French Bioinformatic Institute : Challenges and opportunities in Open Science

EOSC France meeting - june 2023

Jean-François Dufayard - cirad - UMR AGAP - BIOS departement
Co-head of open science, and mutualization with infrastructures actions in IFB

Let's talk about…

I. What is IFB ?

II. IFB as a thematic reference center for Recherche Data Gouv

III. Example of resources

I. What is IFB ?

French Bioinformatics Institute

- National Infrastructure in Life Sciences and Health
  - 2013-2024: PIA1 RENABI-IFB (22.8M€)
  - 2021-2029: PIA3 MUDIS4LS (16.5M €)

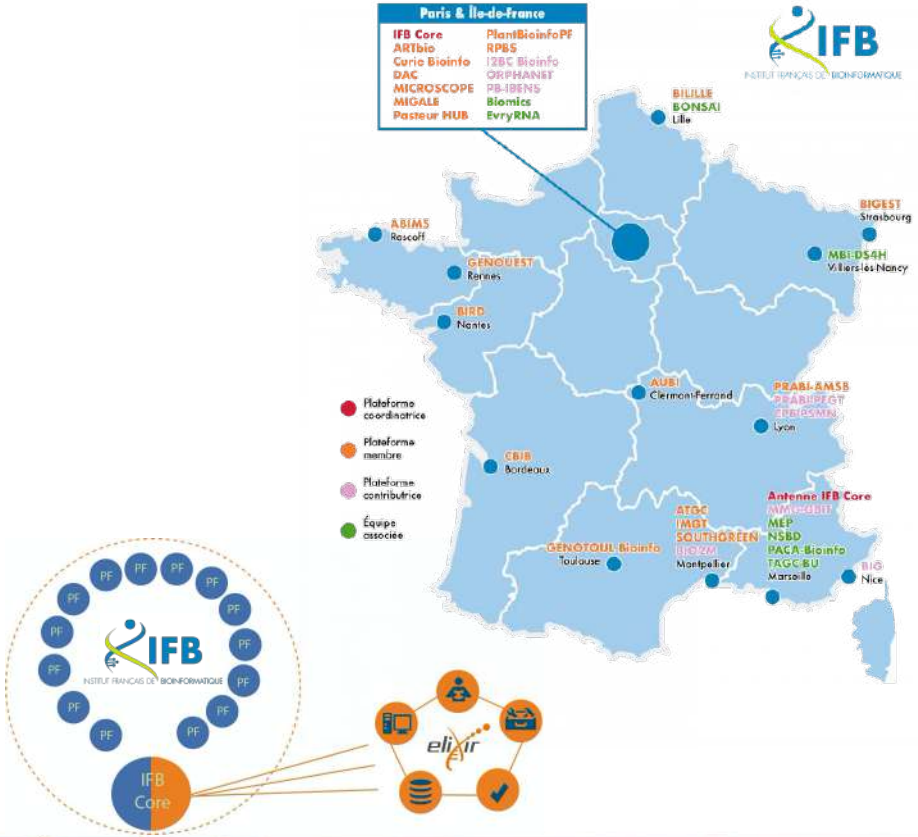
- Missions for life science communities
  - Support to research teams
  - Service deployment
  - Training management
  - Innovative developments

https://www.france-bioinformatique.fr

IFB : actors and leadership

- Actors
  - 29 platforms and 7 associated teams
  - IFB-Core (UAR 3601) : support to research unit
    - *Coordination*
    - *Mutualisation*
    - *Methodological development*
    - *User support*
    - *Training*
    - *French national node in ELIXIR*
- Leading institutes

  CNRS, Inserm, INRAE, CEA, Universités, etc.



