



**FAIR-IMPACT**

Expanding FAIR solutions across EOSC

# FAIR Implementation Workshop : French Data Management Clusters

## Recherche Data Gouv, lessons learned from INRAE

Dimitri SZABO, INRAE, April 2024



Funded by  
the European Union

# INRAE in a nutshell

France's National Research Institute for Agriculture, Food and Environment

## Structure

**18**

centers

**14**

scientific divisions

**267**

research, experimental research and support units

## Staff\*

**8229**

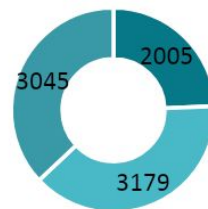
permanent staff

**2798**

contractual staff

**11027**

total



■ researchers  
■ engineers and assistant engineers  
■ technicians

\* 2021 numbers



In  eosoc

**FAIR-IMPACT**  
Expanding FAIR solutions across EOSC

- WP4 & T4.5 lead
- Participation in every WP

**INRAE**

Recherche Data Gouv, lessons learned from INRAE

April 2024 — Dimitri Szabo

# Recherche Data Gouv before Recherche Data Gouv

## Global view of the ecosystem

### 19+ DATA MANAGEMENT CLUSTERS

*Generalist expertise which is geographically close to the research teams it supports on all data issues*

### 4 INSTITUTIONAL REFERENCE CENTRES

*Provide specific guidance according to institutional policies and strategies*

### 1 DATA REPOSITORY

*A mutualised repository for all institutions to deposit and publish data*

### 6+ THEMATIC REFERENCE CENTRES

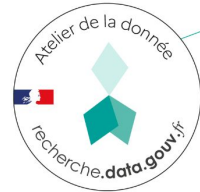
*Expertise in a given scientific field*

### 4 RESOURCE CENTRES

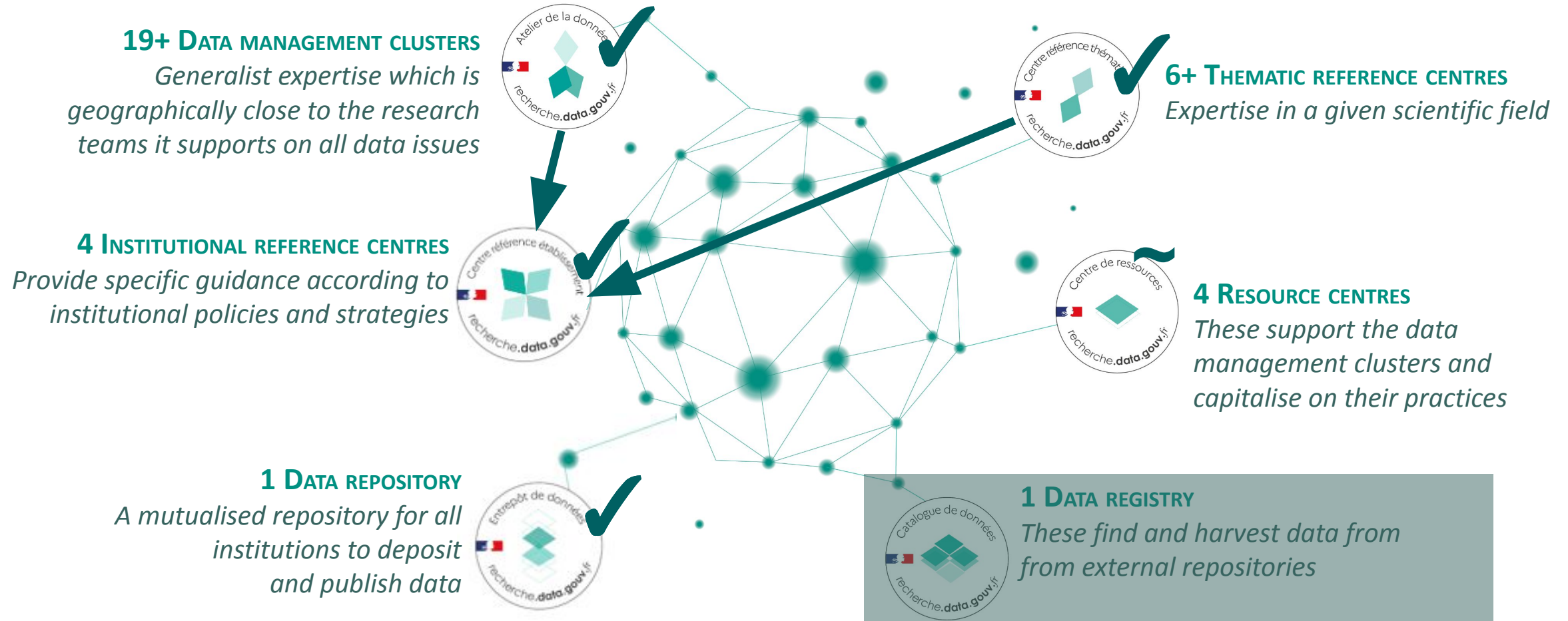
*These support the data management clusters and capitalise on their practices*

