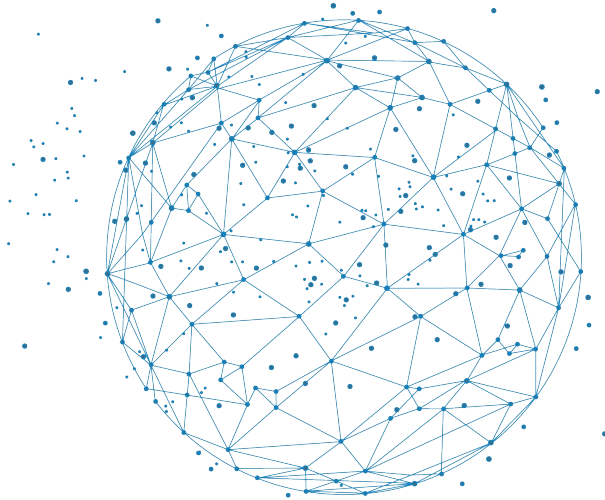


Atelier EOSC-FR, 25 janvier 2023, Strasbourg

L'Institut Français de Bioinformatique et EOSC



Christophe Blanchet (CNRS IFB-core UAR3601)
et les membres des activités mentionnées.



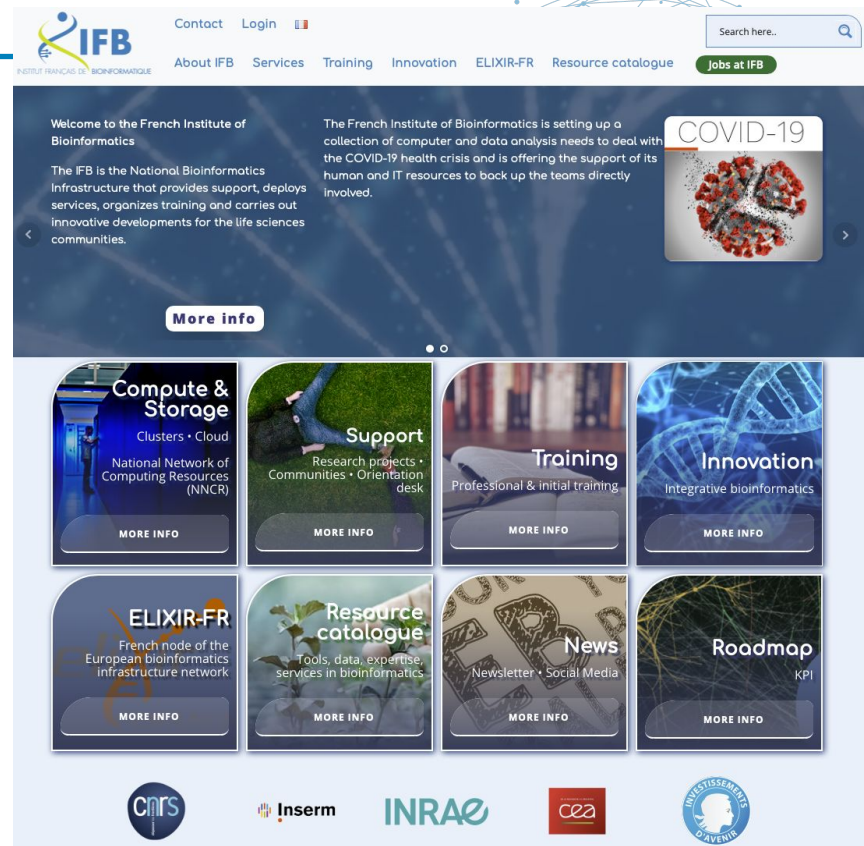
Develop a national service to provide

- core infrastructure of bioinformatic services to the French life science communities;
- training in bioinformatics for life scientists and bioinformaticians;
- support to communities in biology, health, agronomy and environment research programs;
- access to state-of-the-art technologies, high-level expertises and mutualized skills.

