


French Bioinformatic Institute : Challenges and opportunities in Open Science

EOSC France meeting - june 2023

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



INSTITUT FRANÇAIS DE BIOINFORMATIQUE



INRAE



Inserm



- 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



Bienvenue à l'Institut Français de Bioinformatique

L'IFB est l'infrastructure nationale de bioinformatique qui assure un support, déploie des services, organise des formations et réalise des développements innovants pour les communautés des sciences du vivant.

L'Institut Français de Bioinformatique met en place une collecte des besoins en matière d'informatique et d'analyse de données pour affronter la crise sanitaire COVID-19 et offre l'appui de ses ressources humaines et informatiques pour épauler les équipes directement impliquées.

COVID-19

EN SAVOIR +

Calcul & Stockage
Clusters • Cloud
National Network of Computing Resources (NNCR)
EN SAVOIR +

Accompagnement
Projets de recherche • Communautés • Guichet d'orientation
EN SAVOIR +

Formation
Formation professionnelle & initiale
EN SAVOIR +

Innovation
Bioinformatique intégrative
EN SAVOIR +

ELIXIR-FR
Nœud français du réseau européen d'infrastructures bioinformatiques
EN SAVOIR +

Catalogue de ressources
Outils, données, expertises, services en bioinformatique
EN SAVOIR +

Actualités
Newsletter • Réseaux sociaux
EN SAVOIR +

Roadmap
Feuille de route • KPI
EN SAVOIR +

CNRS | Inserm | INRAE | CEA | Investissements d'Avenir

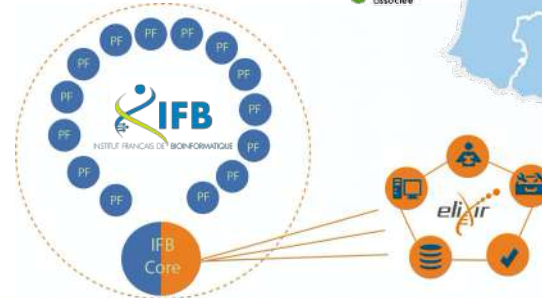
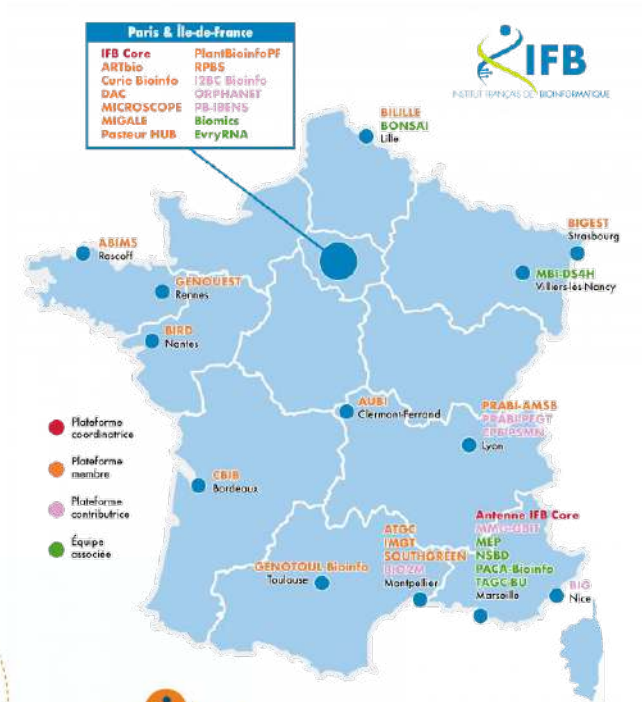
<https://www.france-bioinformatique.fr>

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.