### 1 DATA REGISTRY

*These find and harvest data from from external repositories*



## Global view of the ecosystem with existing previous services at INRAE



## Global view of the ecosystem : Data repository

### 19+ DATA MANAGEMENT CLUSTERS

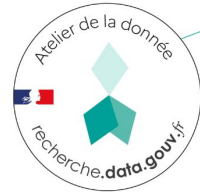
*Generalist expertise which is geographically close to the research teams it supports on all data issues*

### 4 INSTITUTIONAL REFERENCE CENTRES

*Provide specific guidance according to institutional policies and strategies*

### 1 DATA REPOSITORY

*A mutualised repository for all institutions to deposit and publish data*



### 6+ THEMATIC REFERENCE CENTRES

*Expertise in a given scientific field*

### 4 RESOURCE CENTRES

*These support the data management clusters and capitalise on their practices*

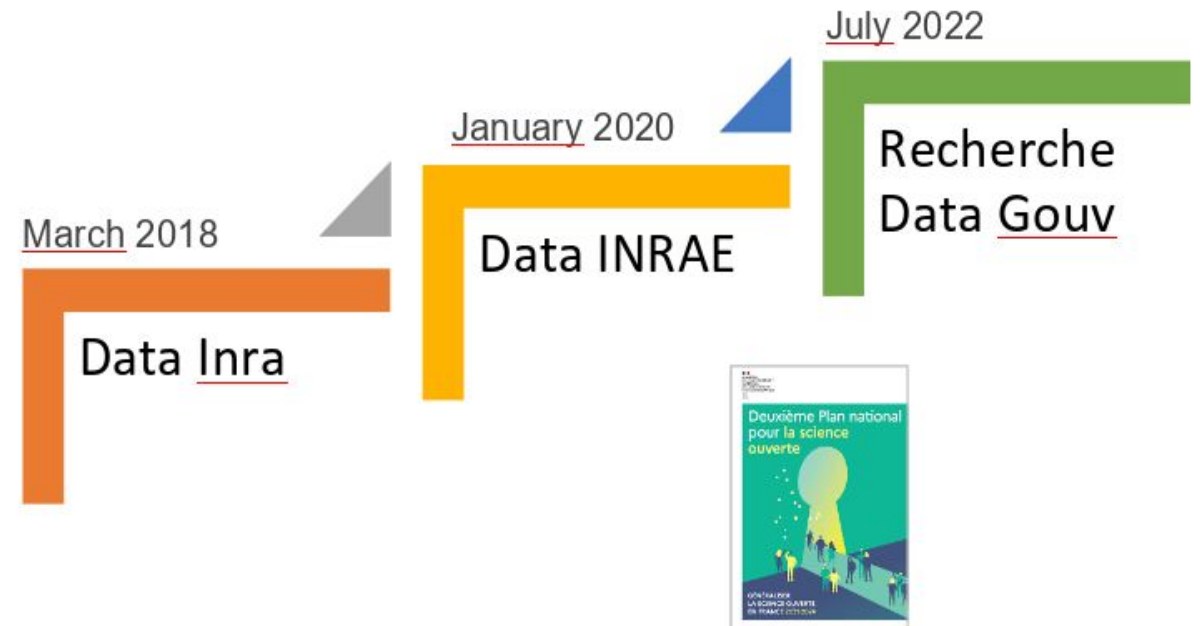
### 1 DATA REGISTRY

*These find and harvest data from external repositories*



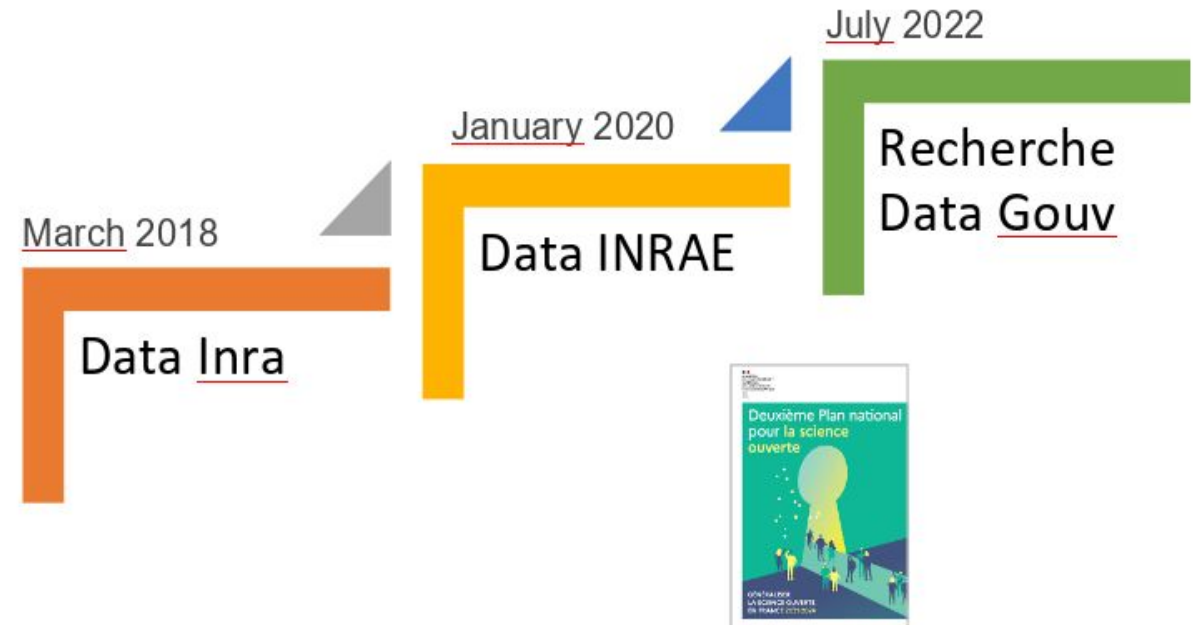
# The former Data Repository : Data INRAE

- Opened in 2018
- Following a 2 years project on Inra's data
- Institutional repository (with partners)
- Technical basis for Recherche Data Gouv



# From Data INRAE to Recherche Data Gouv

- Upscale of ressources
- INRAE's data migration
- Metadata, vocabulary and config. changes
- New environment with new technologies
- (planned) Migration to new infrastructure





# Lessons learned from the transition

- Most « INRAE » of the Work Packages
