IFB/ELIXIR-FR, the french node of ELIXIR

Anne-Françoise Adam-Blondon, Jacques van Helden, Gildas Le Corguillé,
Christophe Blanchet, Hélène Chiapello, Christine Gaspin, David Salgado,
Julien Seiler, Morgane Thomas-Chollier, Yves Vandenbrouck, et al.

ELIXIR and the French Institute of Bioinformatics (IFB)

Two collaborative networks of people building resources and services for research in Life Sciences

ELIXIR: a European distributed infrastructure for Life Science information

https://www.elixir-europe.org/
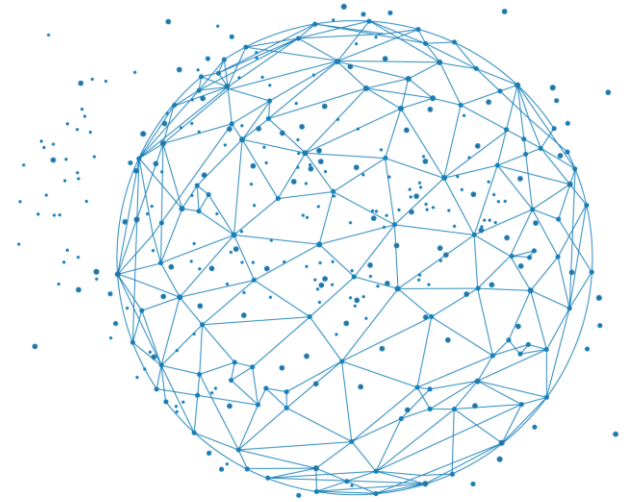
26 nodes: 25 national nodes + EMBL-EBI collaborate in order to:

Enable scientists to access and analyse life science data

Develop people and capacity to benefit science and society

Deliver services to support distributed data management and analytics

Science

People

Technology

Nodes

Equip national Nodes for successful long-term operations

The french Institute of Bioinformatics (IFB) is ELIXIR-FR, the French node of ELIXIR

IFB is a distributed French National Research Infrastructure

- 36 platforms and research teams spread all over France

Coordination: IFB-core

- Multi-research organizations : CNRS, Inserm, INRAE, CEA

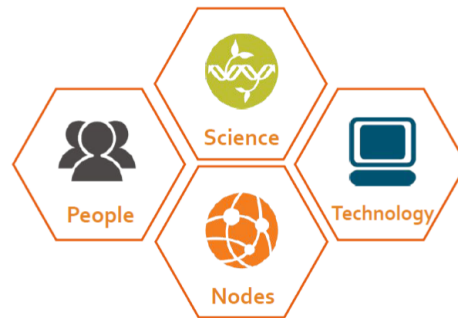**Common resources**

- Supported by the French Ministry of Higher Education and Research since 2012

A large set of expertise in bioinformatics

- >400 experts (~200 FTE)

Paris & Île-de-France

IFB Core	PlantBioinfoPF
ARTbio	RPBS
Curie Bioinfo	I2BC Bioinfo
DAC	ORPHANET
MICROSCOPE	PB-IBENS
MIGALE	Biomics
Pasteur HUB	EvryRNA

- Plateforme coordinatrice
- Plateforme membre
- Plateforme contributrice
- Équipe associée

Map labels: BILILLE BONSAI Lille, BIGEST Strasbourg, ABIMS Roscoff, MBI-DS4H Villiers-lès-Nancy, GENOUEST Rennes, BIRD Nantes, AUBI Clermont-Ferrand, PRABI-AMSB PRABI-PFGT CPB-PSMN Lyon, CBIB Bordeaux, Antenne IFB Core MMG-GBIT MEP NSBD PACA-Bioinfo TAGC-BU Marseille, ATGC IMGT SOUTHGREEN BIO2M Montpellier, GENOTOUL Bioinfo Toulouse, BIG Nice

IFB/ELIXIR-FR has a large set of domain of applications

IFB's domains of application encompass the three scientific priorities of the 2024-2028 programme of ELIXIR