Develop a strategic vision to maintain French research at the highest level of expertise

- Innovation axis: integrative bioinformatics
- Be used as a lever to ambitious national and international research projects in life sciences and their applications

Ensure the international representation of the French bioinformatics community



The screenshot shows the homepage of the French Institute of Bioinformatics (IFB). At the top, there is a navigation bar with links for 'Contact', 'Login', 'About IFB', 'Services', 'Training', 'Innovation', 'ELIXIR-FR', 'Resource catalogue', and 'Jobs at IFB'. A search bar is located on the right. The main content area features a welcome message and a 'COVID-19' banner. Below this, there are eight interactive cards for 'Compute & Storage', 'Support', 'Training', 'Innovation', 'ELIXIR-FR', 'Resource catalogue', 'News', and 'Roadmap', each with a 'MORE INFO' button. At the bottom, logos for partner organizations like CNRS, Inserm, INRAE, cea, and Investissements d'Avenir are displayed.

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

Overview of IFB in 2022

Distributed infrastructure

- 36 platforms and research teams
- widely distributed
- covering all bioinformatics domains

Reach all life sciences

fundamental biology, health, agriculture, environment

Pooled resources

- Fundings:
PIA1 RENABI-IFB (22.8M€), PIA3 MUDIS4LS (16.5M€)
- Task forces organized across platforms

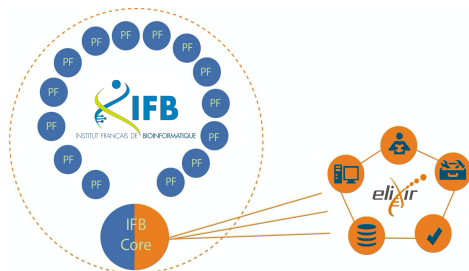
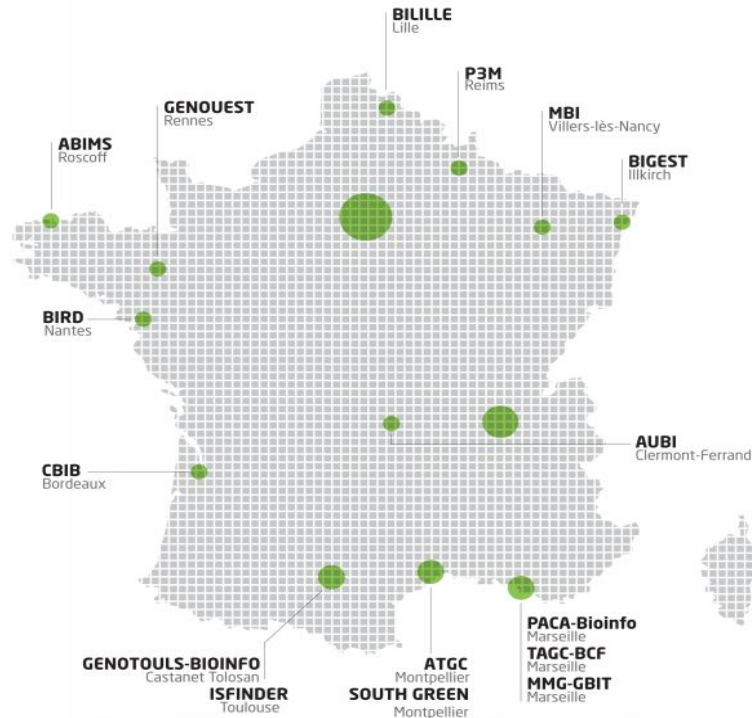
Coordination: IFB-core (UAR 3601)

- Multi-organisms (CNRS, Inserm, INRAE, CEA)
- Mission : coordination and management of shared resources
- **French node of ELIXIR (ELIXIR-FR)**

Common projects with high impact on the organization of the community

- National projects
- European projects

20 member platform
8 contributing PF
8 research teams
>400 experts (~200 FTE)



RÉGION PARISIENNE

EBIO Orsay
INSTITUT CURIE Paris
IGR Villejuif
MICROSCOPE Evry
MIGALE Jouy-en-Josas

