



# Stratégie Data-IA à l'IPC

## Application dans les projets de recherche

Présenté par :

Marc Lambert – Directeur de la transformation du numérique

# Institut Paoli Calmettes

Patients en file active	44 859
Consultations	116 262
Lits	277
Chirurgie	11 461

Imagerie + Imagerie Interventionnelle  
Laboratoire de biologie du cancer  
Radiothérapie  
Médecine Nucléaire

Nombre de patients inclus	1 473
Publications de recherche	411
Nbre de chercheurs (CRCM)	390





# Missions Data-IA du Département de la Transformation du Numérique

- **Rendre les données disponibles aux départements métiers**
- **Assurer la conformité réglementaire de l'accès aux données**
- **Mettre en œuvre les moyens techniques**
  - **Infrastructure (Réseaux ; Stockage ; Puissance de calcul)**
  - **Environnements logiciels (Python ; R ; Jupiter ; SAS ...)**
- **La qualité des données ?    ➡    C'est l'affaire de tous !**

# Stratégie Data-IA de l'IPC pour la recherche

- **Stratégie Data-IA : un pilier du Projet d'Etablissement**
- **Gouvernance et Ethique**
- **Data Factory et projet EDS**
- **Recherche : CRCM, MIB, EDS Méditerranée, Anapath Digital, 4DOmics...**
- **Pose les bases pour le développement du jumeau numérique**

# Les Données de l'IPC

- 3 Domaines de données
- Cycle de vie de la donnée
- Typologie et réglementation spécifiques



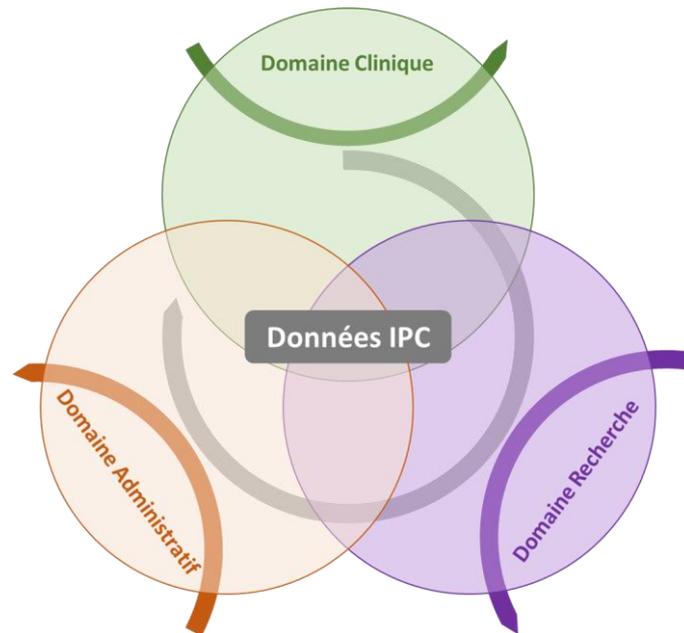
Base de données



Rapports, comptes rendus, publications...



Imagerie



# Stratégie Data Factory

The logo for LORAH features the word "LORAH" in a bold, black, sans-serif font. The letter "O" is replaced by a blue circle with a white arrow pointing to the right, and a small orange triangle is positioned at the bottom left of the circle.

Large Oncologic Repository  
for Advanced Healthcare

by Institut Paoli-Calmettes

The logo for OSCAR features the word "OSCAR" in a bold, black, sans-serif font. The letter "O" is replaced by a blue circle with a white arrow pointing to the right, and a small orange triangle is positioned at the bottom left of the circle.

Oncology Science for Care,  
Analysis & Research

by Institut Paoli-Calmettes

The logo for AIOLI features the word "AIOLI" in a bold, black, sans-serif font. The letter "O" is replaced by a blue circle with a white arrow pointing to the right, and a small orange triangle is positioned at the bottom left of the circle.