Enabling and responding to advances in cellular and molecular research will drive life science

Connect the latest developments and established data resources to realise the potential of cellular and molecular biology

Biodiversity, food security and pathogens are interrelated societal challenges

Mobilise and integrate molecular data to support transnational research programmes in biodiversity, food security and pathogens

The promise of genomics research to improve health and disease outcomes

Provide the infrastructure to support the discovery, access, sharing and analysis of human genomics data and linked data on a massive scale

Better deliver the data, services and expertise necessary to address important societal challenges

- **Contribution to transversal activities addressing common challenges**

  - Multi-research organizations governance

  - FAIR data management and data interoperability

- **Partnership in large national projects**

- **Enhance partnerships in ELIXIR activities and in European projects**

ELIXIR-FR is actively contributing to ELIXIR's activities
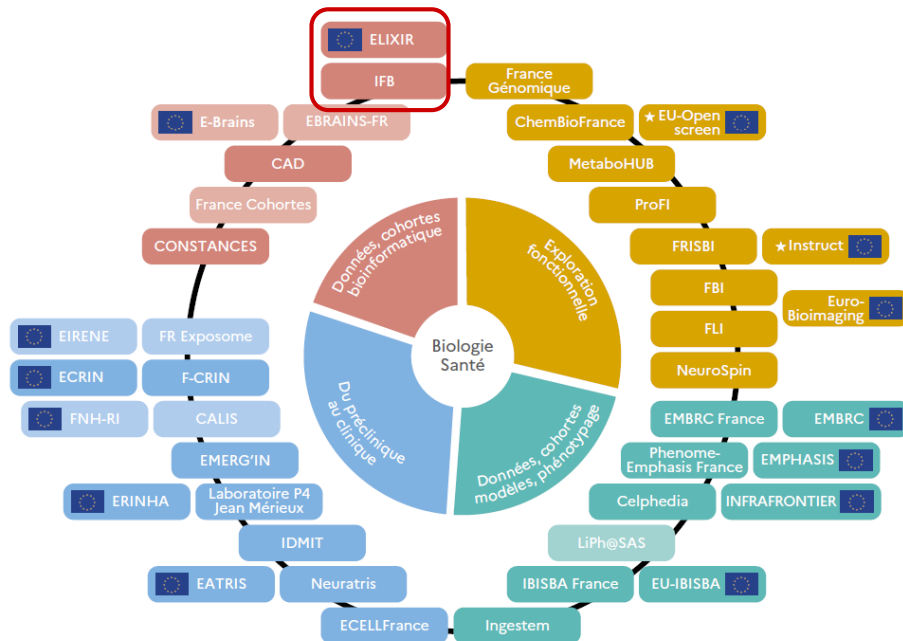
- **Liaising and networking** with all platforms and most of the communities and contributing to their activities
- **Co-coordinating and contributing to the strategic planning of some of them** FR
- Contributing to European projects**: 9 since 2020

Platforms

Technology

ELIXIR Platforms bring together experts from Nodes to develop ELIXIR's technical vision and coordinate activities in defined technical areas. There are five Platforms: Data, Tools, Interoperability, Compute and Training.

Data Platform
Aims to identify key data resources across Europe and support the linkages between data and literature.

Tools Platform FR
Helps researchers find the best software tools to analyse their data.

Compute Platform
Develops services to make it easier to store, share and analyse large datasets.

Interoperability Platform
Develops and encourages the adoption of standards to describe life science data.

Training Platform
Helps scientists and developers find the training they need, and also provides that training.

