

# The ANR SO-DRIIHM project:

To guide and assist an interdisciplinary community in searching and depositing data in a jungle of available repositories

E. Lerigoleur, P. Dayre



“Driving open science within the LabEx DRIIHM” (2020-2024)



**USE CASE: PLATEFORME "SCIENCE OUVERTE" POUR UNE COMMUNAUTÉ INTERDISCIPLINAIRE**

**USE CASE: PLATEFORME "SCIENCE OUVERTE" POUR UNE COMMUNAUTE INTERDISCIPLINAIRE**

**Coordination :** UMR 5602 GEODE (Géographie de l'Environnement), unité mixte CNRS (INEE) / Université Toulouse 2 Jean Jaurès

**Contact :** Emilie Lerigoleur

**Partenaires académiques pluridisciplinaires :**

- (Informatique) l'IRIT
- (Communauté d'utilisateurs et communauté de support) LabEx DRIIHM
- 13 Observatoires Hommes-Milieus (OHM), dont 6 sont en métropole, 2 en Outre-mer

<https://datanoos.univ-toulouse.fr/fr/use-case-plateforme-science-ouverte-pour-une-communaute-interdisciplinaire>

**DATA NOOSPHERE**

- USE CASE: MODÉLISATION AÉRODYNAMIQUE ENRICHIE ET INTELLIGENTE
- USE CASE: PERSONNALISATION DE DONNÉES ENVIRONNEMENTALES
- INACCESSIBLE
- USE CASE: PLATEFORME "SCIENCE OUVERTE" POUR UNE COMMUNAUTÉ INTERDISCIPLINAIRE
- DATACORE: GROUPE D'EXPERTISES TRANSDISCIPLINAIRES
- ANIMATIONS

# Context: The LabEx DRIIHM and its 13 Observatories Hommes-Milieux (OHM)



Pardo 2019

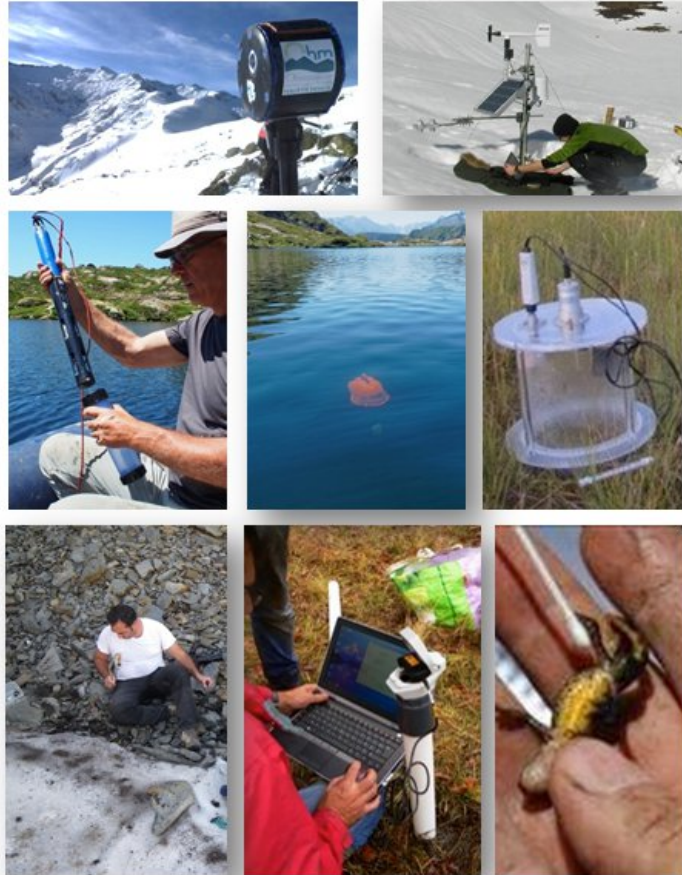
CNRS-INEE research network since 2007, funded by LabEx since 2012

- ~ **1250 scientists** / 200 labs / 133 EPST-universities-graduate schools
- 866 projects funded (5 to 15k€)
- 73 disciplinary fields
- very heterogeneous data from the "long tail"
- Staff : 1 director + 1 responsible for governance + 1 computer scientist (non-permanent staff) + **data-DRIIHM** group (~10 pers.)

**A data infrastructure built brick by brick**  
geocatalogs, web GIS, photo library

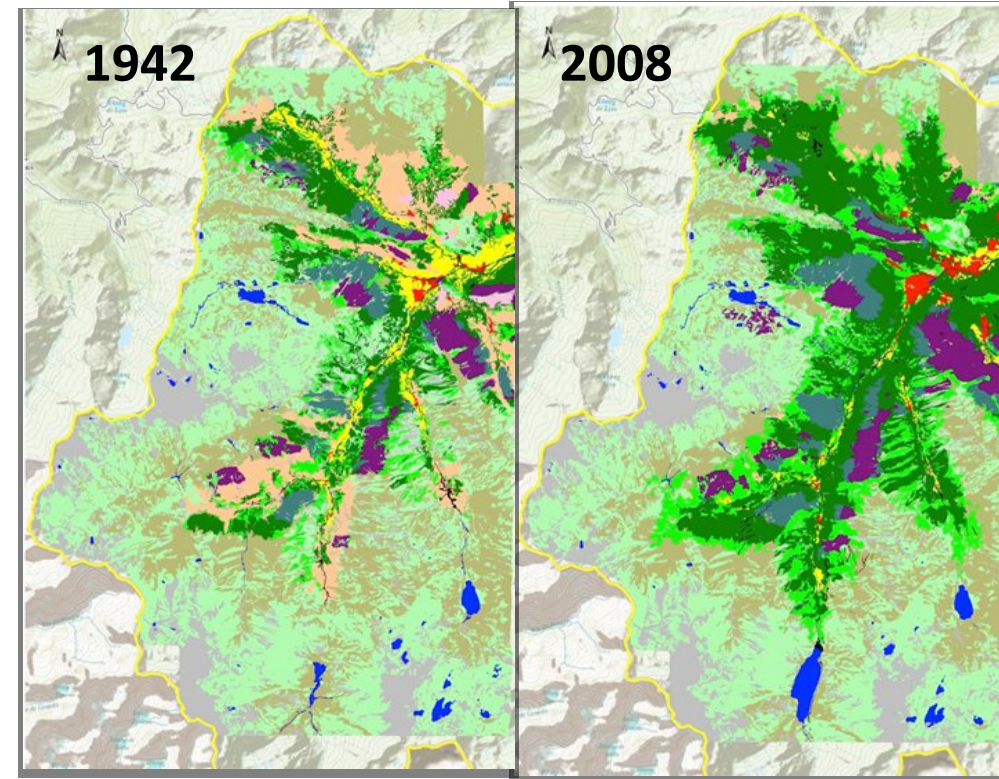
# Context: The different activities and types of data collected/generated

Example: OHM Pyrénées Haut Vicdessos



Galop 2016

- Hydro-climatic monitoring
- Lake monitoring
- Snow cover monitoring
- Vegetation monitoring
- Floristic and faunistic inventories
- Socio-economic surveys
- Paleoenvironmental studies
- Archaeology
- Historical and eco-historical studies
- Remote sensing