IFB services related to open science

Open Science and Interoperability

- Data management
- Standards, ontologies…
- DMPs
- Interaction with data producing structures

User support

- Unique entry point, the “guichet”
- Project support, from the redaction to the end
- **RDMkit:** https://rdmkit.elixir-europe.org

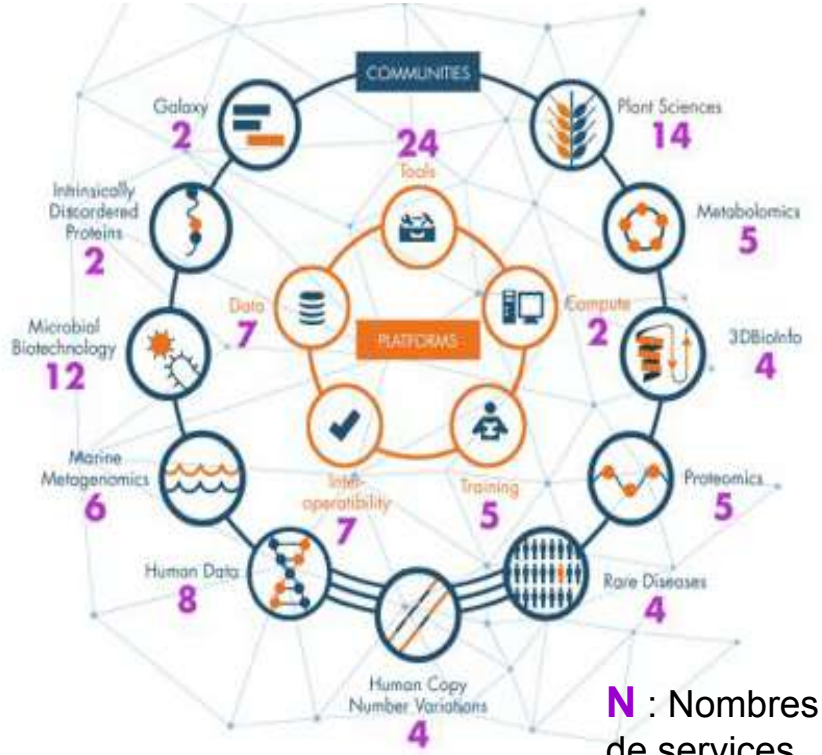
Digital infrastructure (11Po, 20 000 cores)

- National Network of Computing Resources (**NNCR**)
- 2 solutions: **cluster** + **cloud**
- **Project spaces for communities**

Training catalog

- 136 trainings (>2 500 trainies)
- **training FAIR-data and FAIR-bioinfo**

- ELIXIR:
  - International organisation
  - Coordinates resources in 22 countries
  - https://elixir-europe.org/about-us
- 5 service platforms (working groups)
  - TOOLS, COMPUTE, DATA, INTEROPERABILITY, TRAINING
- 11 scientific communities
  - 3D-BioInfo, Intrinsically Disordered Proteins, Galaxy, Metabolomics, Human Data, etc.
  - New community: research data managers