Communities

Science

3D-Bioinfo	Biodiversity
Food and nutrition	Galaxy
Intrinsically Disordered Proteins	Marine Metagenomics FR
Metabolomics	Microbial Biotechnology
Plant Sciences FR	Proteomics
Single-Cell Omics	Systems Biology
Toxicology FR	

Human Data Communities

Federated Human Data	Human Copy Number Variation FR
Rare diseases	Cancer Data (focus group) FR

Guy Perrière
ELIXIR-FR Head of Node

Anne-Françoise Adam-Blondon
ELIXIR-FR Head of Node

Angela Sáenz
ELIXIR-FR Node coordinator

Development of the 2022-2025 roadmap

…and brainstorming on the next roadmap

Claudine Médigue
Director of IFB
Head of Node ELIXIR-FR
2017-2022

Gildas Le Corguillé
Interim director of IFB
2022-2023

Jacques van Helden
IFB's scientific coordinator
Head of Node ELIXIR-FR

Anne-Françoise Adam-Blondon
Deputy Head of Node ELIXIR-FR
2018-2023

New road map and new organisation of IFB activities

Infrastructure
- National Network of Computing Resources (**NNCR**)
- Two modalities: **cluster** + **cloud**
- Compliance to law (**GDPR**)

Tools
- **Development** of dedicated tools
- **Packaging** (conda)
- **Virtualisation** (appliances, containers)
- **Software development practices**
- **Collaborative development, open code**

Support to users
- One stop portal and contact address
- Co-development of the bioinformatic/statistical part of the project
- **DMP**
- Development of **workflows**
- **Data Science**

Knowledge bases
- Development of data and knowledge bases
- Quality and data standards
- Technical support to developers
- Deployment of databases developed up in France

Task Forces

Training
- **Thematic schools** (NGS, multi-omics, phylogeny, biostat, programming, workflows…)
- **Bring Your Own Data** (BYOD) trainings
- Webinars, MOOC
- **FAIR-data et FAIR-bioinfo training and capacity building**

Prospective and innovation
- Identification of emerging needs in bioinformatics
- Funding of Pilot-projects

Open Science & Interoperability
- Data Management
- Data Standards, Ontologies, semantics
- Data brockering
- Machine actionable DMP
- Contribution to the network of RIs

11

A flavour of IFB/ELIXIR-FR in operations in 2023 - Synergies with ELIXIR

Distributed infrastructure leveraging IFB's platforms forces
- www.france-bioinformatique.fr/en/compute-and-storage
- 9 platforms involved
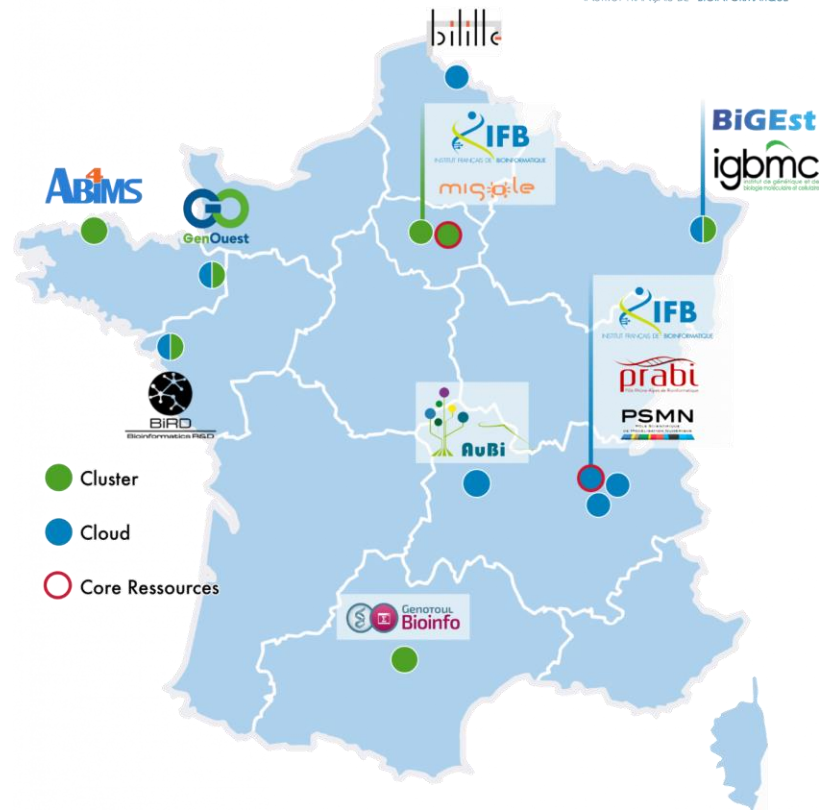- 27,000+ HT-cores, 135 To RAM, 14 Po storage