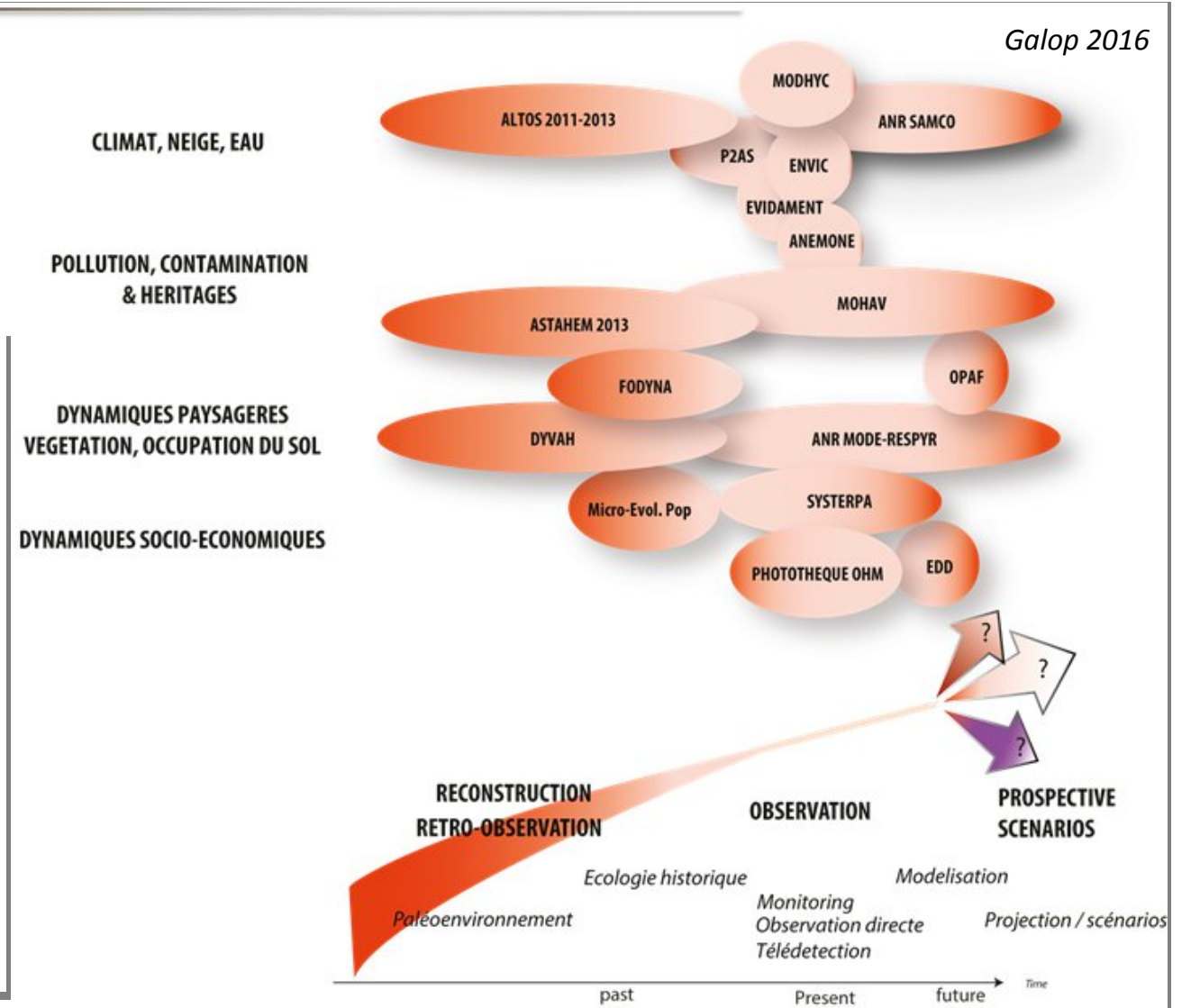
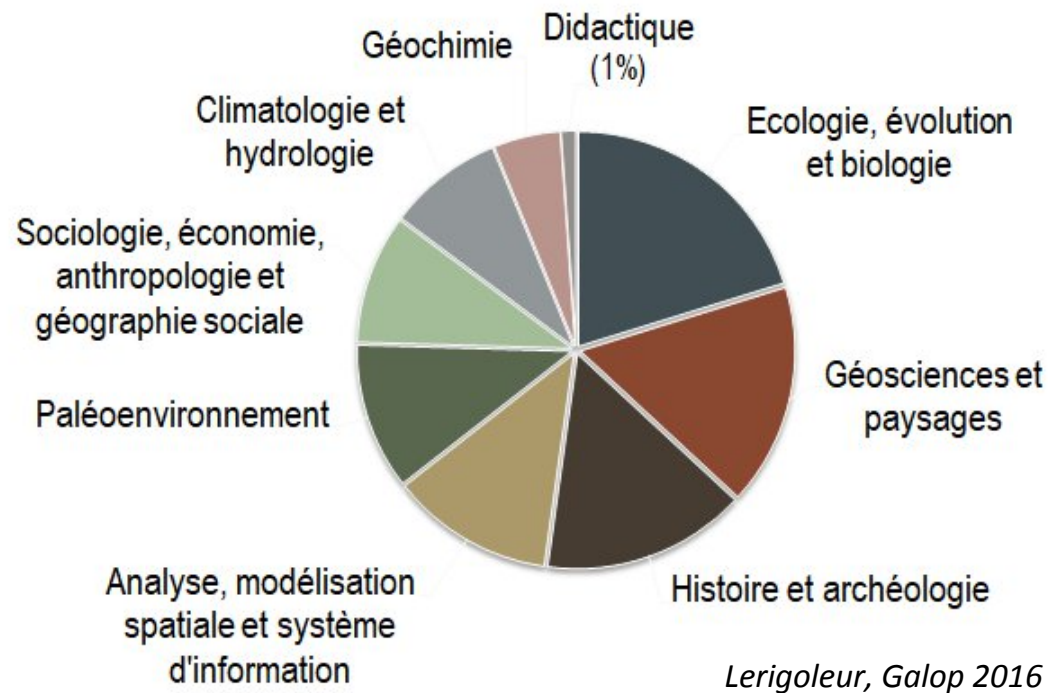
Houet 2014

# Context: Interdisciplinarity with interconnected research projects

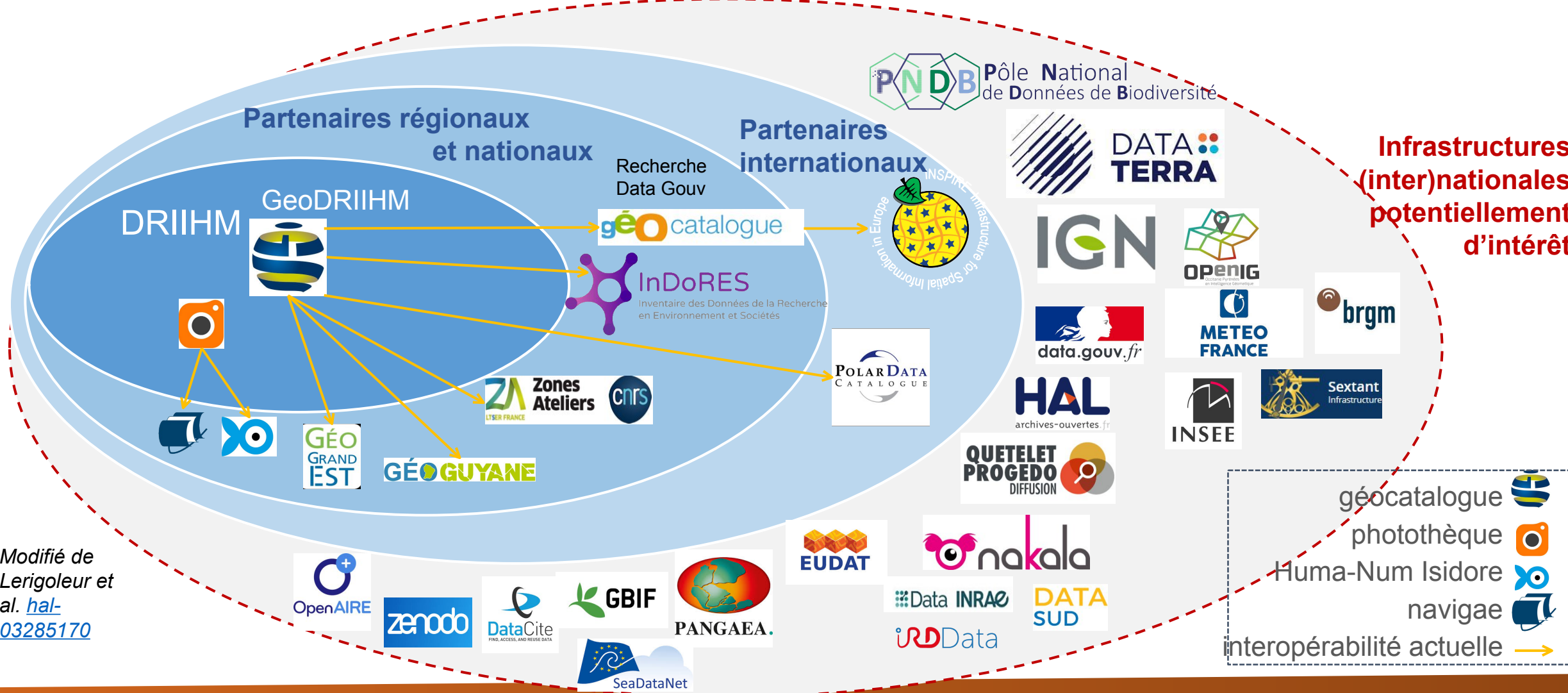
**Example : 18 annual projects funded by the OHM Pyrénées Haut Vicdessos**

→ many disciplines

→ 3 temporal scales



# Complex interoperability: how to connect the OHM platform with disciplinary and institutional e-infrastructures?



Modifié de  
Lerigoleur et  
al. [hal-03285170](https://hal.archives-ouvertes.fr/hal-03285170)

# ANR SO-DRIIHM project (oct. 2020 - mars 2023 2024)

**Findings:** dispersion of OHM publications, poor data sharing, potential loss of data, non-ergonomic cataloguing tool, poor interdisciplinarity

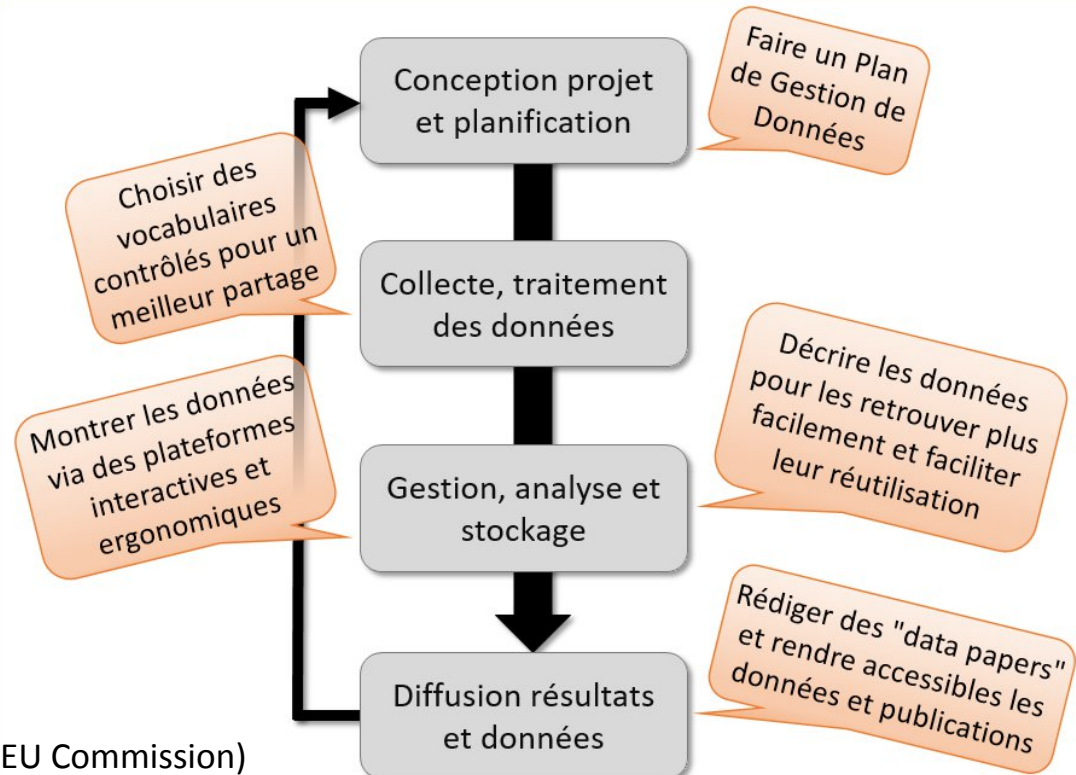
2 usage surveys (2016 and 2018) + **article** about the state of play [Arnaud et al 2020](#)