- Transparency is difficult
- Confusion between previous repository and current collection
- More budget but same team for more Use Cases and expectations



## Global view of the ecosystem : Resource centers

### 19+ DATA MANAGEMENT CLUSTERS

*Generalist expertise which is geographically close to the research teams it supports on all data issues*

### 4 INSTITUTIONAL REFERENCE CENTRES

*Provide specific guidance according to institutional policies and strategies*

### 1 DATA REPOSITORY

*A mutualised repository for all institutions to deposit and publish data*

### 6+ THEMATIC REFERENCE CENTRES

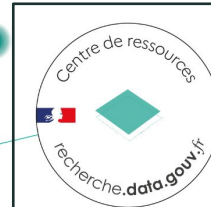
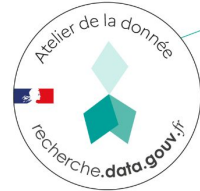
*Expertise in a given scientific field*

### 4 RESOURCE CENTRES

*These support the data management clusters and capitalise on their practices*

### 1 DATA REGISTRY

*These find and harvest data from from external repositories*



# Data Support Services

- Data repository support team
- Online & on-site trainings
- And much more (DMP, Publications, Semantics, ...)
- Still existing despite « double involvement »



## Lessons learned from the transition

- Training the trainers... takes time
- Keep new/other parallel services in mind



## Global view of the ecosystem : Research support modules

### 19+ DATA MANAGEMENT CLUSTERS

*Generalist expertise which is geographically close to the research teams it supports on all data issues*

