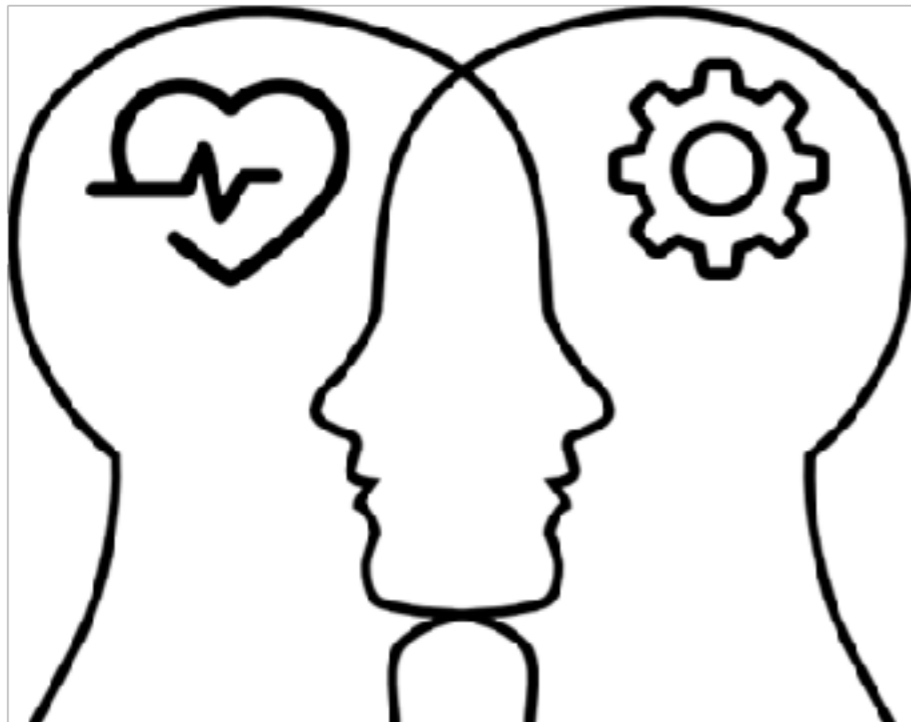


# Datathon inter-CHU



# Datathon

- Format habituel :
  - 48h dans une ville
  - Une base de données de recherche (MIMIC par ex)
- Ex : <http://blogs.aphp.fr/dat-icu/>



- Exemples de problématiques soulevées
  - Prédiction de la durée de séjour en ICU
  - Prédiction de l'évolution de l'insuffisance rénale en ICU
  - Codage automatique des CRH ...

# Centres

- Pour le moment

[vianney.jouhet@chu-bordeaux.fr](mailto:vianney.jouhet@chu-bordeaux.fr)



[prosper.burq@gmail.com](mailto:prosper.burq@gmail.com)



[nicolas.paris@aphp.fr](mailto:nicolas.paris@aphp.fr)  
[adrien.parrot@caramail.fr](mailto:adrien.parrot@caramail.fr)

- Et bien d'autres nous l'espérons...



**Propositions à discuter**

# Quand et où ?



- 1/ Lors d'un **week-end** et **multicentrique** :  
Impose que tous les centres partenaires aient un cluster de calcul opérationnel  
Présentation des résultats lors de MedInfo à Lyon