## Agir tout au long du cycle de vie des projets de recherche

SO-DRIIHM : 8 partners / 98 k€

2 main aims :

- **Promoting open science:** showing the benefits of sharing and opening data, explaining the FAIR principles
- **Co-construction of an ergonomic and interoperable e-infrastructure,** integrating existing tools and supporting researchers in the "FAIRisation" of data



"As open as possible, as closed as necessary" (H2020 EU Commission)

# SO-DRIIHM : first actions

21

productions  
dans HAL

- 1 data paper
- 14 com. orales
- 3 posters
- 3 autres productions



- 8 communications séminaires DRIIHM et OHM/OHMi
- 1 café du DRIIHM RGPD\* et PGD\*
- animation d'ateliers
- interventions dans séminaires externes et école thématique

Interventions dans des séminaires et animation d'événements



Enquête sur les usages

- 107 répondants
- 1 synthèse [hal-03697372](https://hal.archives-ouvertes.fr/hal-03697372)
- 1 dépôt des données pseudonymisées [DOI:10.34847/nkl.db933xfo](https://doi.org/10.34847/nkl.db933xfo)

FAIR

Promotion de la science ouverte et co-conception du portail DATADRIIHM

Outil de recherche de ressources sur la science ouverte ORRSO

- 1 projet tutoré Master "Information et Médiation Scientifique et Technique" Lyon 1
- 1 stage pour la création de l'outil ORRSO  
→ Martin-Mennetrier *et al.* 2022 [hal-03720513](https://hal.archives-ouvertes.fr/hal-03720513)  
Resp.: C. Pardo, ML. Trémélo, M. Massaviol

Accompagnements de projets pilotes

- Formation au PGD\* et suivi
- Travail de standardisation des données pour OHM Lit. Caraïbe, Oyapock et Vicdessos
- Dépôt des données dans les entrepôts et catalogage des métadonnées associées
- Conseils RGPD (post-docs et programmes transverses)

Conception, maquettes et prototypes

- 1 stage
- 1 groupe de travail
- Partenariat avec [datanoos](https://datanoos.org/)
- Recherche de financements complémentaires



# SO-DRIIHM : first actions

21

productions  
dans HAL

- 1 data paper
- 14 com. orales
- 3 posters
- 3 autres productions



- 8 communications séminaires DRIIHM et OHM/OHMi
- 1 café du DRIIHM RGPD\* et PGD\*
- animation d'ateliers
- interventions dans séminaires externes et école thématique

Interventions dans des séminaires et animation d'événements



Enquête sur les usages

- 107 répondants
- 1 synthèse [hal-03697372](https://hal.archives-ouvertes.fr/hal-03697372)
- 1 dépôt des données pseudonymisées [DOI:10.34847/nkl.db933xfo](https://doi.org/10.34847/nkl.db933xfo)

FAIR

Outil de recherche de ressources sur la science ouverte ORRSO

- 1 projet tutoré Master "Information et Médiation Scientifique et Technique" Lyon 1
- 1 stage pour la création de l'outil [ORRSO](https://hal.archives-ouvertes.fr/hal-03720513)  
→ Martin-Mennetrier *et al.* 2022 [hal-03720513](https://hal.archives-ouvertes.fr/hal-03720513)  
Resp.: C. Pardo, ML. Trémélo, M. Massaviol

Promotion de la science ouverte et co-conception du portail [DATADRIIHM](https://datadriihm.fr)

Accompagnements de projets pilotes

- Formation au PGD\* et suivi
- Travail de standardisation des données pour OHM Lit. Caraïbe, Oyapock et Vicdessos
- Dépôt des données dans les entrepôts et catalogage des métadonnées associées
- Conseils RGPD (post-docs et programmes transverses)

Conception, maquettes et prototypes

- 1 stage
- 1 groupe de travail
- Partenariat avec [datanoos](https://datanoos.fr)
- Recherche de financements

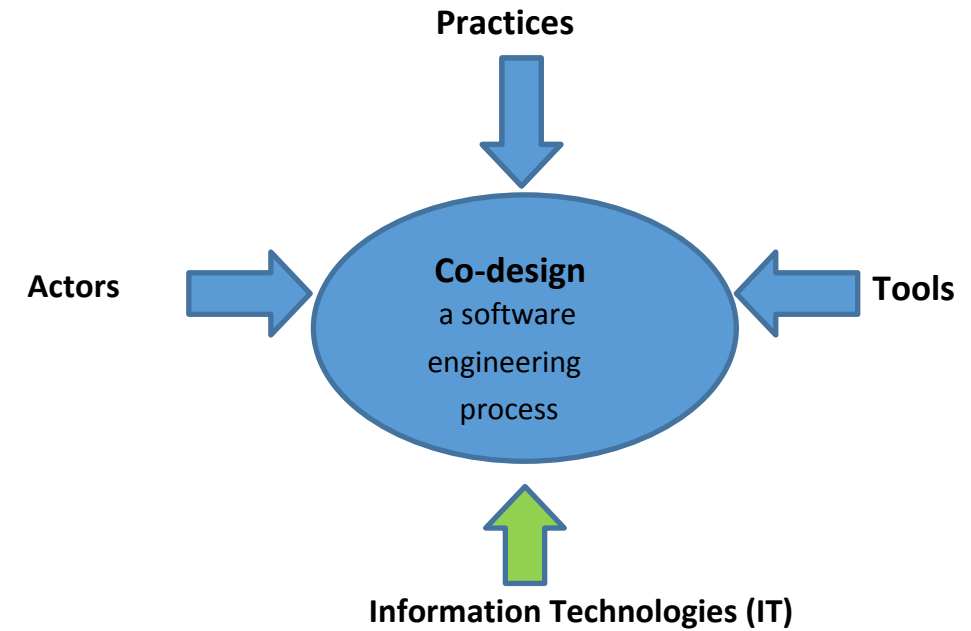
==> Plan France Relance Emplois R&D  
Collab. IRIT-Autonoms (Makina corpus)  
Nov. 2022 - Oct. 2024  
2 CDD dev. back-end + front-end

# SO-DRIIHM : zoom on the co-design of the DATA-DRIIHM platform

## Co-design software engineering process

« The aim of this process is to manage dynamic and integrative view on tools and practices to elaborate expectations, requirements and project management. The design approach is the idea of co-evolution of tools, practices and actors. »

*Allert et al., 2007. KP-LAB Knowledge Practices Laboratory... (hal-00593146)*

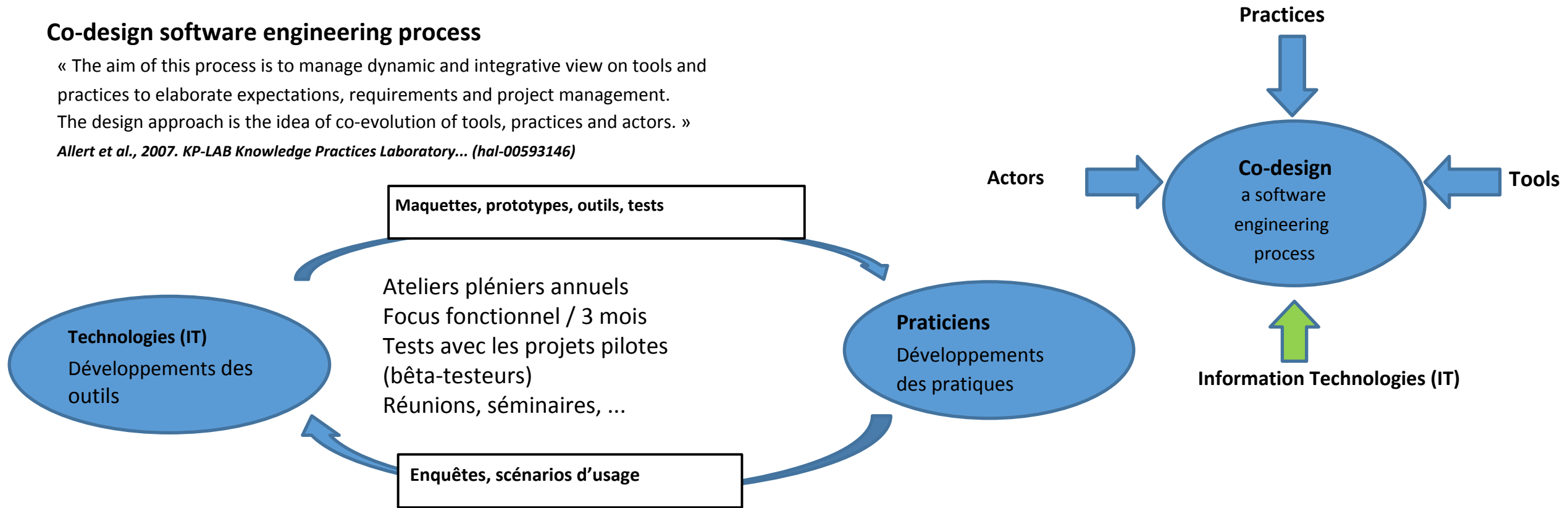


# SO-DRIIHM : zoom on the co-design of the DATA-DRIIHM platform

## Co-design software engineering process

« The aim of this process is to manage dynamic and integrative view on tools and practices to elaborate expectations, requirements and project management. The design approach is the idea of co-evolution of tools, practices and actors. »

