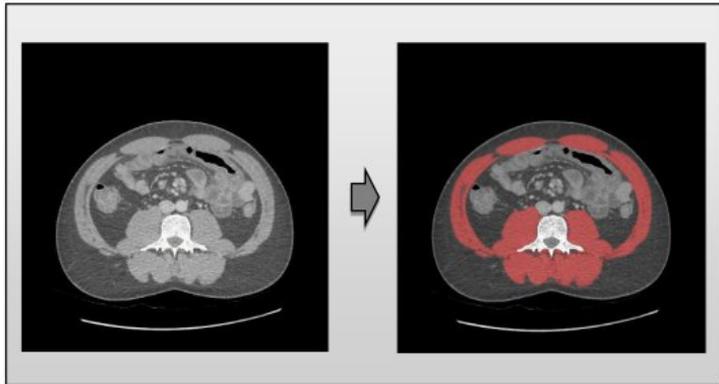


# ODIASP

**Pr Eric Fontaine**  
Dr Cécile Bétry  
Pr Gilbert Ferretti  
Pr Ivan Bricault



Examen scanner



Outil IA ODIASP

ISM

## PREDIMED

Examens  
scanners

ISM

Données  
biologiques

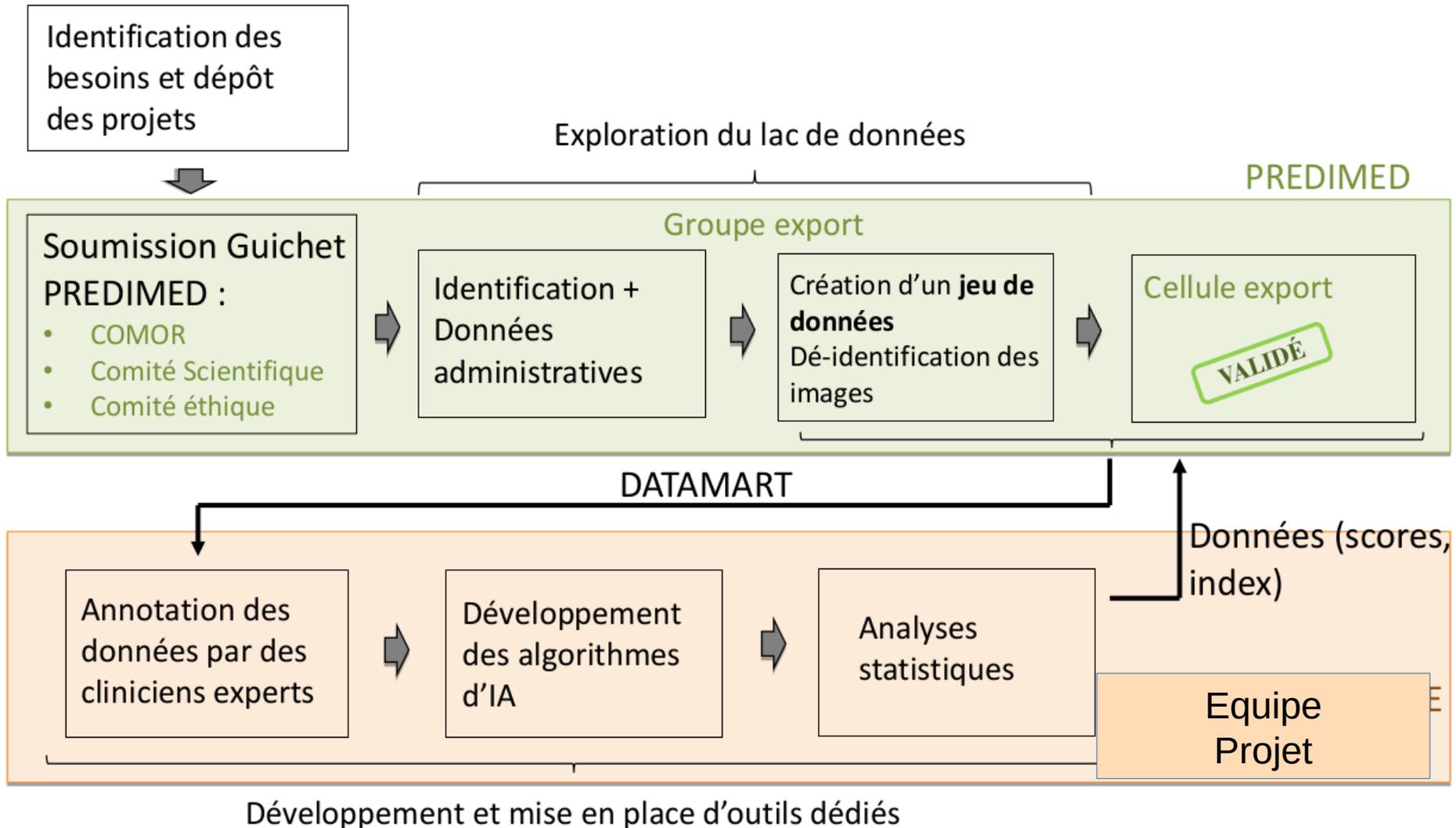
Diagnostiques

...

Nouveaux  
marqueurs de  
la sarcopénie

Automatiser la détermination de l'ISM

# Exemples de projets



# En pratique

- Un fonctionnement par projet
- Nécessité d'anticipation +++
- Un rôdage en cours

## Contacts :

- [Svetlana.artemova@chu-grenoble.fr](mailto:Svetlana.artemova@chu-grenoble.fr)
- [Christian.Lenne@univ-grenoble-alpes.fr](mailto:Christian.Lenne@univ-grenoble-alpes.fr)