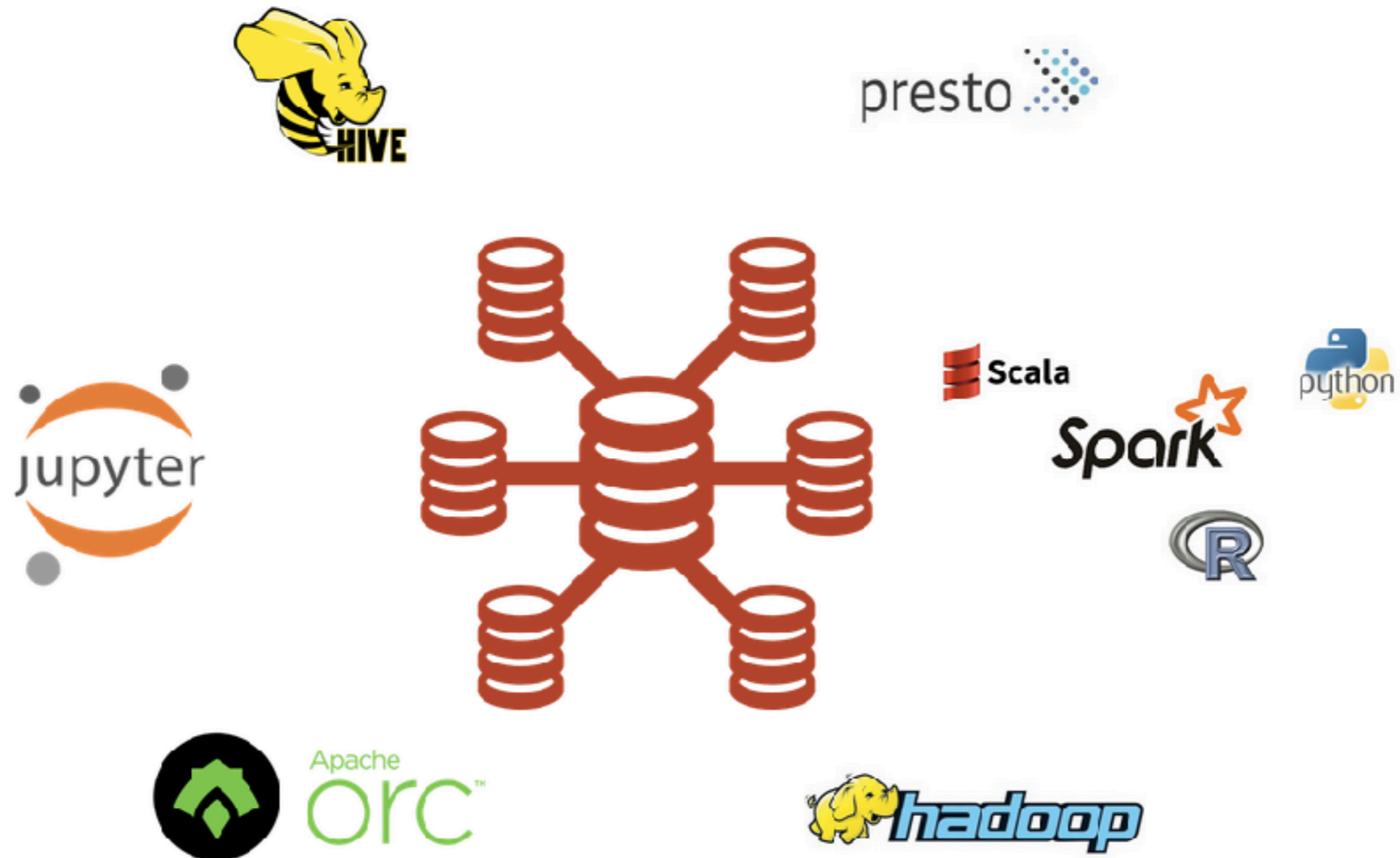


- 2/ Lors d'un **week-end** et dans un lieu **unique** : AP-HP  
Sur le cluster de calcul déjà en place



- 3/ Sur **plusieurs mois** et **multicentrique** : par exemple  
Accès à un dataset publique 3-4 mois avant  
Et présentation des résultats lors de MedInfo à Lyon

# Exemple de cluster



# Format des données

- OMOP

<https://www.ohdsi.org/data-standardization/the-common-data-model/>

Pourquoi : standardisation des concepts, librairies d'analyse variées

- Champs d'action

Démographie => âge, genre

Conditions => infarctus du myocarde

Medicaments => paracétamol 1g x 4 / J PO

Laboratoires. => Hb 10 g/dL

+/- Physiologie => Fréquence cardiaque 100 / min

# Jeu de données



1/ Données fictives

+/- mise en ligne sur : <https://www.data.gouv.fr/fr/>

Ressource : <https://github.com/EDS-APHP/pySyntheticDatasetGenerator>



2/ Jeu de données américain de recherche

Ex : MIMIC-OMOP

Ressource : <https://github.com/MIT-LCP/mimic-omop>



3/ Datas pseudonymisées

Ressources : aucune



# Travail à effectuer

1/ Mapping des concepts locaux => concepts standards

Ex : icd9 => SNOMED

Partage des terminologies sur un GitHub : <https://framagit.org/interchu/conceptual-mapping>

Ressources **actuelles** : 10 ETP (équipe référentielle APHP)

2/ Extract, Transform, Load (ETL)

Ex: ORBIS, Dx CARE => OMOP

Ressources **à financer** : équipes de <https://ohdsi.org/>

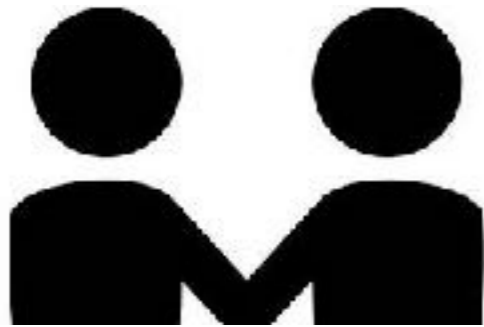
3/ Si données pseudonymisées

Evaluation du risque de ré-identification

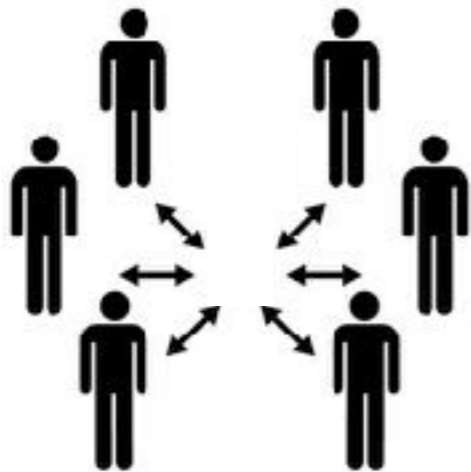
Ressources **actuelles** : 2 ETP (AP-HP et Privacy Analytics)



**Intérets**



- Se rencontrer  
Faire évaluer les pratiques communes



- Décloisonner, échanger, se rencontrer



- Améliorer la recherche médicale, penser l'IA avec les médecins, améliorer la prise en charge des patients

Rejoignez nous sur framateam :  
[framateam.org/interchu](https://framateam.org/interchu)  
(invitation nécessaire)