

# Explore

## Connaître et découvrir l'activité scientifique de l'IRD

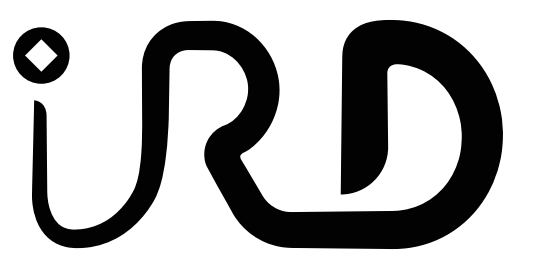
Daniel Salas<sup>1</sup>, Pierre Le Corre<sup>2</sup>,  
<sup>1</sup>IRD MSO, <sup>2</sup>IRD DDUNI



mnemotix  
Donnons du sens à vos données



DATAACTIVIST



Déployer la recherche  
Partager la science  
Transformer l'avenir

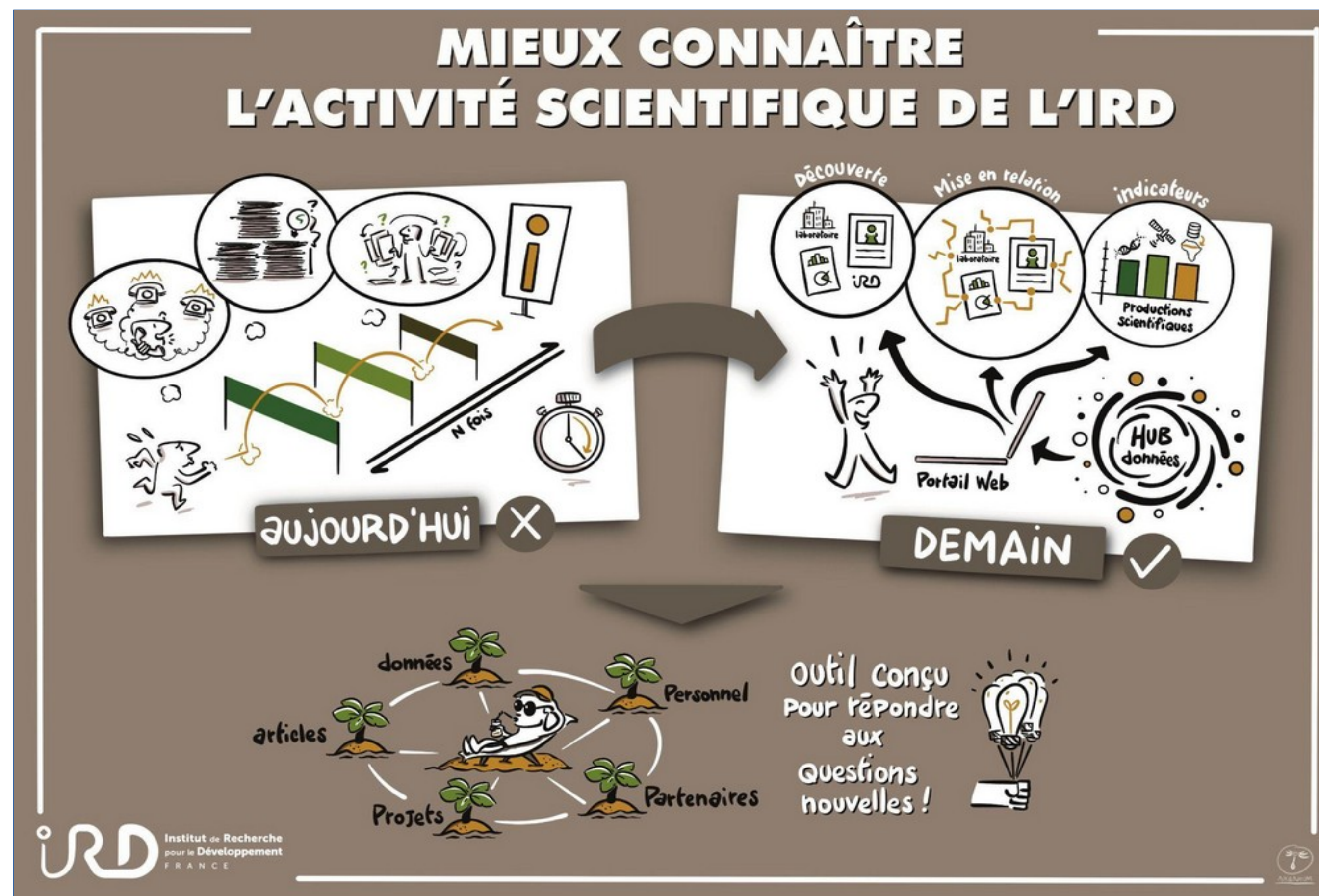