N : Nombres de services ELIXIR-FR

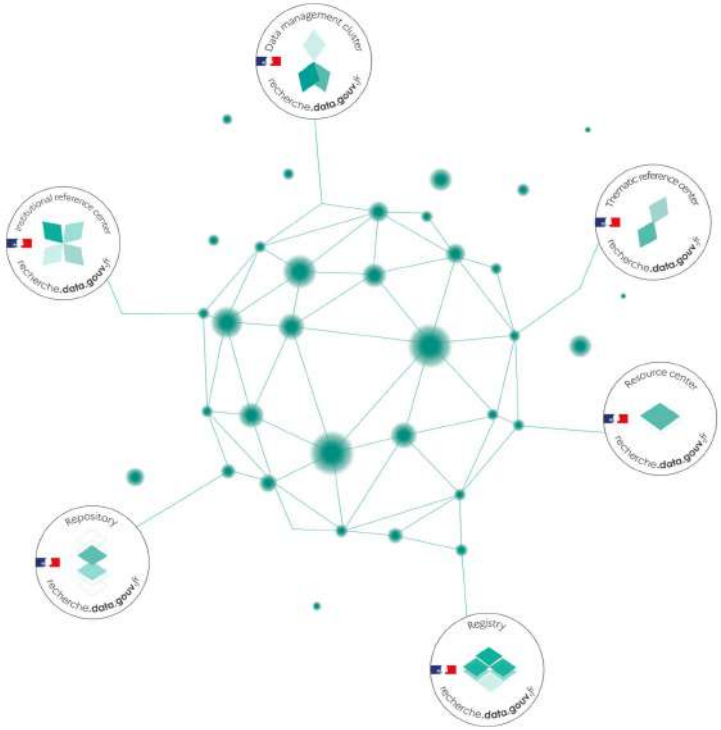
MINISTÈRE
DE L'ENSEIGNEMENT
SUPÉRIEUR
ET DE LA RECHERCHE

Liberté
Égalité
Fraternité

II. IFB as a Thematic Reference Center

Role in Recherche Data Gouv

RÉPUBLIQUE FRANÇAISE recherche.data.gouv.fr



Part of an ecosystem

Thematic reference center:

➔ Define and disseminate:

- Description standards
- Sharing standards
- Good practices
- Recommended repositories

➔ Propose open science specific tools and services to support communities

Context - which communities, which data ?

Much more than the “historical” bioinformatic domain:

- Agronomy
- Environnement
- Microbiology
- Biomedical
- …

Large panorama of heterogeneous data:

- ➔ Omics (meta, geno, transcrito, metabo, proteo…
- ➔ Many types of imagery
- ➔ Many types of tabulated data

Interoperability with other domains (ecology, human sciences…)

Communities are legion, very variable in term of maturity, and are hard to unit in common groups
Very strong role of animation for the IFB

Animation of communities

National level : national infrastructures in life sciences and health (INBS), structured around the GIS IBISA : united in a “Club des INBS”

Creation of a “research data” working group :

- ➔ Inventory of data types and repositories
- ➔ Inventory and state of the art of existing international standards
- ➔ Preparing the coconstruction of FAIRisation good practices

European level: involvement in the ELIXIR Research Data Manager Community -> aligning french and european activities and initiatives

II. Examples of resources for Open Science in IFB

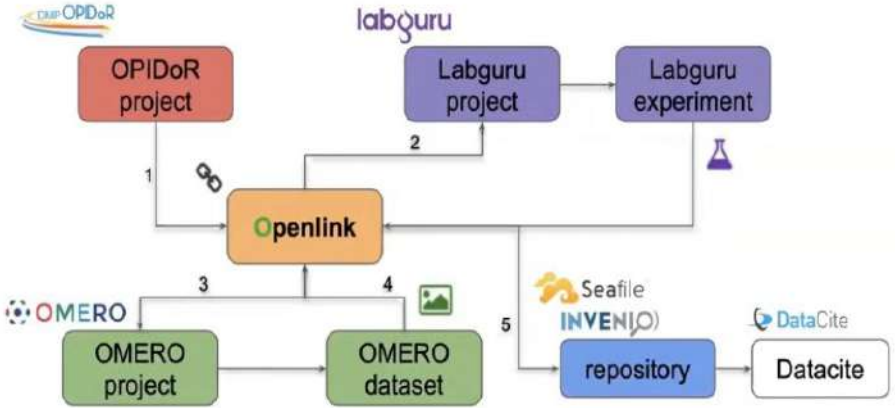
Thematic modular DMP templates

Structured and modular "entity/structure"DMP templates, domain specific

- ➔ Imagery
- ➔ Proteomics
- ➔ Genomics
- ➔ Metabolomics…

- **Modularity to compose customized DMPs**
- **Ready for "machine actionable", information flow between entity/structure DMP and project DMP**
- **Interoperability between Data Stewardship Wizard (easy template edition) and DMP-OPIDoR**