C3BI

Paris
RPBS Paris
URGI Versailles
ORPHANET Paris
ICONICS Paris
IFB CORE Evry

RÉGION LYONNAISE

INCA-SLC Lyon
PRABI-HCL Lyon
PRABI-AMSB Villeurbanne
PRABI-Lyon-Grenoble Villeurbanne
PRABI-Lyon-Gerland Lyon
CBP-PSMN Lyon

Infrastructure

- Clusters (SLURM) and Clouds federation (Biosphere, RAINBio, Standard API)
- MUDIS4LS (PIA3), ELIXIR (ESFRI), EOSC-Life (EU H2020), EOSC Pillar (EU H2020)

Tools and Workflows

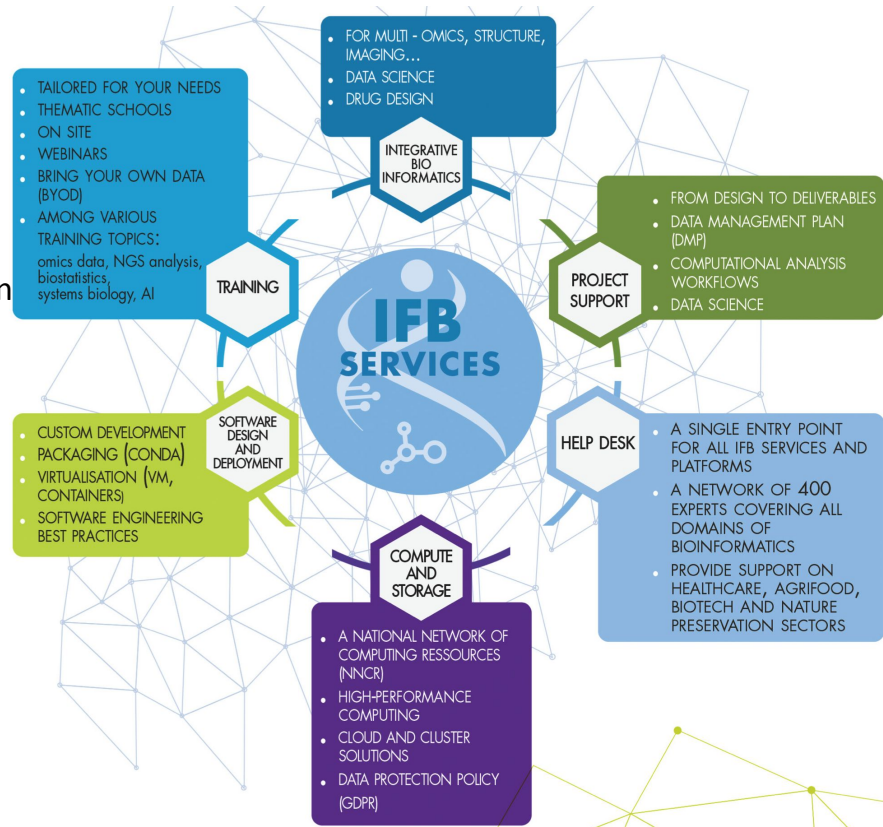
- Service catalogue and git repositories (IaC) for tools installation and pre-defined environments (clusters and clouds).
- Nextflow, snakemake, CWL, usegalaxy.fr.

Data

- DMP, FAIR-checker
- maDMP4LS (ANR), ELIXIR-CONVERGE (EU H2020)
- Centre de référence thématique Biologie-Santé (Recherche Data Gov)

Training

- Catalogue and trainings FAIR-Bioinfo, Integrative Biology...