Artificial Intelligence for Oncologic  
and Lifescience Innovation

by Institut Paoli-Calmettes

## Data Factory du domaine clinique & administratif

- Regroupe toutes les données au domaine hospitalier & administratif
- Fusion et qualification des données de santé produites par les activités hospitalières
- Préparation et mis en forme des données pour tous les usages ultérieurs
- Protection maximale des données identifiantes
- Outils de cohorting (Consore, Medexprim)
- Analytics et BI

## Data Factory du domaine recherche

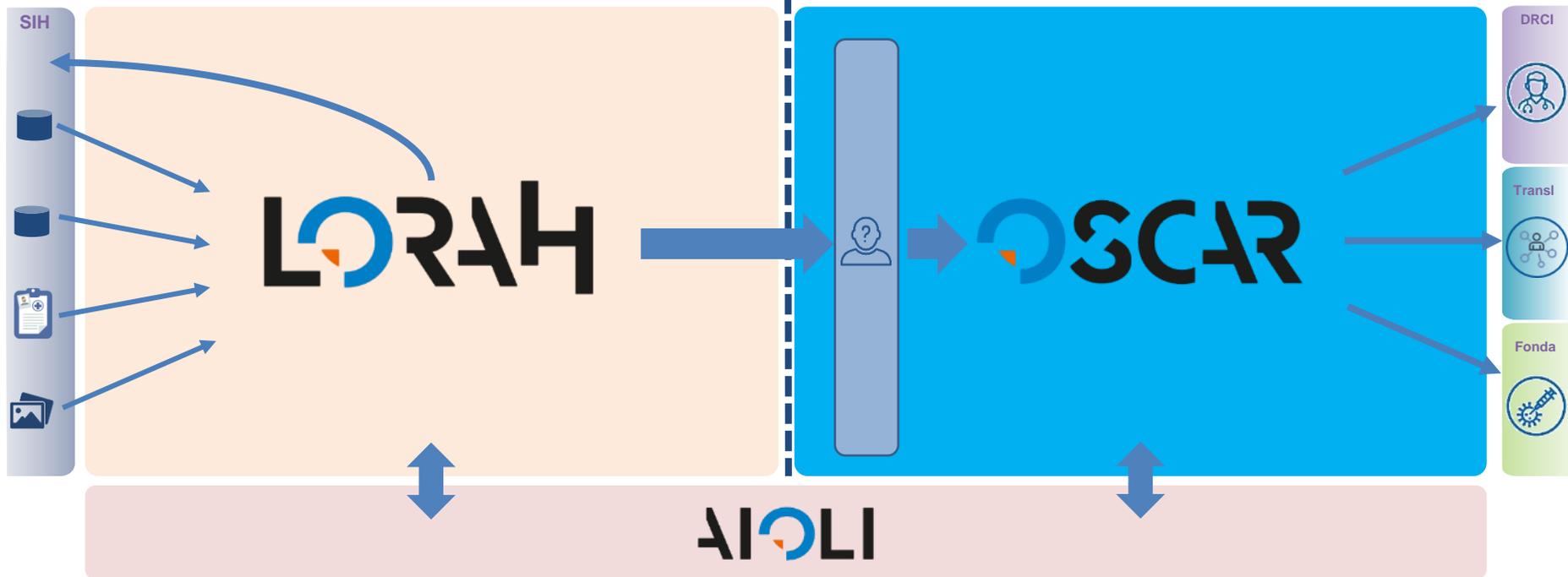
- Regroupe toutes les données appartenant au domaine de la recherche en conformité avec la CNIL
- Héberge l'EDS ainsi que les espaces projets sécurisés
- Pseudonymise les données provenant de LORAH
- Gère le partage des données externes utiles aux projets de recherche (CRCM, autres...)
- Exporte les publications ou le partage de données des projets multicentriques.