## 1. Contexte

### Stratégie numérique IRD

- Valoriser les données et les activités des scientifiques
- Amplifier la **transparence** des activités institutionnelles

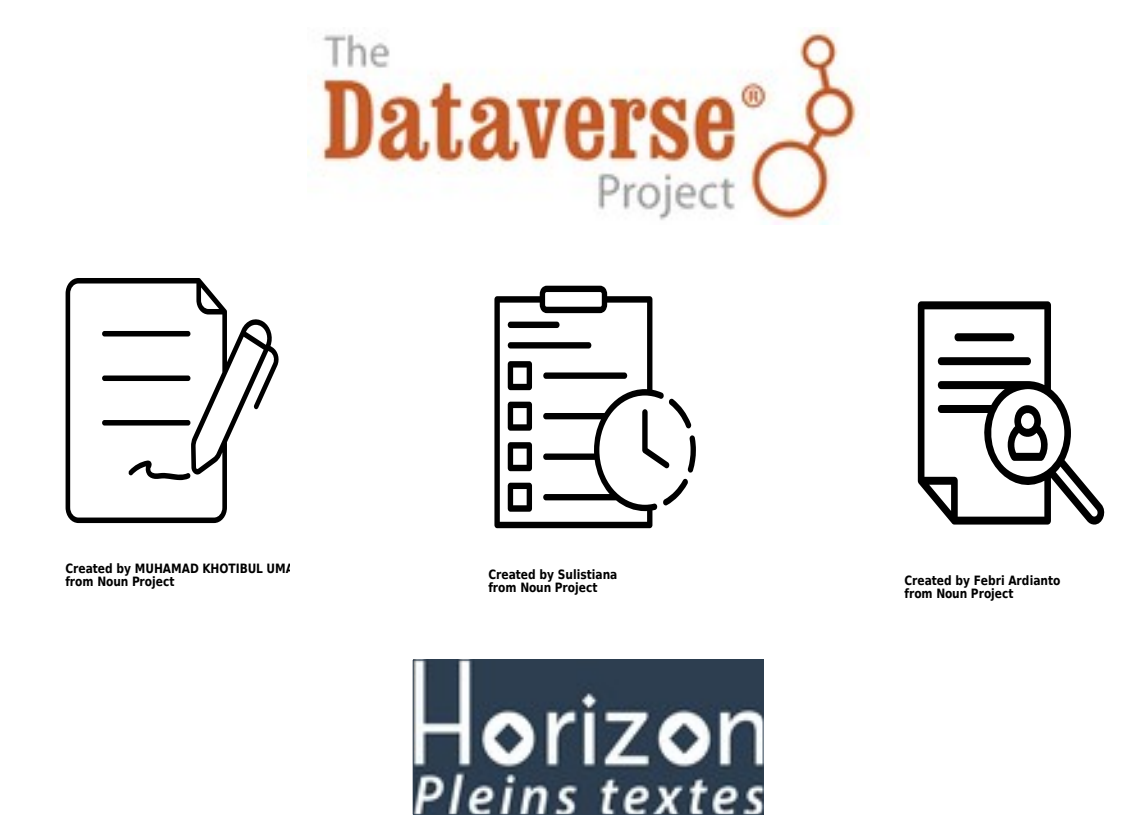
### Enjeux et objectifs du projet

- Désiloter des données dispersées : connecter les sources de données
- Augmenter la **découvrabilité** des données : utiliser des référentiels partagés
- Exposer des informations complètes et à jour : créer une interface de consultation