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)



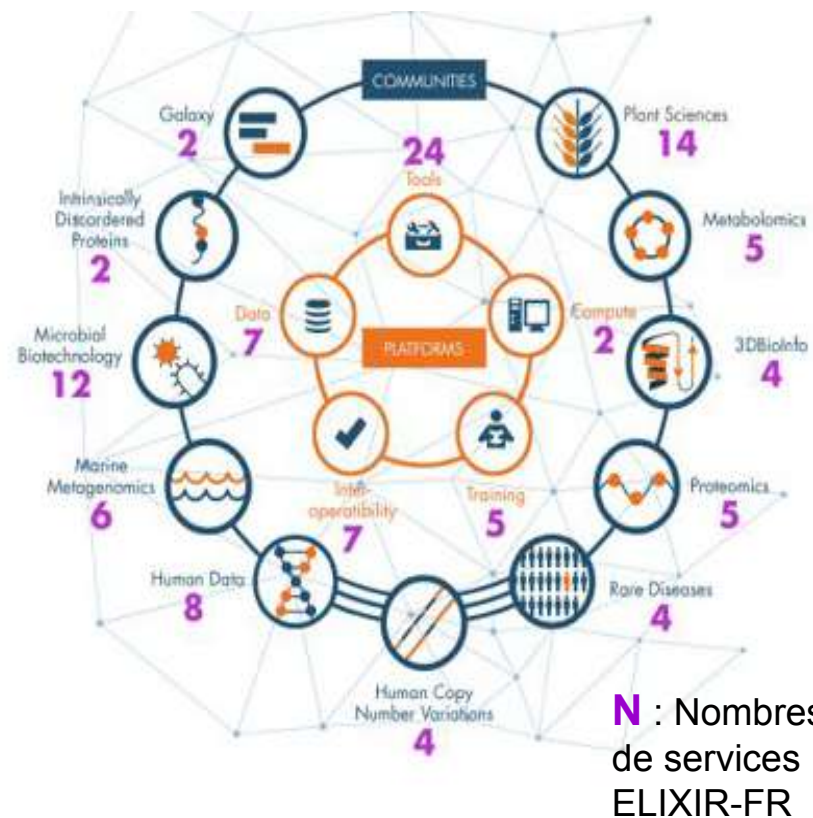
- National Network of Computing Resources (NNCR)
- 2 solutions: **cluster + cloud**
- **Project spaces for communities**

Training catalog



- 136 trainings (>2 500 trainees)
- **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





**MINISTÈRE
DE L'ENSEIGNEMENT
SUPÉRIEUR
ET DE LA RECHERCHE**

*Liberté
Égalité
Fraternité*

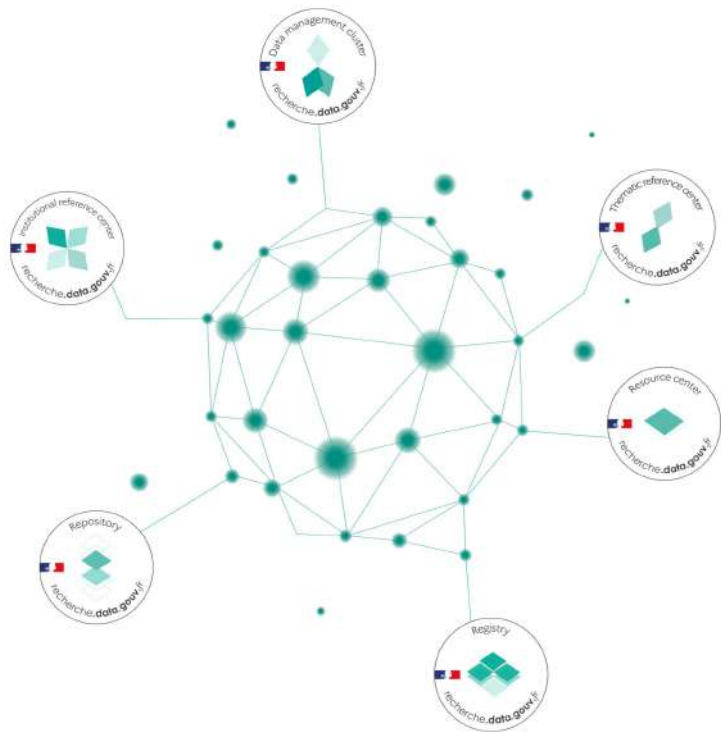
II. IFB as a Thematic Reference Center

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

Much more than the “historical” bioinformatic domain:

- Agronomy
- Environnement
- Microbiology
- Biomedical
- ...