*Allert et al., 2007. KP-LAB Knowledge Practices Laboratory... (hal-00593146)*



- Maquettage de l'interface utilisateur / espace personnel
- Ateliers avec les utilisateurs pour affiner les scénarios d'usage de la plateforme et écrire de nouvelles spécifications
- Développement d'un prototype « Prove Of Concept » de l'architecture technique
- Développement d'un démonstrateur des scénarios de recherche et de dépôts de données

# SO-DRIIHM : zoom on the co-design of the DATA-DRIIHM platform

A multi-scale co-design process

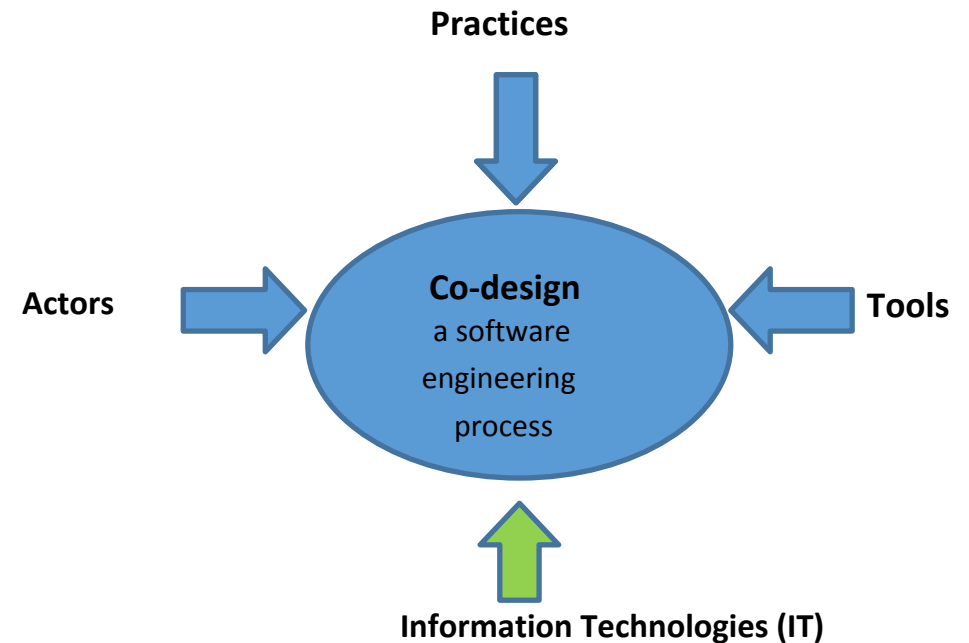
## A doubly integrated eco-system:

- integration of digital tools
- integration of actors

## The actors:

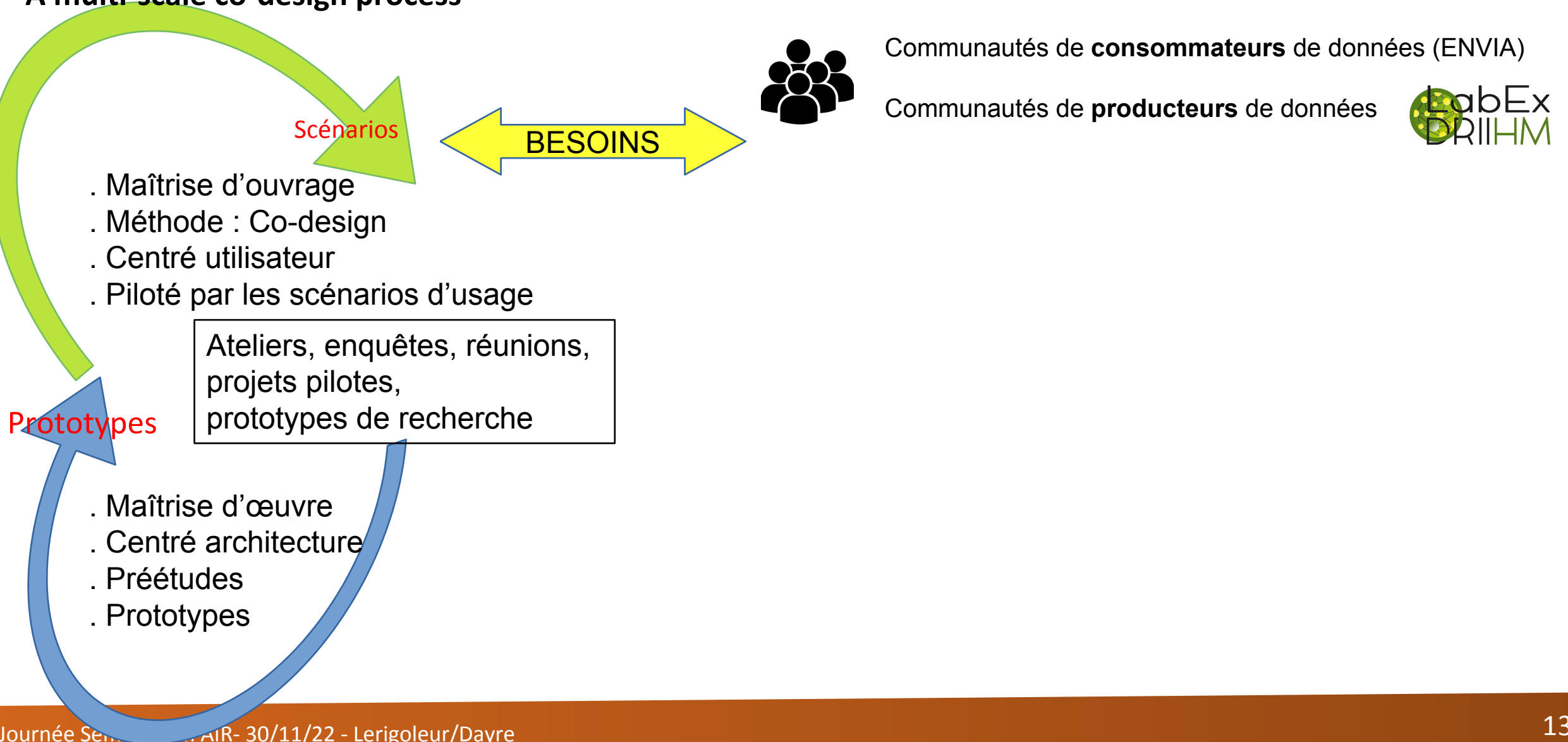
- end users (1250 researchers)
- support staff (10 engineers)
- administrators
- developers (frontend, backend, web data)
- the IT research community

=> A multi-scale co-design process



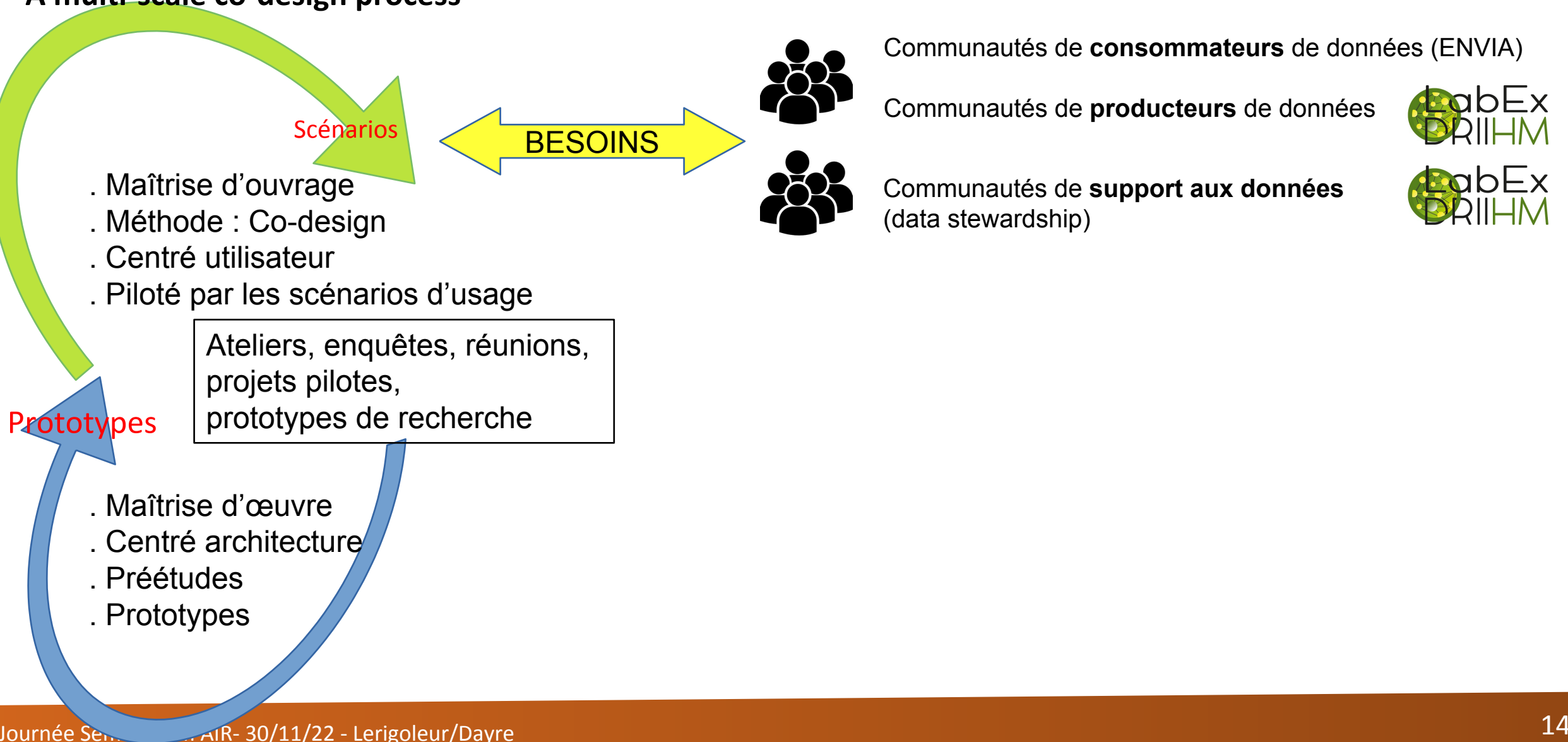
# SO-DRIIHM : zoom on the co-design of the DATA-DRIIHM platform