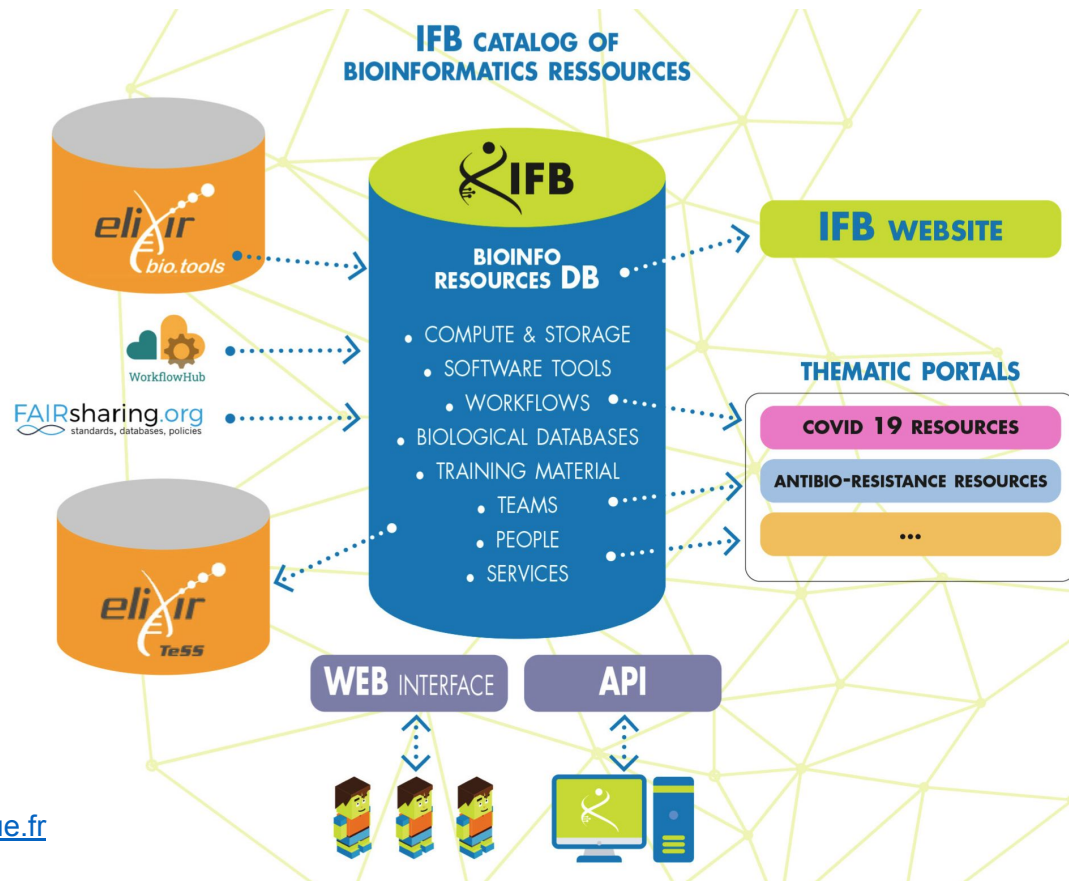
Catalogue of services for stakeholders, researchers, teachers, students, partner institutions...

Contents : data, tools, workflows, infrastructures, trainings... provided by IFB.

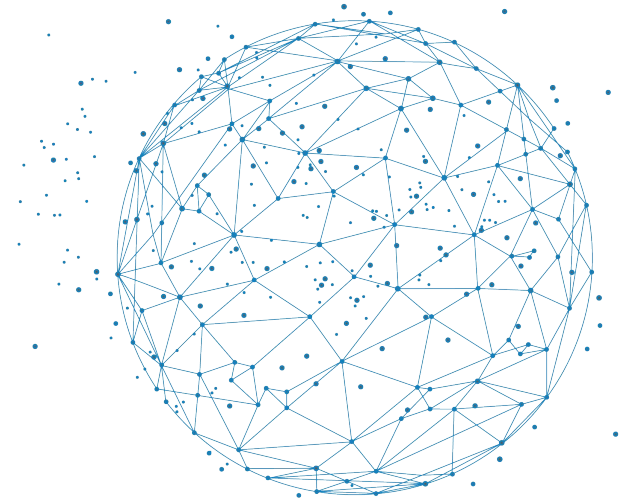
Available online through web pages and APIs (OPenAPI, swaggerAPI).