Operation using **complementary technologies**
- 7 **clusters** (SLURM)
- 8 **clouds** (Openstack)

Managed by a **distributed team of experts from IFB's platforms (+ 17 persons)**
- Sharing of experience, best practices, development

Collaborations with other **regional and national computing centers**

- ● Cluster
- ● Cloud
- ○ Core Ressources

Julien Seiler

Gildas Le Corguillé

Christophe Blanchet

Olivier Collin
ELIXIR Tech Coordinator

A few indicators	2021	2022
Total number of active accounts	10 900	10 329
New accounts	2 884	4 272

Virtual and reproducible environments of analysis

National activities

- **Software development**
    - Sharing good practices for development
    - Sharing recipes for deployment

- **Deployment of environments for analysis**
    - E.g. Usegalaxy.fr (metabolomics, proteomics, covid19, …)

Tools development and deployment	2021	2022
New tools	19	16
Number of unique users of the platforms web tools	573 244	352 644
Number of uploads of the tools developed by the platforms	16 912	24 307

Vincent Lefort

Hervé Ménager
ELIXIR Tools PF coord
EDAM steering committee

Anthony Bretaudeau
Elixir Galaxy community

Olivier Sallou
Elixir Tech Coord

International/ELIXIR activities

- **Service Delivery Plan (SDP) d'ELIXIR:** labelling of French **services**

- **Involvement in the coordination of the ELIXIR Tools Platform** (Hervé Ménager): contribution to the strategic planning, coordination of activities, outreach, good practices

- **Involvement in strategic projects for the ELIXIR Tools platform: BioContainers** (Olivier Sallou) Dynamic deployment of virtual environments

- Involvement in the **ELIXIR Galaxy community** (Anthony Bretaudeau): contribution to usegalaxy.eu

14

National activities

- **IFB web portal:** www.france-bioinformatique.fr

- **Catalogue of the French resources** (in progress)
    - Computing and storage resources, tools, knowledge bases
    - Training material and events
    - Expertises
    - IFB platforms

- **Support of user's projects** by IFB's platforms: data analysis and data valorisation

Sylvain Milanesi   **Valérie Cognat**   **Christine Gaspin**

Project Support	2021	2022
Total number	255	309
International	19	37
Private companies	13	13

International activities (carried out by the Tools + the training Task Forces)

- **Automatic synchronisation** IFB's catalogue with ELIXIR's registries:
    - **Bio.tools**: softwares
    - **TeSS**: training material and events
- **Improvement of searches by users, thematic collections of resources, cross referencing: EDAM**, bioinformatic ontology co-supported by ELIXIR-FR and ELIXIR-Norway
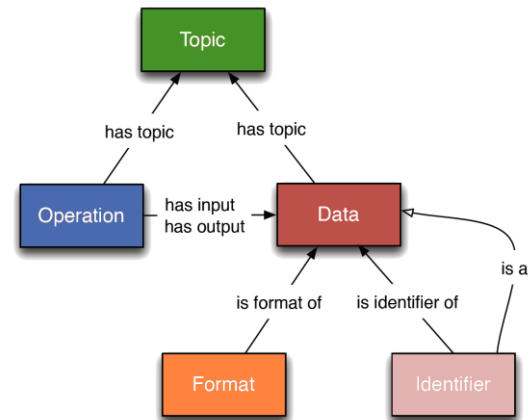
- **Development of services supporting FAIR data management and reproducible data analysis**
 - Transversal to all research communities : MUDIS4LS (FR), EOSC-Life (EU), ELIXIR-CONVERGE (EU), ELIXIR-STEERS,
  - Dedicated to specific domains of application
    - SARS-CoV-2 surveillance: EMERGEN (FR), ELIXIR-CONVERGE (EU)
    - Antibioresistance surveillance: ABRomics (FR)
    - Agroecology: AGROSERV (EU)
    - Plant Phenotyping: PHENET (EU)
    - Food systems: SAMS – Cloud4SAMS (FR)
- **Leverage ELIXIR/ELIXIR-FR "on-shelf" services to support data management and analysis in users projects**
  - Agroecology: AgroNum-BReIF (FR)
  - Biodiversity: AGENT (EU), ATLASea-BYTE-Sea (FR)