A multi-scale co-design process



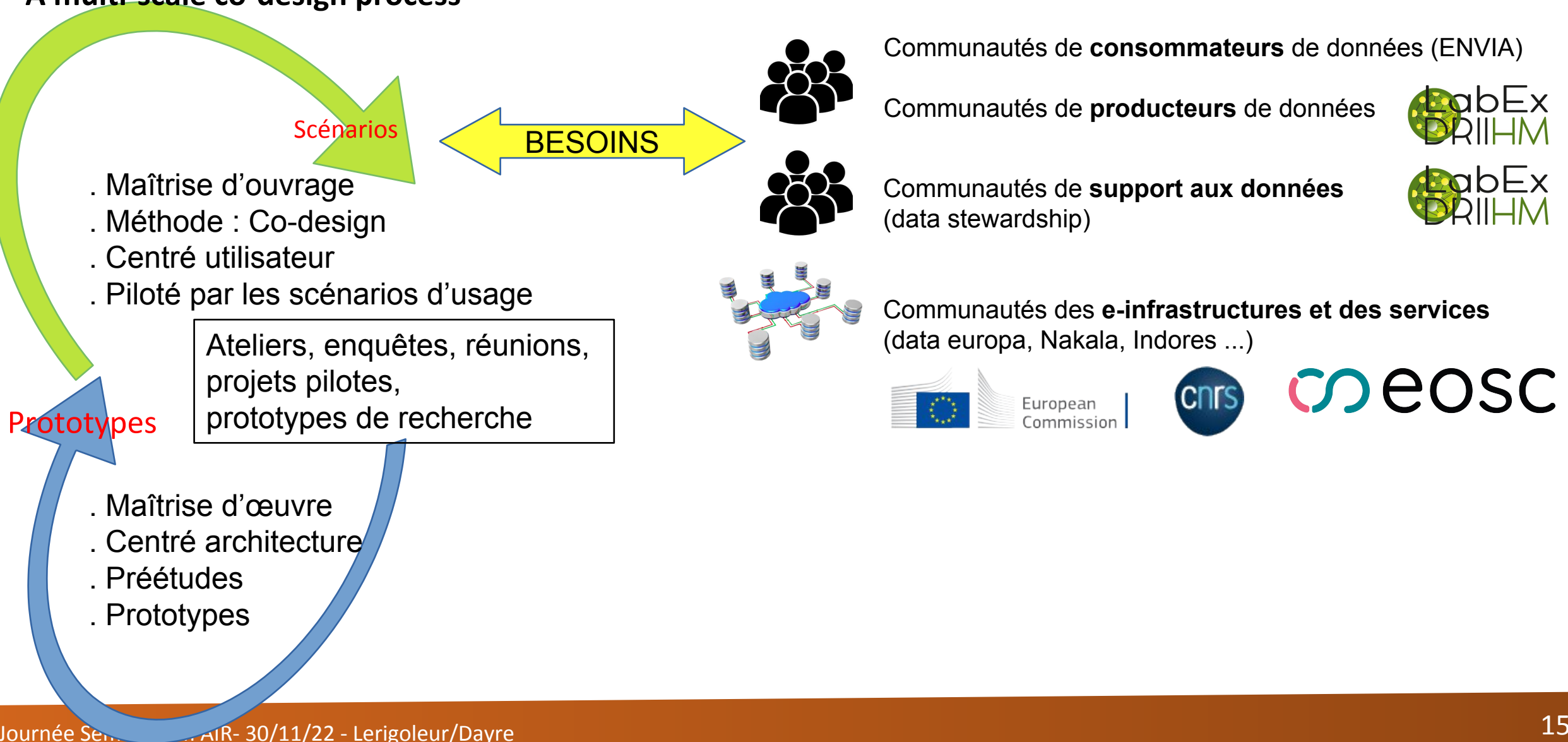
# SO-DRIIHM : zoom on the co-design of the DATA-DRIIHM platform

A multi-scale co-design process



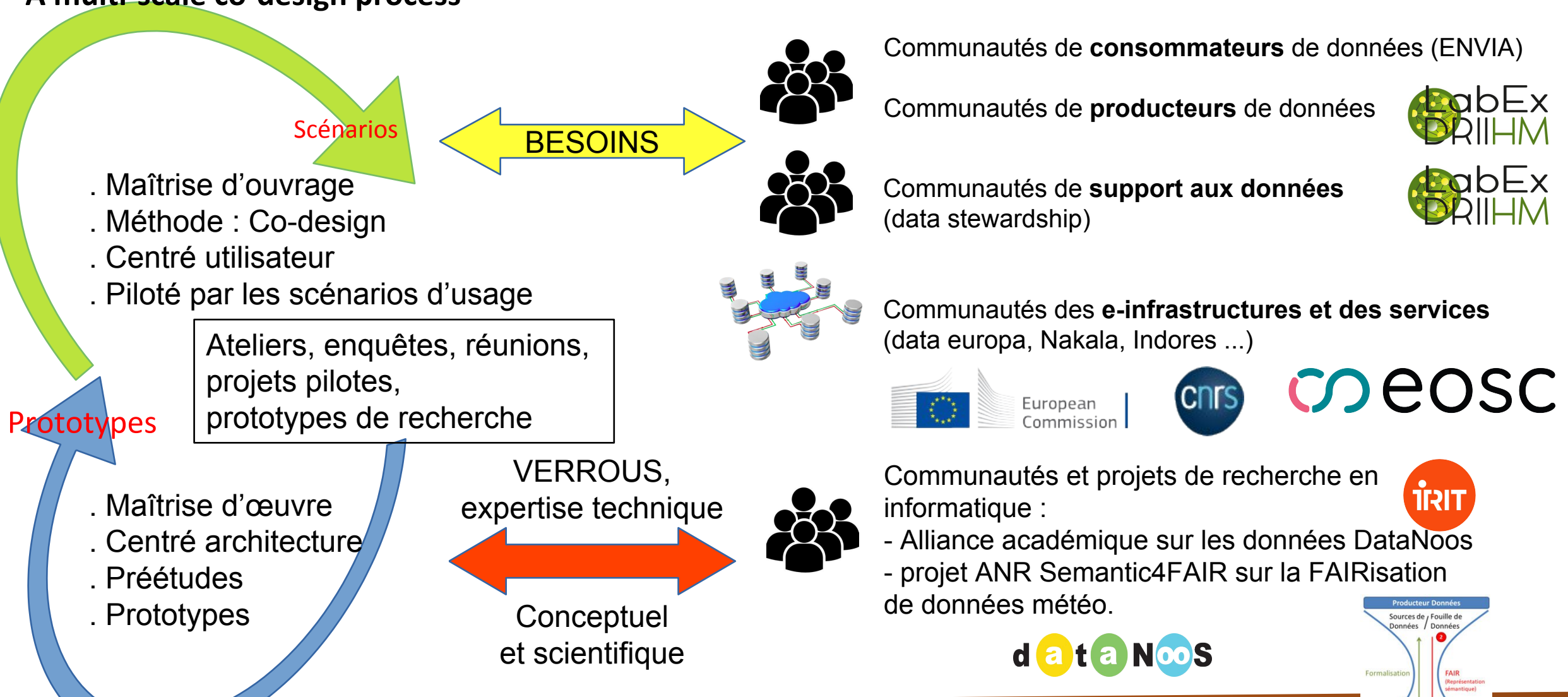
# SO-DRIIHM : zoom on the co-design of the DATA-DRIIHM platform

A multi-scale co-design process



# SO-DRIIHM : zoom on the co-design of the DATA-DRIIHM platform

A multi-scale co-design process



- . Maîtrise d'ouvrage
- . Méthode : Co-design
- . Centré utilisateur
- . Piloté par les scénarios d'usage

- . Maîtrise d'œuvre
- . Centré architecture
- . Prétudes
- . Prototypes

Ateliers, enquêtes, réunions, projets pilotes, prototypes de recherche

VERROUS, expertise technique

Conceptuel et scientifique

Communities of **consommateurs** de données (ENVIA)

Communities of **producteurs** de données

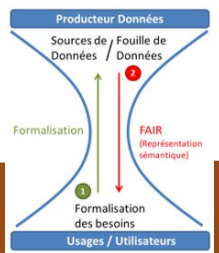
Communities of **support aux données** (data stewardship)

Communities of **e-infrastructures et des services** (data europa, Nakala, Indores ...)