## Data Factory du cluster de calcul IA

- Regroupe les ressources technologiques de l'IPC dédiées à l'Intelligence Artificielle
- Optimise l'utilisation des ressources dans un espace sécurisé
- Attribution dynamique des ressources de calcul auprès de LORAH ou OSCAR
- Planification et distribution de l'utilisation des ressources
- Optimisation du cout et valorisation des ressources de calcul

## Domaine Clinique & Administratif

## Domaine de la Recherche



# FHIR, au cœur de notre stratégie Data

- FHIR = Fast Healthcare Interoperability Resources
- Conçu pour tous les domaines de données des hôpitaux
- Open Standard porté par ANSI et ISO dans le groupe HL7
- Conversion simple vers OMOP, OSIRIS, CDISC, OpenEHR
- Bidirectionnel RESTful API, Web native API

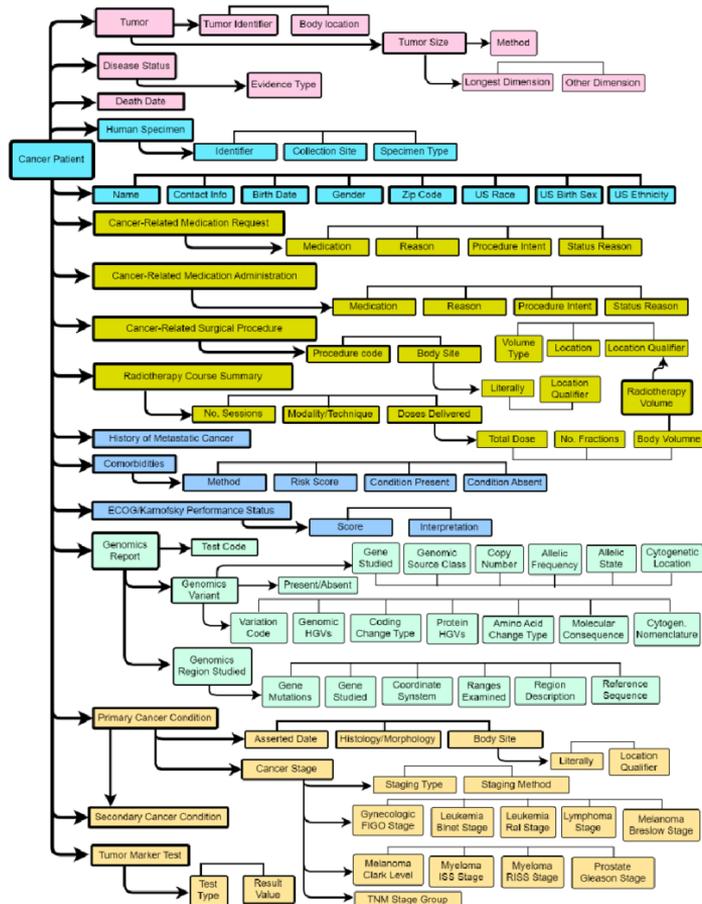
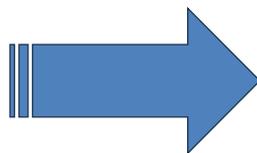




+



+



# Entrepôt de données de santé

Patients référencés	>150 000
Nombre de séjours	> 7 000 000
Références en Data Cat.	19 237
Capacité de stockage	250 To