ELIXIR collaborative one-stop portal of guidelines and resources to support FAIR data management: https://rdmkit.elixir -europe.org/
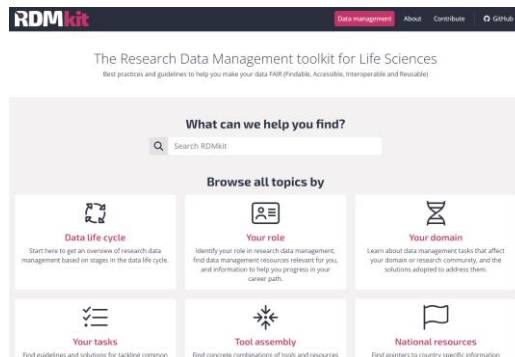
Operational interoperability and discoverability of multi-stakeholders data flows
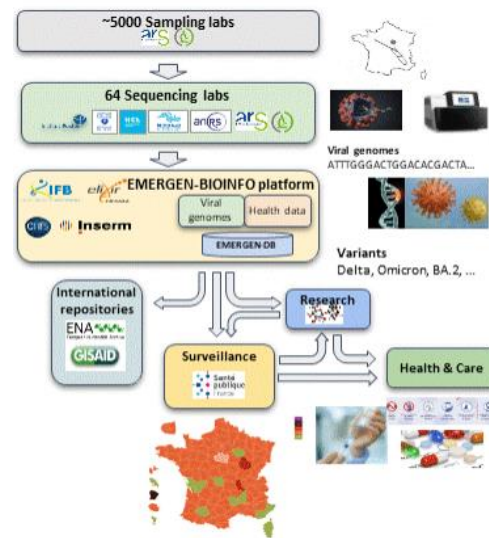
- ELIXIR coordination towards the development of a European Health Data space based on access to a federation of databases
  - **Beyond One Million Genomes (B1MG)**
  - **Genome Data Infrastructure (GDI)**
  - EOSC4Cancer

GDI main objectives :

- Support the EU 1+Million Genomes (1+MG) initiative (Digital Europe policy) ambition
- Establishing a federated, sustainable, interoperable and secure infrastructure based on open community standards.
- Each 20 1+MG signatory countries in GDI
  - Will provide a node or Data Hub
  - Each country manages their own data (e.g. regional hubs):
  - Data hubs provide cross-border data analysis

France:

- October 2022: signature by the French Ministry of Research of the 1+MG consortium agreement
- By 2027 : set-up of an operational national node giving access to genomic and clinical data
  - Leader: the French genomic medicine infrastructure (CAD)
  - IFB/ELIXIR-FR leverage its services benefiting to the data node

European Genomic Data Infrastructure

National activities

- **Trainings provided under different formats to best address the diversity of training needs**
    - **Schools** for researchers in collaboration with research institutes: EBAII, ETBii 2023
    - **University Diploma** in collaboration with Université Paris Diderot (2019-2021): **DUBii**
    - **Trainings** for capacity building on FAIR DM and Reproducible analysis: FAIR data and FAIR bioinfo
    - **E-trainings** (*in progress*)

- **Common training resources**
    - IFB's training catalog synchronized with TeSS
    - Management of the trainings and training materials: Moodle, IFB's GITlab

International activities

- Contribution to ELIXIR **Training Platform** and **FAIR Training Focus Group** transversal activities
    - Training Coordinators Group
    - Best practices, FAIRification of the training material: Manuel FAIR Training
    - Improvement of the discoverability and interoperability of training resources registered in TeSS
    - Train-the-Trainer

- "Integrative bioinformatics summer school" organized **in collaboration with ELIXIR-CH** (CAES d'Aussois, FR, sept 2023)

Hélène Chiapello
ELIXIR Training Coord.