Communities and research projects in **informatique** :

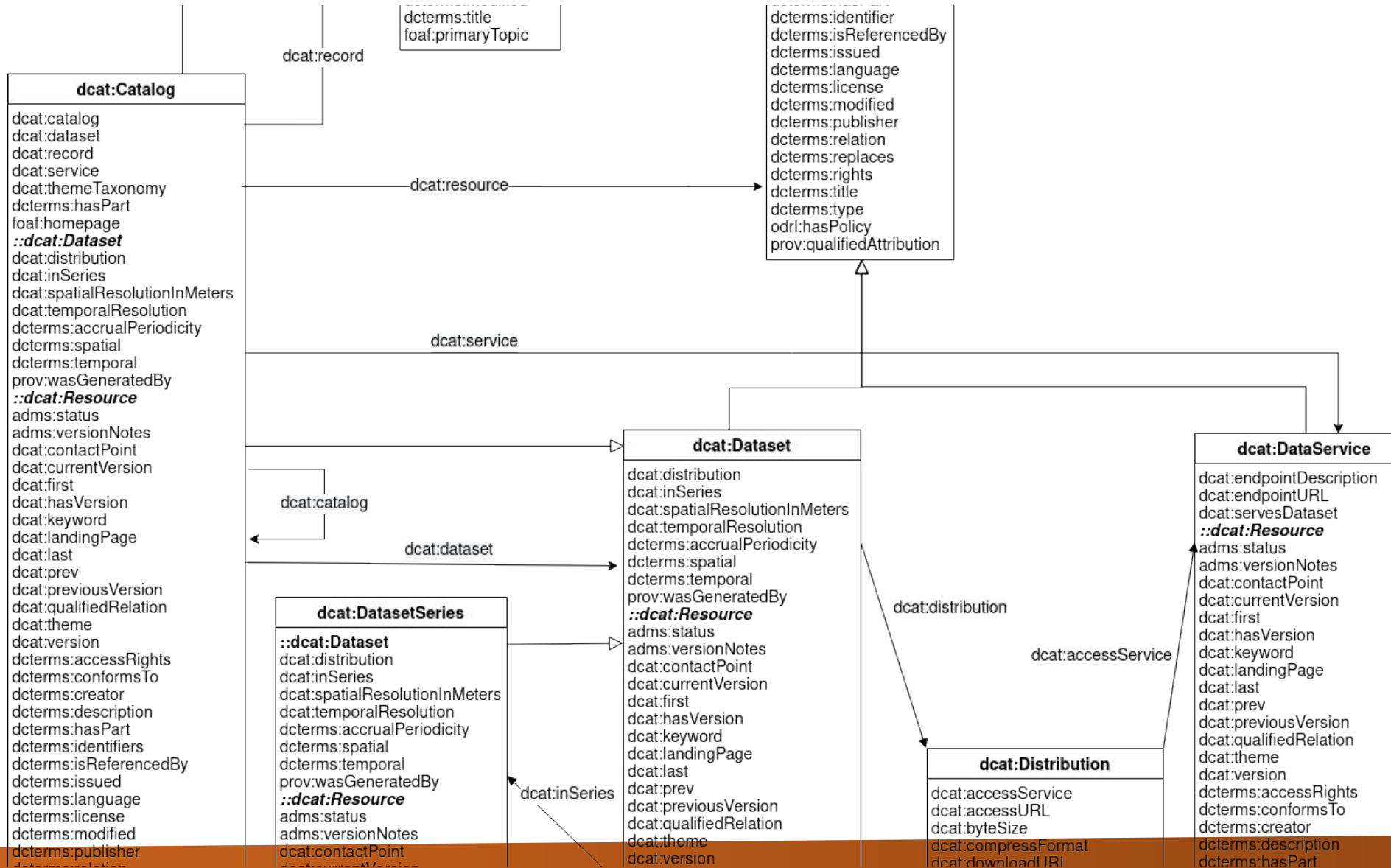
- Alliance académique sur les données DataNoos
- projet ANR Semantic4FAIR sur la FAIRisation de données météo.