- [Alexandre.moreau-gaudry@univ-grenoble-alpes.fr](mailto:Alexandre.moreau-gaudry@univ-grenoble-alpes.fr)

Merci de votre attention!



# Questions

**1<sup>er</sup>** employeur du bassin grenoblois  
avec plus de **9 500** professionnels

Plus de **2 100** lits et places

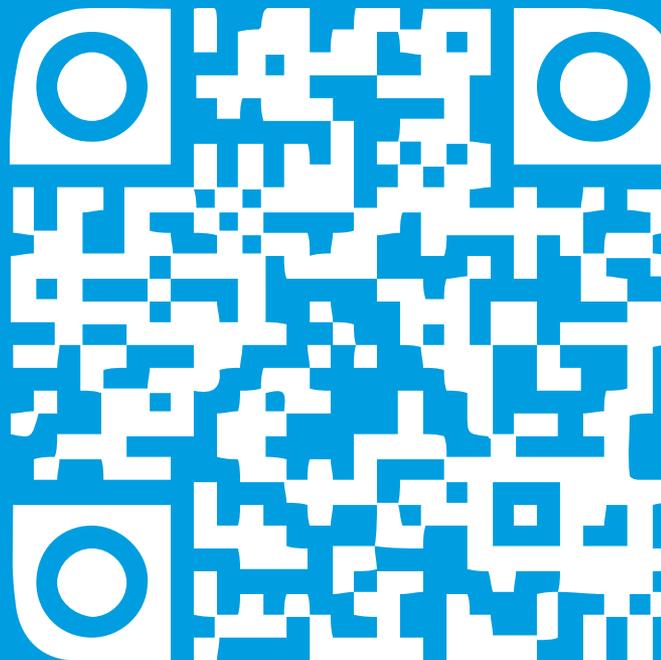
Le **1<sup>er</sup>** *trauma center* de France

Plus de **2 400** patients accueillis par jour

**9** instituts de formation d'excellence  
avec **1200** élèves formés par an

**1 400** études cliniques en cours

**8<sup>e</sup>** Délégation à la Recherche Clinique  
et à l'innovation (DRCI) en France



[www.chu-grenoble.fr](http://www.chu-grenoble.fr)



**04 76 76 75 75**

CHU Grenoble Alpes, Bd de la Chantourne, 38700 La Tronche