Entrepôt format FHIR R4

Domaine sécurisé & pseudonymisé

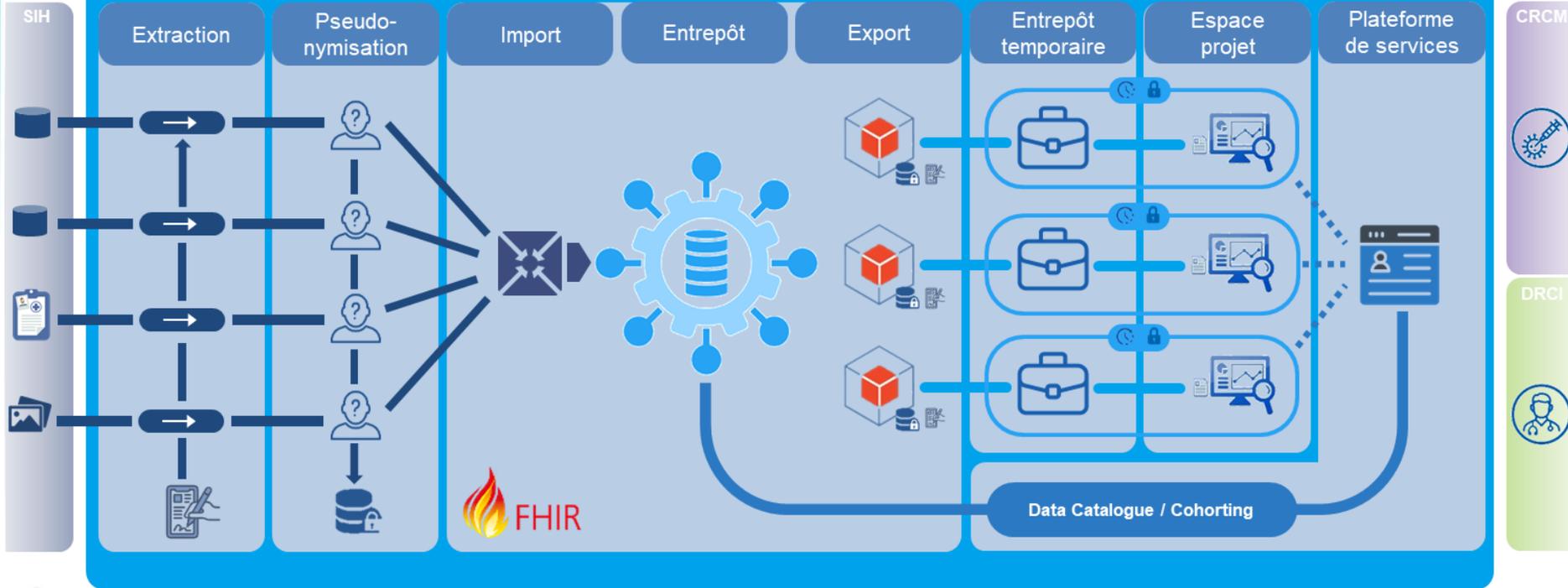
Espace projet isolé

Cluster de calcul HPC IA

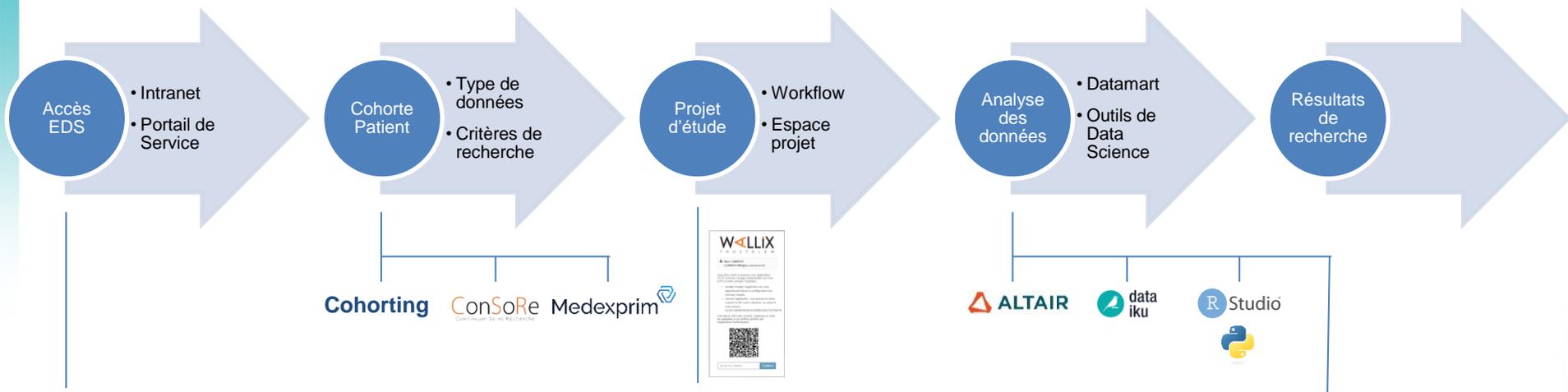
Nombre de cœurs CPU	576
Cartes GPU L40S	4
Stockage HP dédié	56 To



## Entrepôt de Données de Santé



# L'utilisation de l'EDS en pratique



Cohorting ConSoRe Medexprim



ALTAIR data iku R Studio



Together,  
shaping the **FUTURE** of oncology



[www.institutpaolicalmettes.fr](http://www.institutpaolicalmettes.fr)