Olivier Sand
ELIXIR Training Coord.

Trainings	2021	2022
Total number of training events	132	160
Total number of trainees	4 458	4 936
Satisfaction rate	95%	92%

18

Open Science and Interoperability

National activities: development of tool supporting FAIR data management

- **FAIRchecker: e**valuation of the FAIRness of web resources

- Planning (*in progress*)
  - Machine actionnable transfer of information between DMPs and compute infrastructures
  - Improvement of **DMP knowledge models** for projects and for the national infrastructures
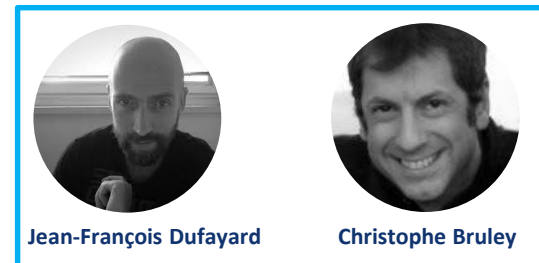 - **Interoperability between DMP softwares**

- Data brockering (*in progress*)

International activities: development of guidelines, standards and ontologies

- Participation to the emergence of the **ELIXIR Data Management community** : definition of its missions, contribution to the network of data stewards
- Contribution to some **Research Data Alliance** activities
- Standards : **Bioschemas**

Frédéric de Lamotte

Marie-Dominique Devignes

Alban Gaignard
Bioschemas steering committee

Jean-François Dufayard   **Christophe Bruley**

Liaison with the network of French Research Infrastructures for Life Sciences

- 2022 the French Ministry of Research and Higher Education launches **Recherche Data Gouv**
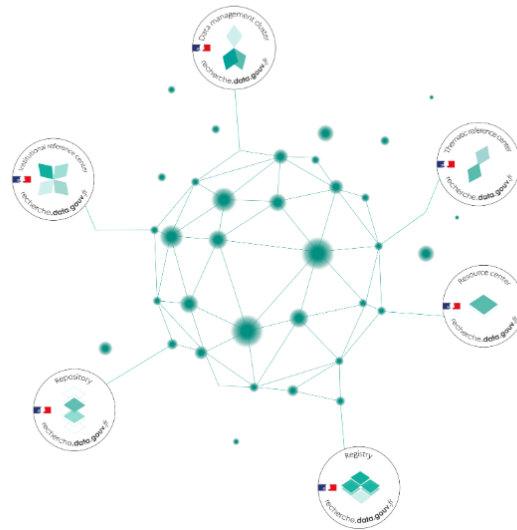
- **IFB/ELIXIR-FR is mandated to organize a « thematic reference center » for Life Science data :**
  - Development and dissemination of data standards
  - Support to data management and data curation
  - Data brokering
  - Harvesting data for the national data catalog

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

recherche.**data**.**gouv**.*fr*

Recherche Data Gouv at a glance

Modules to support research teams on all issues linked to data:
- Data management clustres
- Thematic reference centres
- Resources centres
- Institutional reference centers

Modules to deposit, publish and disseminate data:
- A repository to deposit and disseminate data
- A registry to search for data published in the repository itself or other external repositories

With a board of 3 deputy directors

Director	International coordination	Platform/service coordination	Scientific coordination

Guy Perrière
ELIXIR-FR Head of Node

Anne-Françoise Adam-Blondon
ELIXIR-FR Head of Node

Morgane Thomas-Chollier

Jacques van Helden

Thank you for your attention!

https://www.france-bioinformatique.fr/en/elixir-fr/
Contact us: angela.saenz@france-bioinformatique.fr
Follow us:
@IFB_Bioinfo   @ELIXIREurope
institut-français-de-bioinformatique/

ELIXIR director recruitment open for applications

Deadline for applications is 9:00 BST on Monday 7 August 2023

https://elixir-europe.org/news/director-recruitment-open