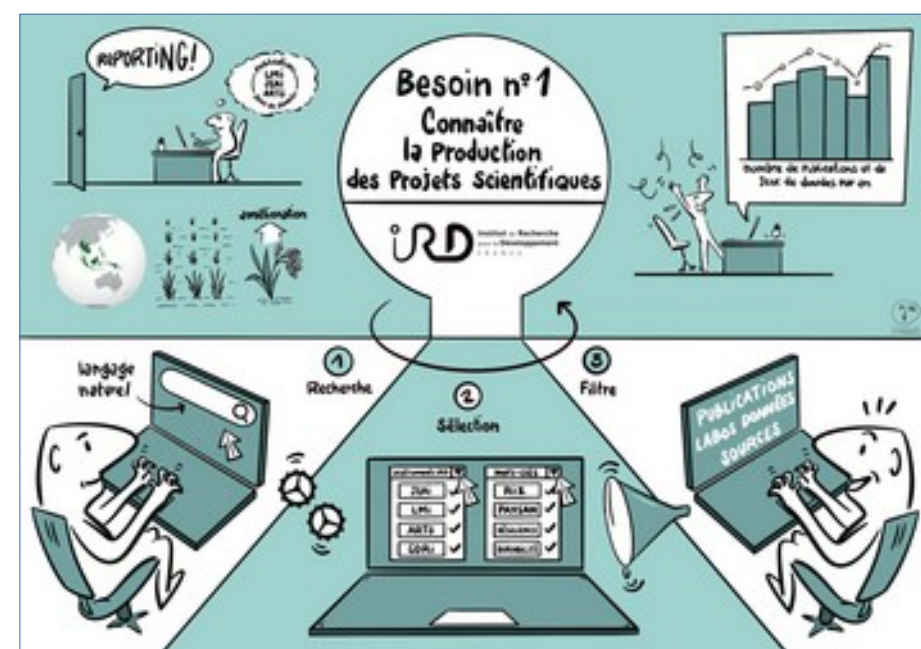


### Sources de données

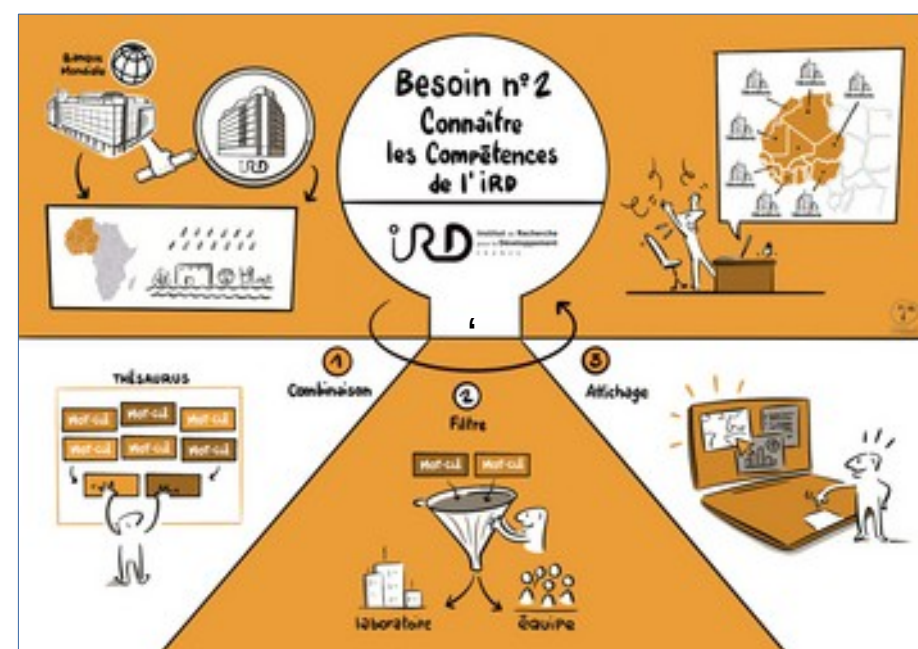
- **Horizon** : archive ouverte des publications IRD
- **DataSuds** : entrepôt de données IRD
- **IRD Ref** : base RH (structures, personnes)
- **IRD Contrat** : contrats
- **FAA** : fiches annuelles d'activités