Large panorama of heterogeneous data:

- Omics (meta, geno, transcrit, 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***

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



...
20+

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

Structured and modular “entity/structure”DMP templates, domain specific

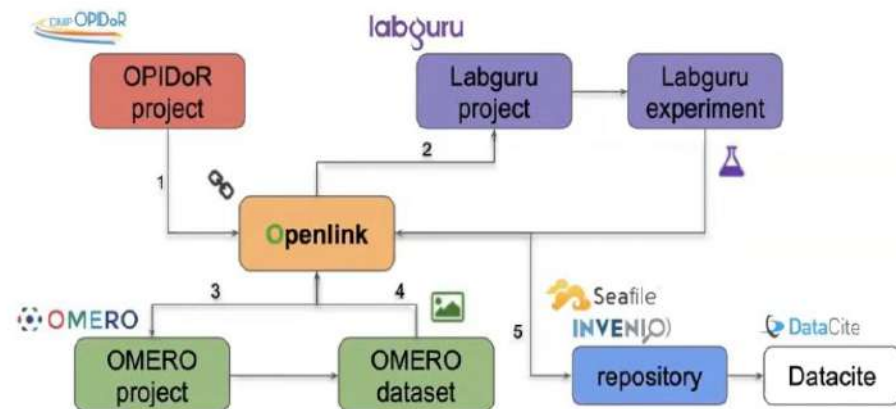
- Imagery
- Proteomics
- Genomics
- Metabolomics...



- 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



- 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

How FAIR is my resource ?

Enter resource identifier (URL/DOI)

<https://doi.pangaea.de/10.1594/PANGAEA.914331>

The input contains the following COCs that you can also see: [View metadata](#)

[Test all metrics](#)

Progress

100% Done

Radar chart of metrics completion

Principle	Name	Description	Time	Comment	Recommendation	Score	Result	Test
F1	Unique Identifier		0:00:02			1	Success	Check
F1	Identifier Persistence		0:00:02			0	Failure	Check
F1	Data Identifier Persistence		0:00:03			0	Failure	Check
F2	Structured Metadata		0:00:03			1	Success	Check
F2	Grounded Metadata		0:00:03			1	Success	Check
F3	Data Identifier Explicitly In Metadata		0:00:03			0	Failure	Check

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

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





Catalogue des formations

Number of data-items: 7, show on 1 pages.

Training name

Perform a filter using the legend below

Open for registration Upcoming training Past training

-  02 Feb 2023 [Principes FAIR pour la gestion des données d'une plateforme IBISA - session 2023](#)
📅 2023-02-02 | 2023-02-03
📍 Paris France
-  13 Jun 2022 [Principes FAIR dans un projet de bioinformatique - Session 2022](#)
📅 2022-06-13 | 2022-06-15
📍 Paris France
-  22 Mar 2022 [Principes FAIR pour la gestion des données de recherche en sciences de la vie - édition Francilienne - Session 1 \(mars\)](#)
📅 2022-03-22 | 2022-03-23
📍 Paris France
-  [Formation IFB Science Ouverte & PGD - Comment gérer des jeux de données haut débit en sciences de la vie et de la santé - Session 2](#)

- **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 Gov***
 - 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 !

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)
- ...



INSTITUT FRANÇAIS DE BIOINFORMATIQUE



france-bioinformatique.fr



elixir-europe.org



[@IFB_Bioinfo](https://twitter.com/IFB_Bioinfo)



[@InstitutFrancaisDeBioinformatique](https://www.linkedin.com/company/InstitutFrancaisDeBioinformatique)



contact@france-bioinformatique.fr



INRAE



Inserm