Connected to the EU ELIXIR resources: bio.tools, TeSS, EDAM, Tools ecosystem

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



Services and Projects



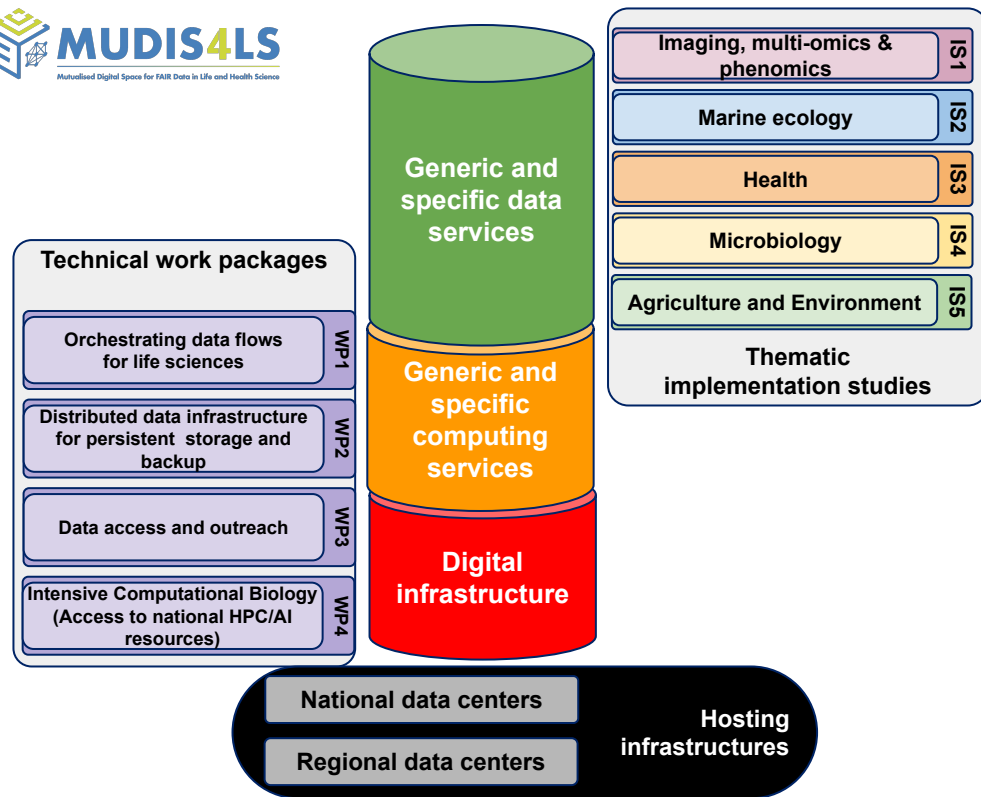


Main axes of the MUDIS4LS project

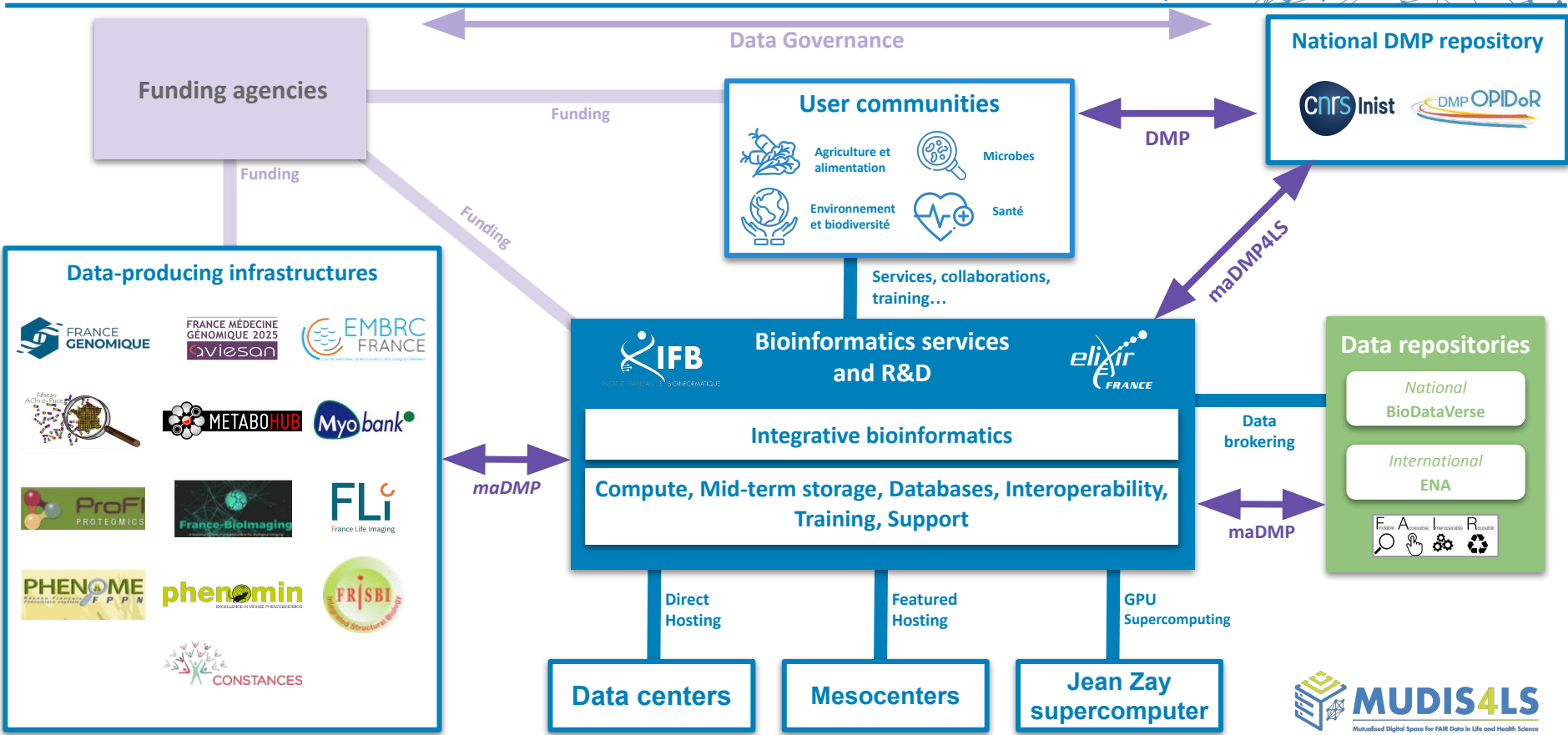
- **Orchestrating data flows** along the whole life cycle
 - maDMP-based management
 - Data brokering
 - Linking to electronic lab book
- Data **FAIR**ification
- **Enlarging the national infrastructure NNCR**
 - Anchoring in national and regional **data centers**
 - Subcontracting with some regional **mesocenters**
- **Data securing** during the projects (mid-term storage)
- **5 thematic implementation studies** ensuring relevance to address user needs

Organisation

- Partners: 39 teams and 14 organisms
- 4 national + 7 regional data centers, 4 partner mesocenters
- 170 persons, 3000 PM = 2.5 person•centuries
- Funding 2021-29: **16,5 M€**



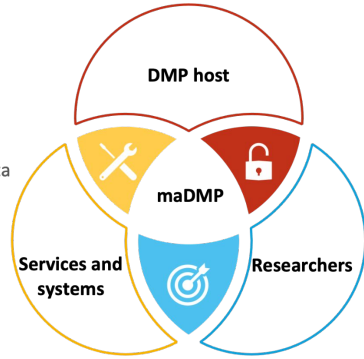
Orchestrer les flux de données tout au long de leur cycle de vie





Data Management

- Better management of resources
- DMP up to date
- Ensuring FAIRness of data



Communication

- Understanding the specific needs of the researcher
- Less work for the researcher (avoiding double capture)
- Better visibility for facilities

FAIR access

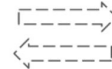
- Fostering safe FAIR principles application
- Share as much as possible
- Protect the data as much as needed