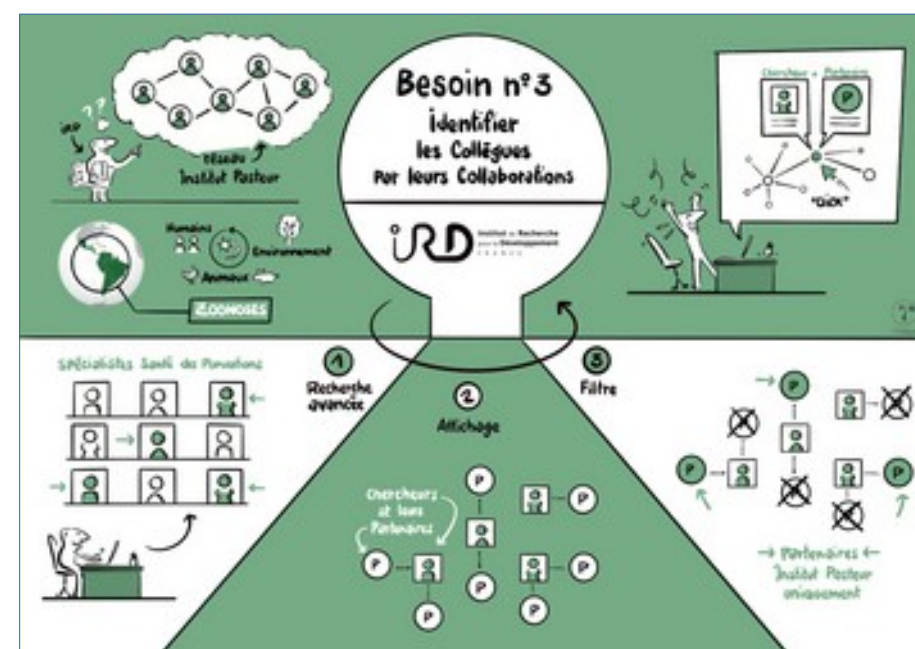
## 2. Cas d'usage



**Gestionnaire de données**  
→ connaître les productions  
→ faire une revue de littérature



**Responsable de structure**  
→ connaître les compétences  
→ répondre aux enquêtes de bailleurs