### 4 INSTITUTIONAL REFERENCE CENTRES

*Provide specific guidance according to institutional policies and strategies*

### 1 DATA REPOSITORY

*A mutualised repository for all institutions to deposit and publish data*

### 6+ THEMATIC REFERENCE CENTRES

*Expertise in a given scientific field*

### 4 RESOURCE CENTRES

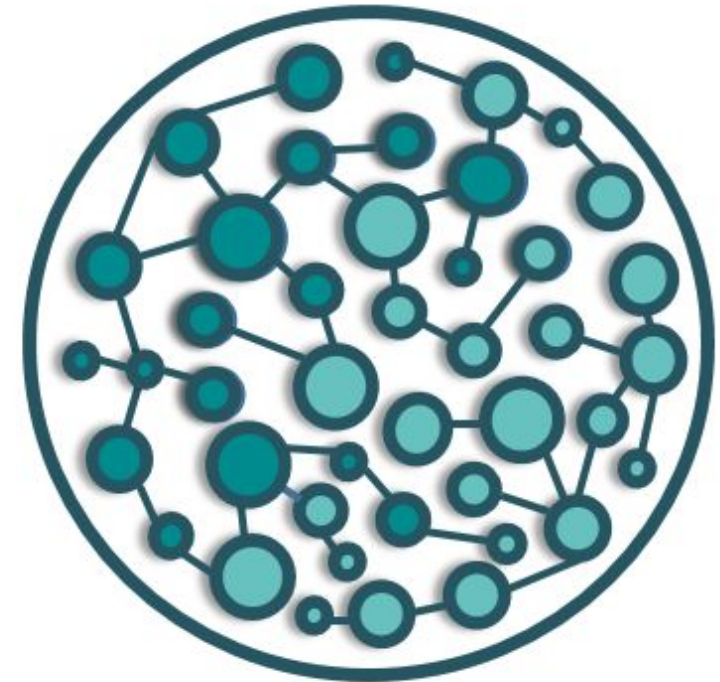
*These support the data management clusters and capitalise on their practices*

### 1 DATA REGISTRY

*These find and harvest data from from external repositories*

# The Data Ambassadors (Référénts Données) Network

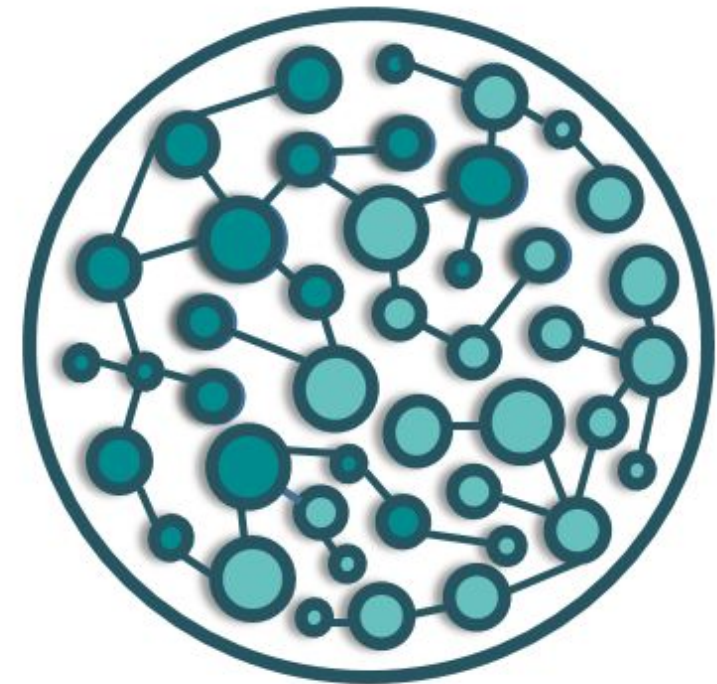
- Institutional network with a national scale
- In the laboratories
- High coverage : almost all research unit, 260+ members
- Assisting researchers in their data management
- Now an Institutional Reference Center





# Lessons learned from the transition

- Interconnection between networks happening
- Difficulties of integrating existing and different networks
- Will to expand the network of networks (RDA)



# Success factors for Recherche Data Gouv as a whole

# Integrate Support Structures

- Holistic approach
- Adapt to specificities
- Promote a culture rather than just a tool
- From network to communities



## Paradigm paradox ?

- Official support is key for endorsement and funding
  - Part of the Open Science Plan
  - Direct ministry support
  - Research Infrastructure inscription
- Integrating User needs and existing services is necessary
  - Previous survey
  - Building upon existing expertise and repository
  - Leveraging the Thematic Reference Centers

**Top Down**



**Bottom Up**

# What now ?

# From project to services

- Dedicated organisation set up
- Harmonise and mutualise between ex. Working Packages
- Research Infrastructure application
- Core Trust Seal certification



# Technical evolutions

- Data Registry : parallel to the other developments and formalisation
- Migration out of INRAE's infrastructure
- Archiving : parallel to INRAE's archiving policy
- Share the expertise and maintain with evolving technologies

# Follow the news

<https://recherche.data.gouv.fr/en>

 @RechercheDataGv

 <https://www.linkedin.com/company/recherche-data-gouv/>

# Thank you!

eosc | FAIR-IMPACT



@fairimpact\_eu /company/fair-impact-eu-project



Funded by  
the European Union



@RechercheDataGv