**dataNoos**

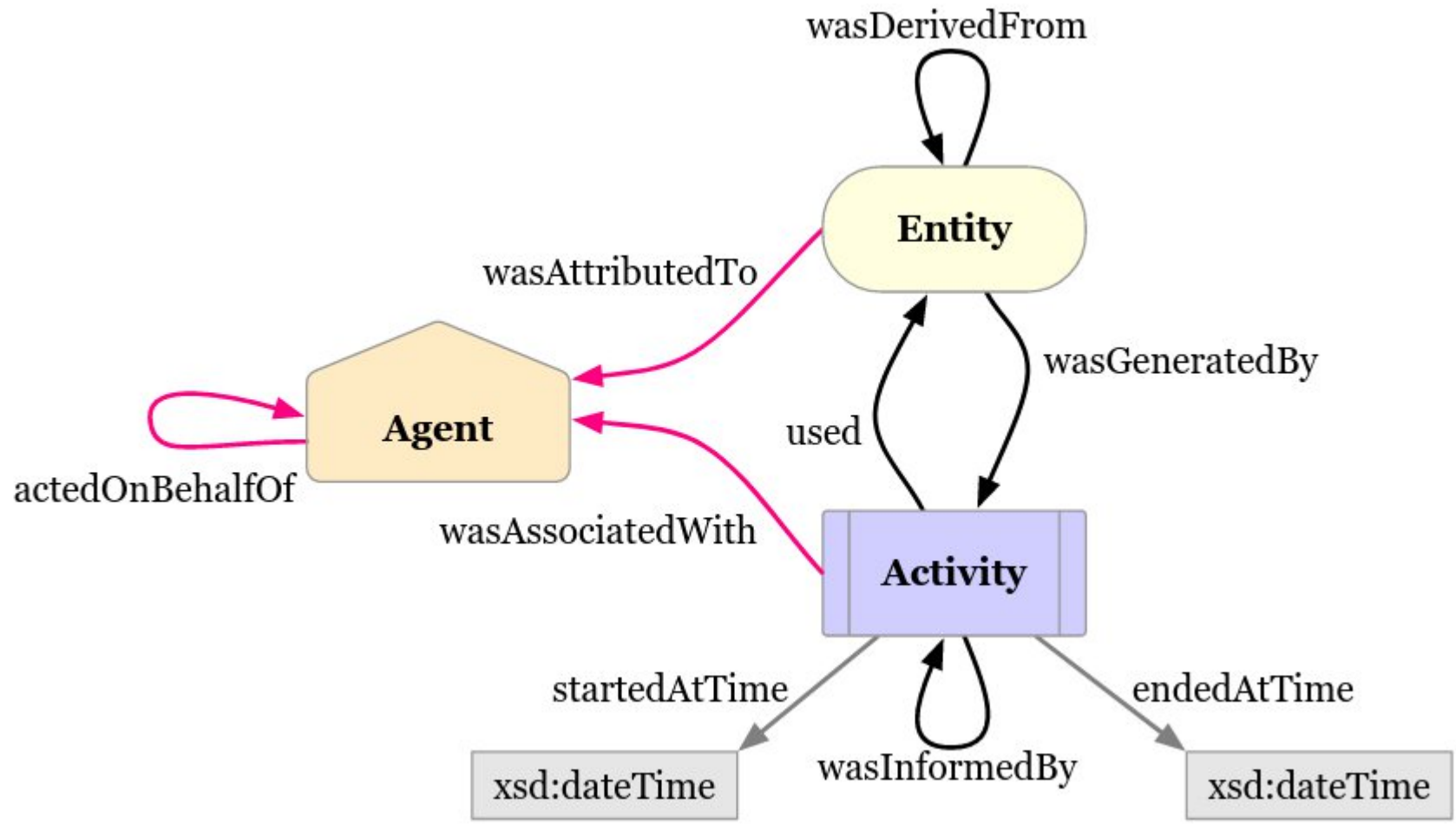




# SO-DRIIHM: geoDCAT-AP used as the standards-based metadata model

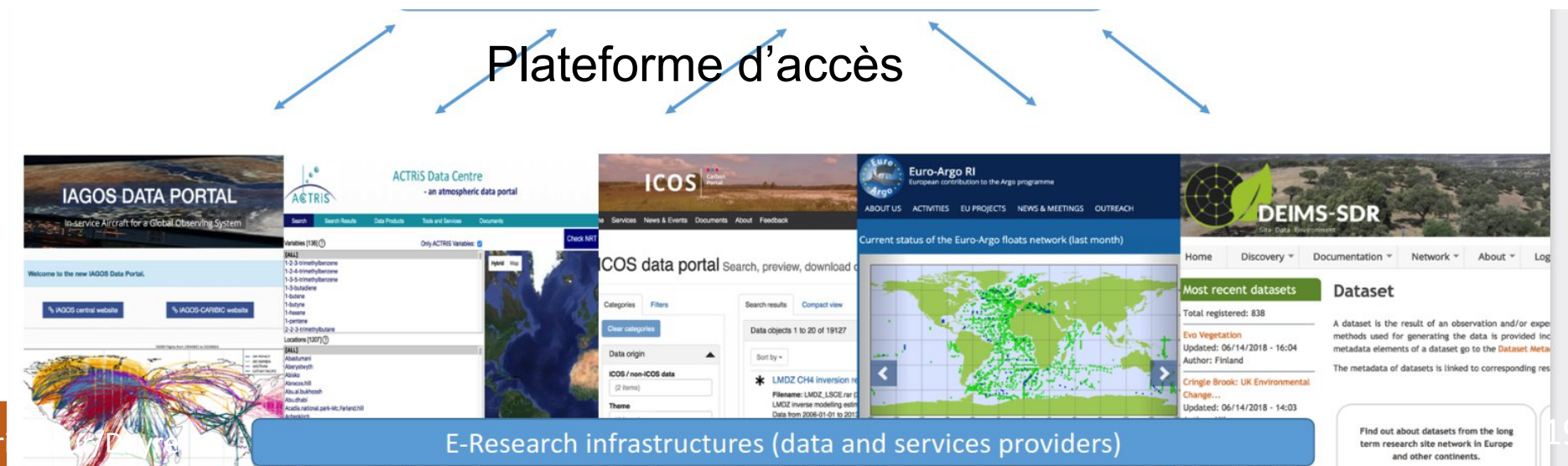


# SO-DRIIHM: PROVO used as the standards-based metadata model



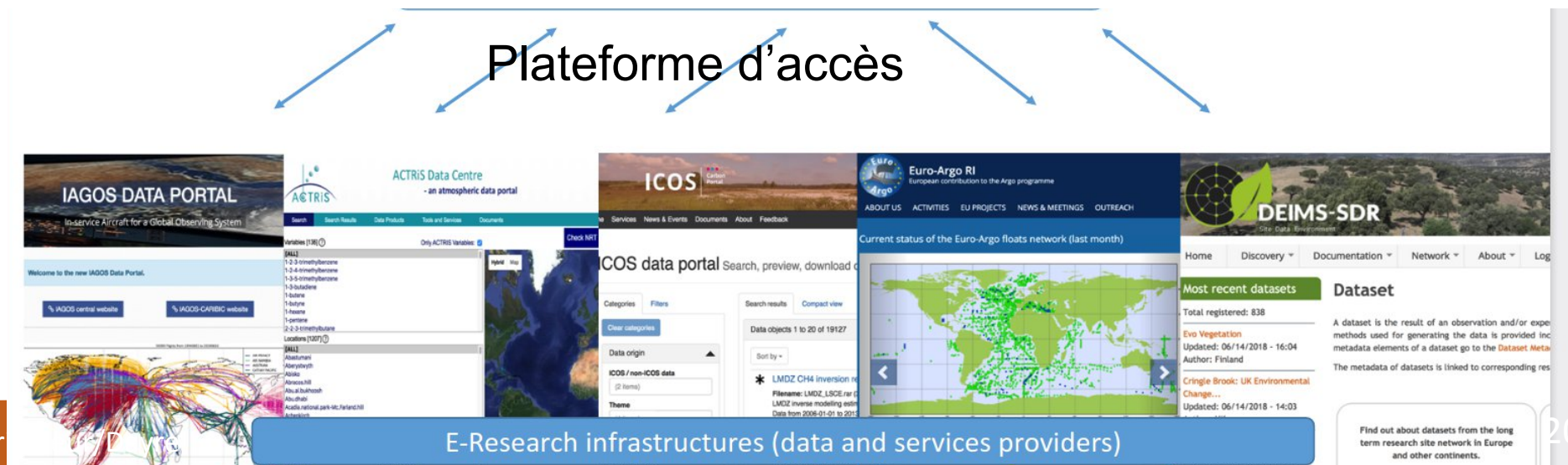
# The service-oriented digital open science eco-system

- Des services d'annuaire de service / de plateforme (registry):
  - pour retrouver des services de catalogage, de stockage de données, de traitement, ...
  - Exemples : [Cat opidor](#), [Re3data](#), [Data.europa.eu](#)
- Des services de catalogage
  - pour exposer et rechercher des jeux de données (décrire, organiser, indexer ) [7]
  - Exemples : la future recherche.data.gouv.fr, data.europa.eu (le portail officiel de l'Europe avec 170 catalogue de producteur de données, de 36 pays, 1,4 millions de jeux de données)

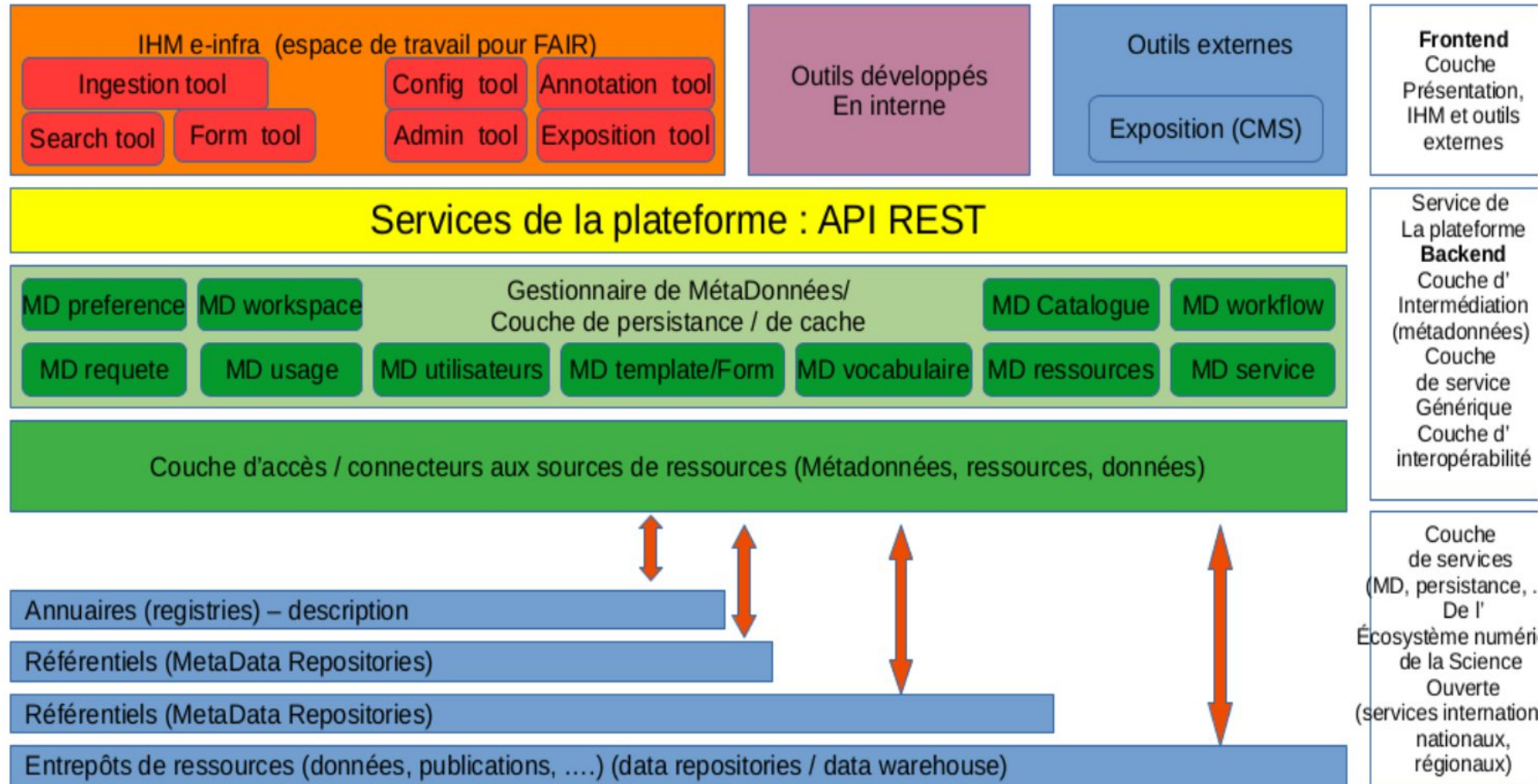


# The service-oriented digital open science eco-system

- Des service d'entrepôt de données (et de distribution)
  - pour stocker physiquement et contrôler les accès (droits, licences) aux données [6]
  - Exemples : [Nakala \(Huma-Num\)](#), [Entrepôt de données de l'INEE \(dataverse InDORES\)](#)
- Les plateformes d'accès aux ressources de la science ouverte
  - pour rendre accessible et enrichir les services à destination des utilisateurs
  - Exemple : [Data Terra](#)
- ...