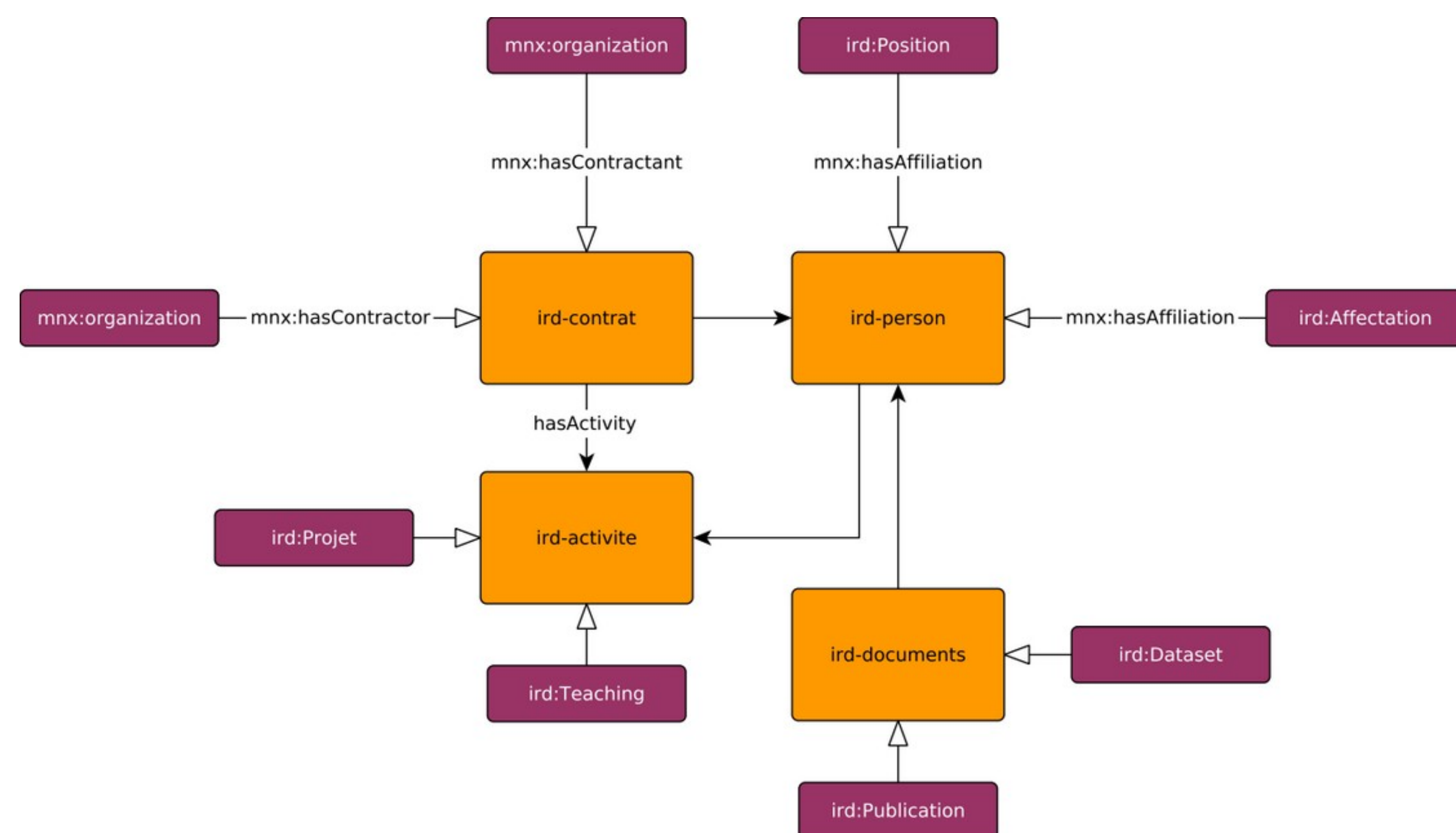


**Producteur de données**  
→ identifier les collaborations  
→ présenter des références à un appel d'offre

## 3. Modélisation simplifiée des données

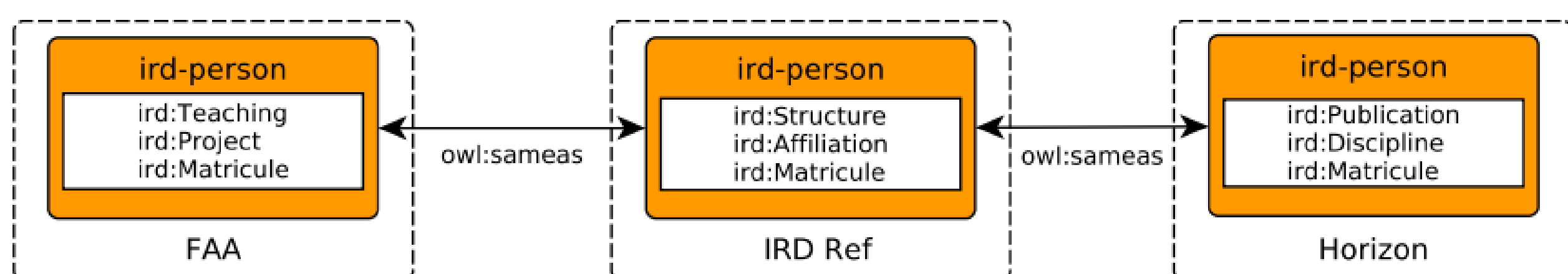
### Quatre entités interconnectées :