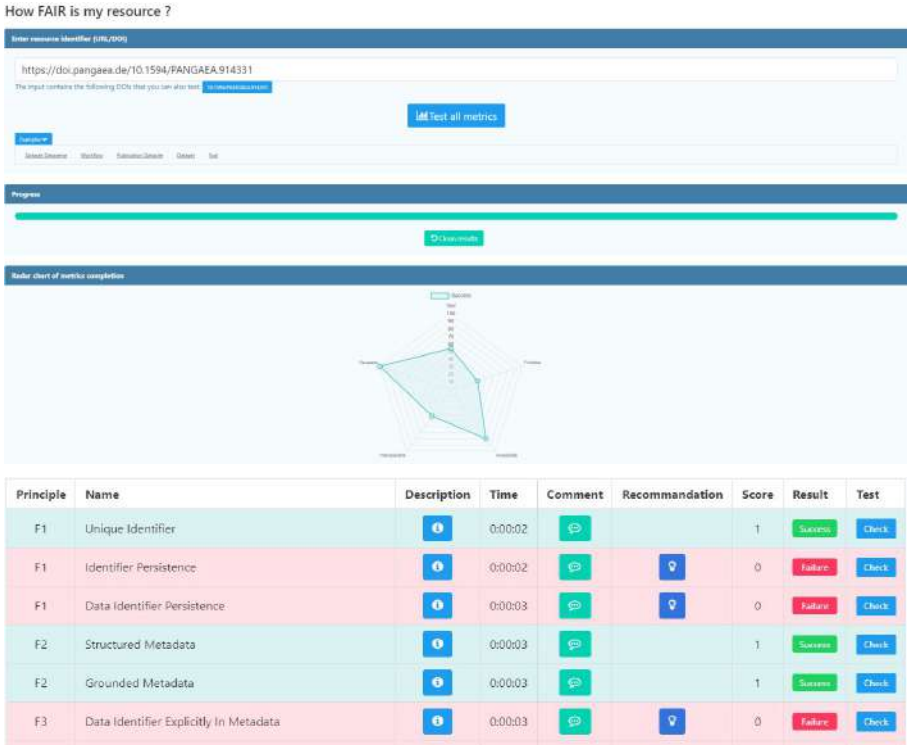
- But :
  - Ease the FAIRification process along the whole data lifecycle
- Machine actionable Data Management Plans (IFB-Core)
  - Structured DMP in OpiDOR
  - Interfaced with service platforms (automatic account and storage space creations)
  - Proof of concept
- OpenLink (IGBMC)
  - Link between DMPs, digital lab books, data repositories and publications
  - Metadata conservation along the data lifecycle
  - Proof of concept in imagery



FAIR-Checker : automatic tool for FAIR assessment

- But
  - Checking the FAIRness of a webpage
  - Automatic test, from an address or a DOI
- Solution : 2 main functionalities
  - CHECK : testing the main metrics
    - *With recommendation if failed*
  - INSPECT : advanced inspection of the webpage metadata
    - *Web semantic technologies (RDF, SPARQL, SHACL)*
    - *Quality control of metadata, if they are or not compatible with standards*

Complementary to other tools like FAIR-Evaluator, FAIR Shake, F-UJI (see FAIRassist.org)



https://fair-checker.france-bioinformatique.fr

Training catalog

CONCLUSION

- IFB : a support unit in life science and health
  - Coordination of bioinformatic platforms
  - Many open science activities
    - *Tools and services*
    - *Training : FAIR, DMPs…*
  - Mutualisation of human and digital resources
- IFB : a thematic reference center for *Recherche Data Gouv*
  - Animation of communities, for the COCONSTRUCTION of FAIRisation standards
    - *Description standards*
    - *Diffusion standards*
    - *Recommended repositories*
  - Interaction with national and european actors

Thanks for your patience, and attention !

THANKS TO

Open Science IFB group

- Marie-Dominique Devignes
- Frédéric de Lamotte
- Alban Gaignard
- Jacques van Helden
- Hélène Chiapello
- Sarah Cohen-Boulakia (-> 2019)
- GT SO&I

IFB dev teal

- Thomas Rosnet
- Paulette Lieby
- Julien Seiler
- Hervé Ménager
- Thomas Denecker
- Vincent Lefort
- Sylvain Milanesi
- Laurent Bourhi
- …

Partners

- INIST (OPiDOR)
- DSW (ELIXIR-CZ)
- EDAM (ELIXIR-NW)
- …

IFB

INSTITUT FRANÇAIS DE BIOINFORMATIQUE

elixir FRANCE

france-bioinformatique.fr

elixir-europe.org

@IFB_Bioinfo

@InstitutFrançaisDeBioinformatiqu

contact@france-bioinformatique.fr

cnrs INRAE Inserm cea