# SO-DRIIHM : a layered service-oriented architecture and micro-services



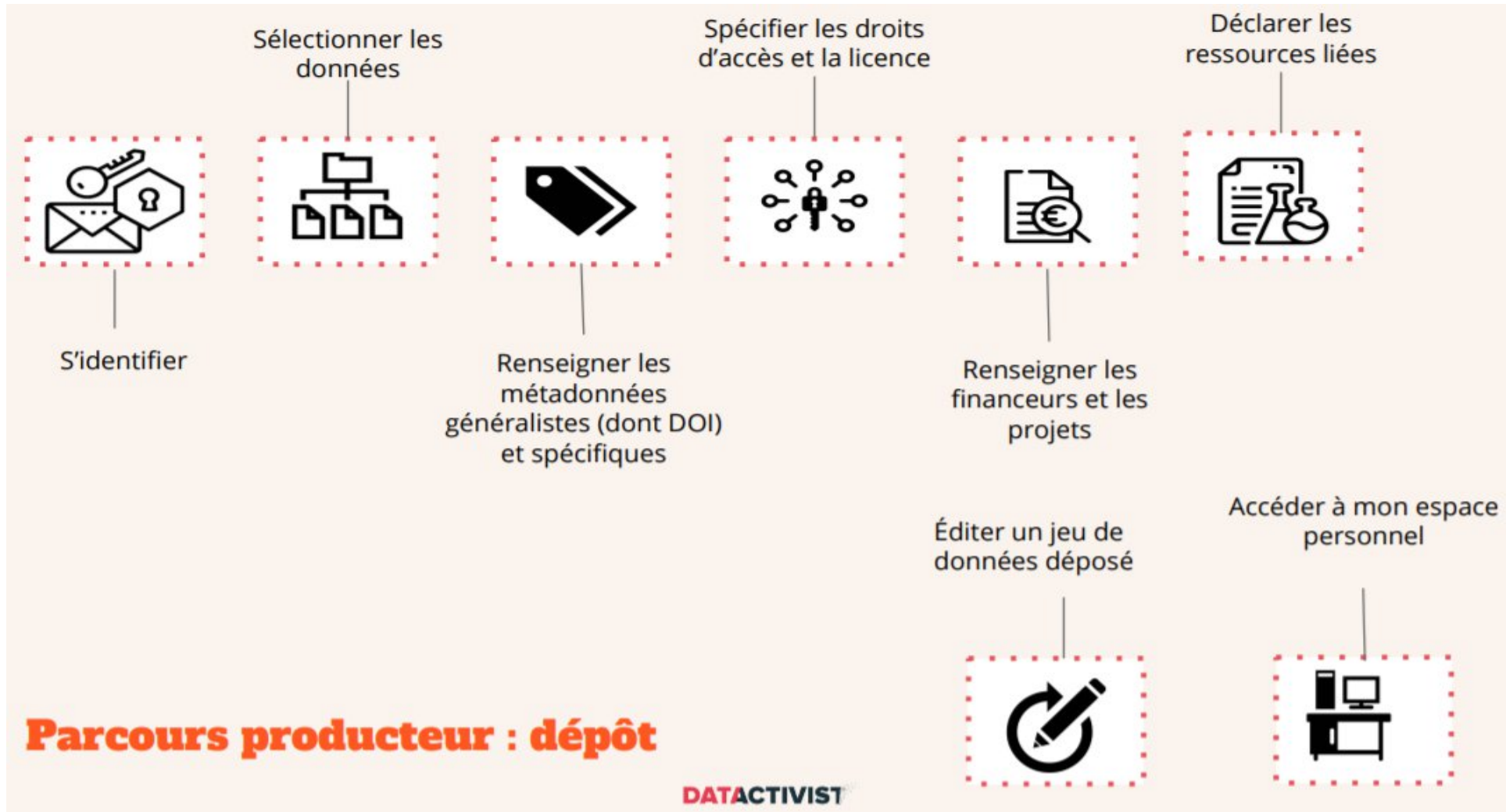
DN.DC20210511

pascal.dayre@irit.fr

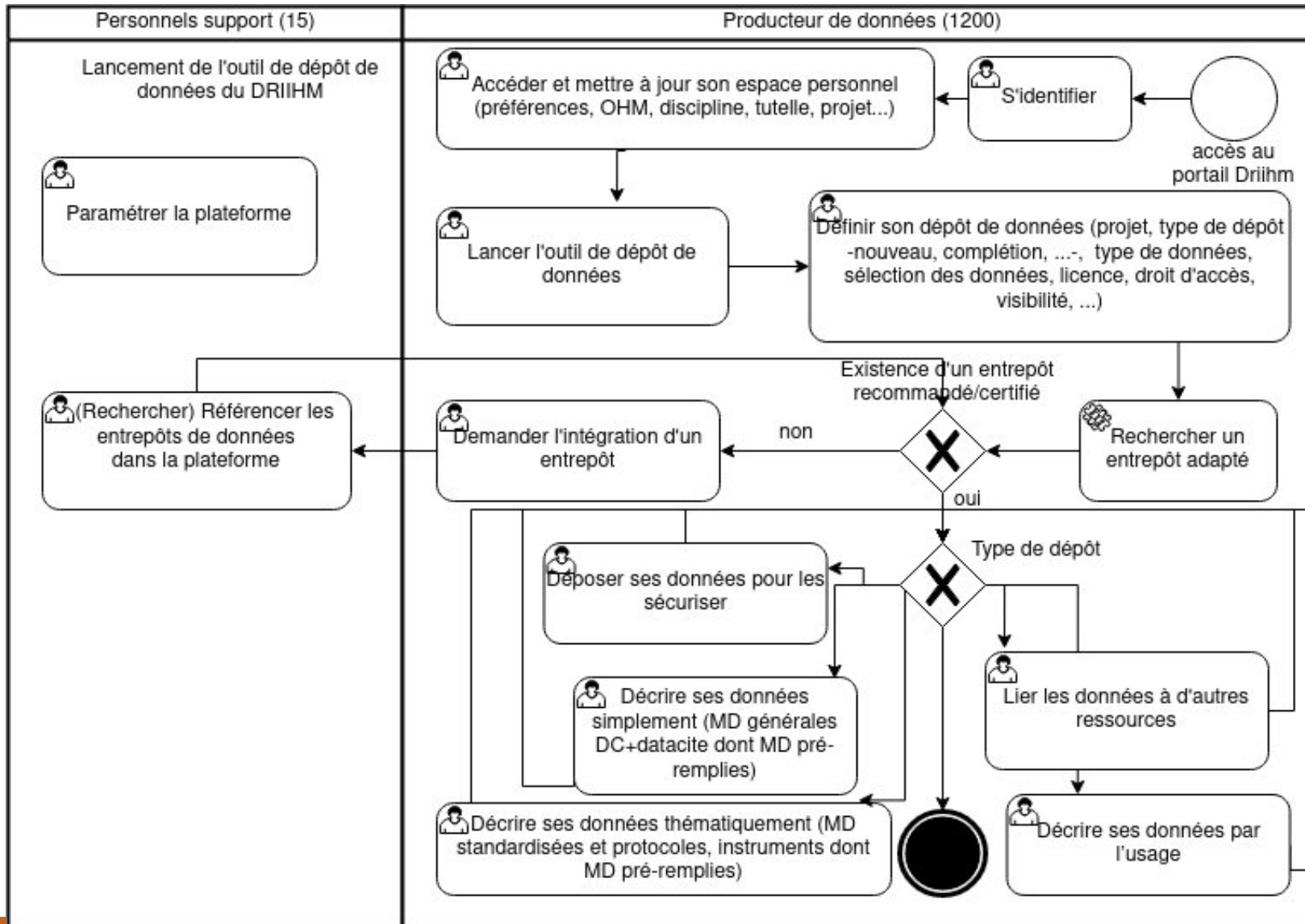
2

# SO-DRIIHM: a process-oriented platform

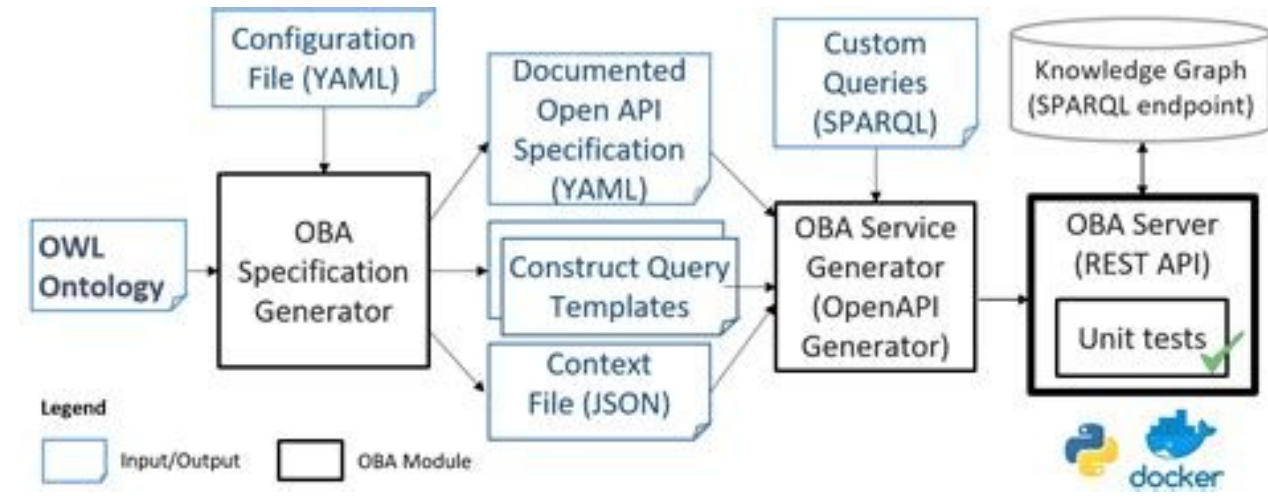
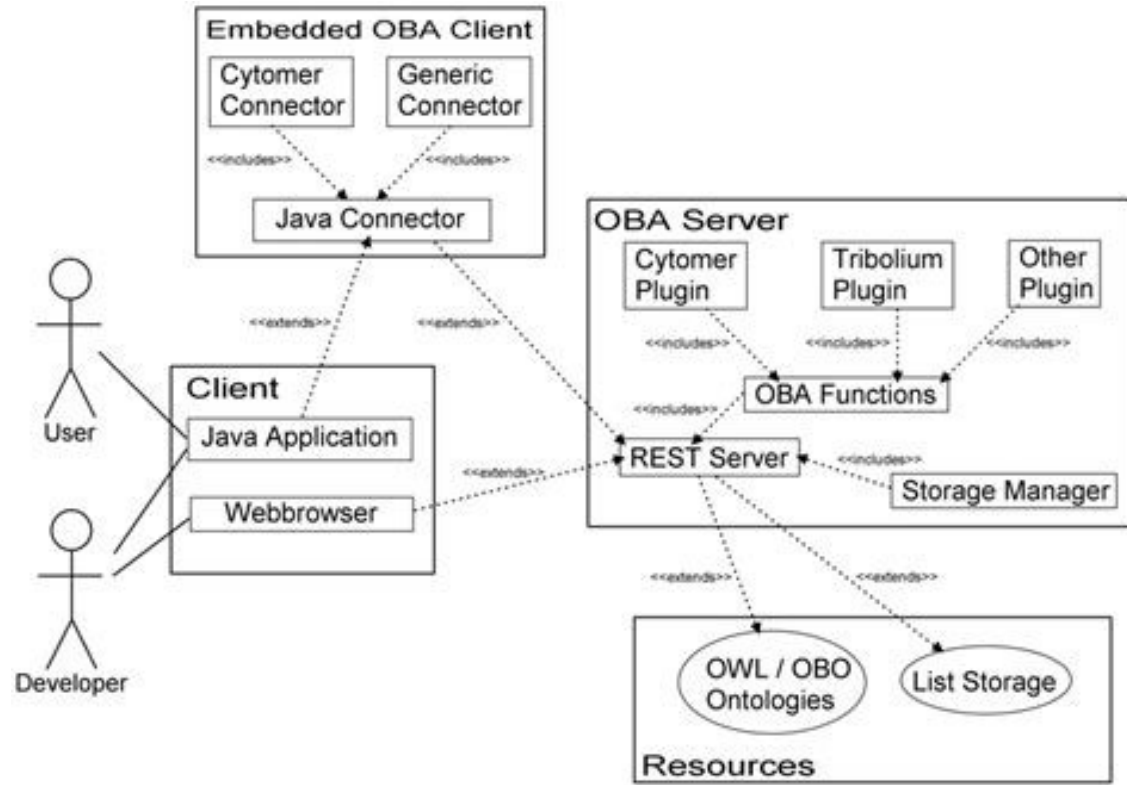
Illustration of the data repository process according to the Committee for Open Science (CoSO)



# SO-DRIIHM: a process-oriented platform



# SO-DRIIHM: OBA (Ontology Based API) oriented development



Garijo, Daniel & Osorio, Maximiliano. (2020). OBA: An Ontology-Based Framework for Creating REST APIs for Knowledge Graphs.



# DATA-DRIIHM: an open source platform

## Call for contributions and collaboration

- development of the core platform
- integration of existing services
- ...

Contact : [pascal.dayre@irit.fr](mailto:pascal.dayre@irit.fr)