- 1) Personnes (chercheurs / ingénieurs)
- 2) Productions (publications / données)
- 3) Contrats
- 4) Activités (enseignement, comités...)



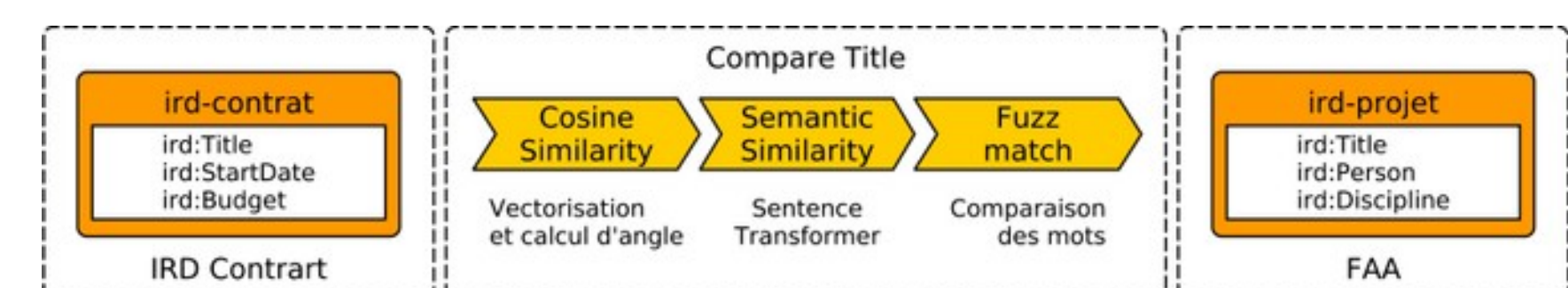
### Données pivot

Les matricules des personnes servent de données pivot entre les publications et les fiches d'activités chercheurs. On note la relation d'équivalence entre les entités ird-person.



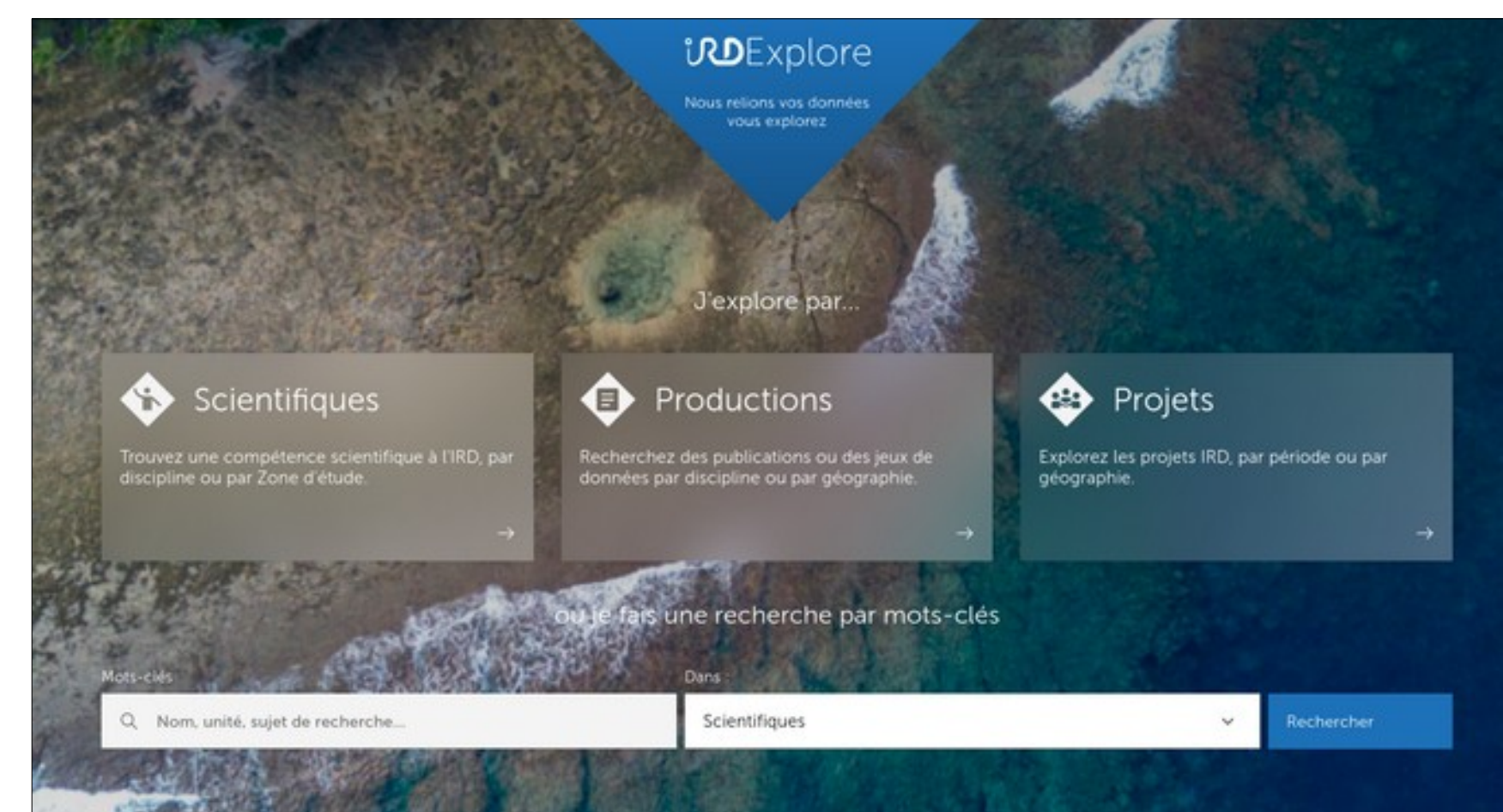
## 4. Implémentation de la base de connaissance

- 1) Création d'un triple Store GraphDB
- 2) Connexion aux bases de données : OAIMPH, GraphQL, scripts scala
- 3) Analyse des données
- 4) Croisement des données (ML)
- 5) Workflow de mise à jour (AirFlow)



Alignement des titres projets et contrats

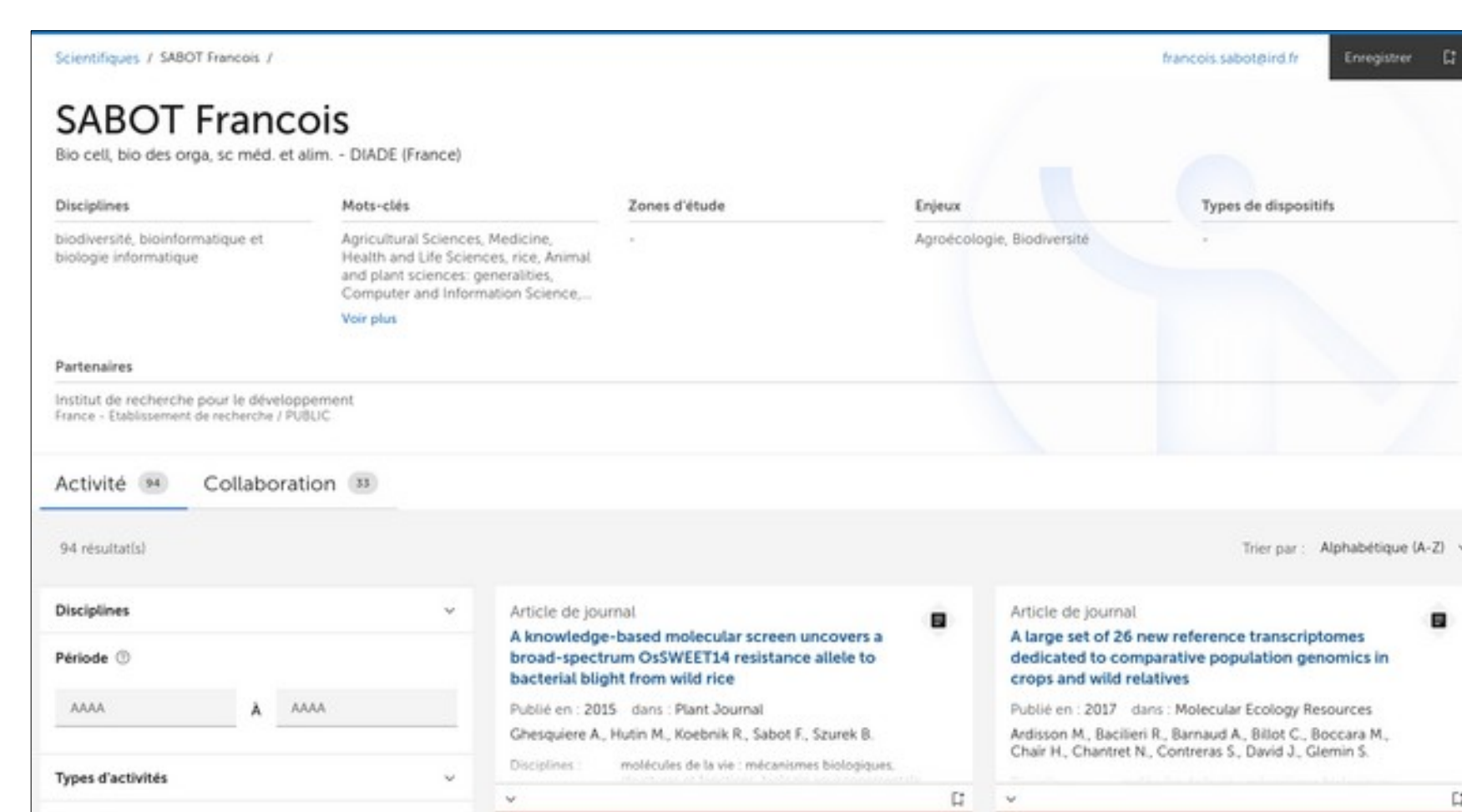
## 5. Portail de consultation



3 entrées depuis la page d'accueil

- Scientifiques
- Productions
- Projets

- Affichage de tous les résultats
- Tri par facettes :
  - Disciplines
  - Zones géographiques
  - Types d'activités...



Page d'un scientifique

- CV
- Onglet activités
- Onglet collaborations

Encart profil similaire



## 6. Bilan

- Création d'un graphe de connaissance IRD et de son API
- Liaison entre les productions / compétences / projets
- Appropriation de la solution technique : stack web sémantique
- Module cartographique peu utilisable
- Données financières non consolidées
- Liaison projets / contrats faible
- Intégration au SI complexe

## 7. Perspectives / évolutions

- Aligner le référentiel géographique Horizon avec le référentiel Unesco
- Proposer des fonctionnalités de filtres avancés (ET / OU)
- Ajouter d'autres sources de données :
  - HAL
  - Fiches de compétences des commissions scientifiques
  - Productions logicielles de la forge
- Créer / utiliser un référentiel projets