2. Get the project informations



3. Get the project going

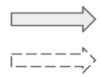
1. Fill the DMP



5. Bidirectionnal data transfer

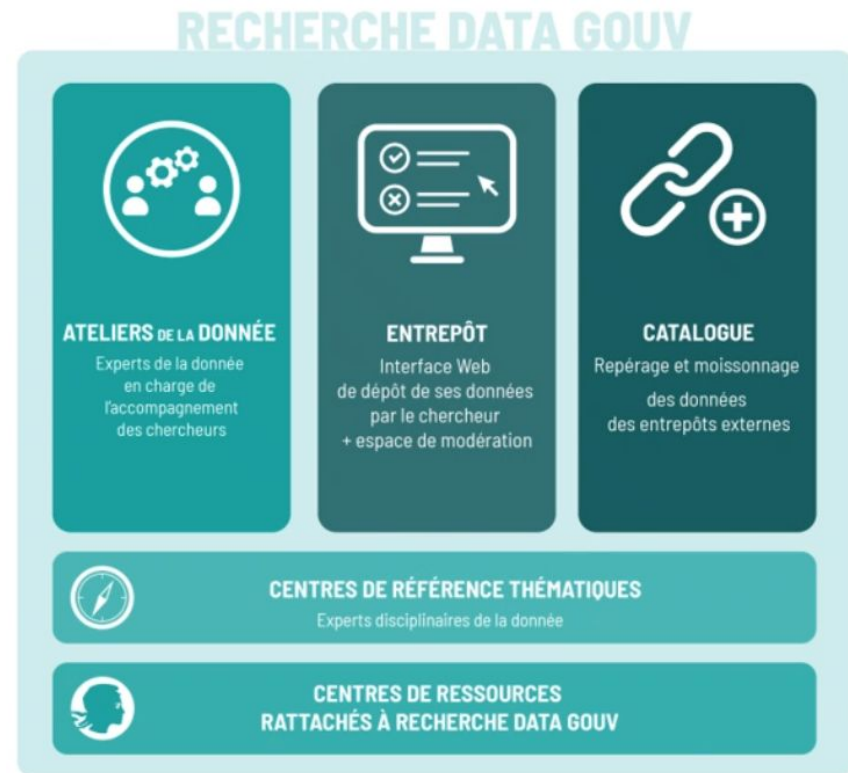
6. DMP integration in SEEK

4. Export and Import to and from RDA format



→ : Implemented feature
 - - - - - : Future feature

- Le MESRI lance le plan [Recherche Data Gouv](#)
 - Composantes
 - Ateliers de la donnée
 - Entrepôt national de données
 - Catalogue des données produites par les équipes françaises
 - Centres de référence thématiques
 - Centres de ressources
 - **L'IFB est désigné Centre de Référence des Données de biologie-santé.**
 - Attentes du MESRI
 - Établir les standards et développer des standards
 - Gestion des données
 - Soins à la donnée
 - Moissonnage des données pour le catalogue
 - Courtage des données
- Animation des communautés biologie santé, des autres INBS, co-construction des standards



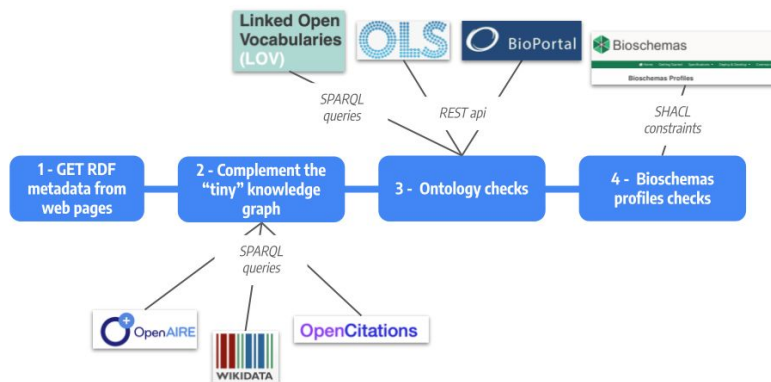
Cet écosystème consiste à fédérer les actions des dispositifs existants et à amplifier leur présence et leurs actions sur tout le territoire et dans tous les domaines thématiques

FAIR-Checker is an online tool

- supporting both data publisher & developers
- leveraging semantic web standards and technologies
- under active development (API, Bioschemas validation, content-negotiation)
- welcoming new contributions (can be easily extended)

Data providers and consumers can check web resources.

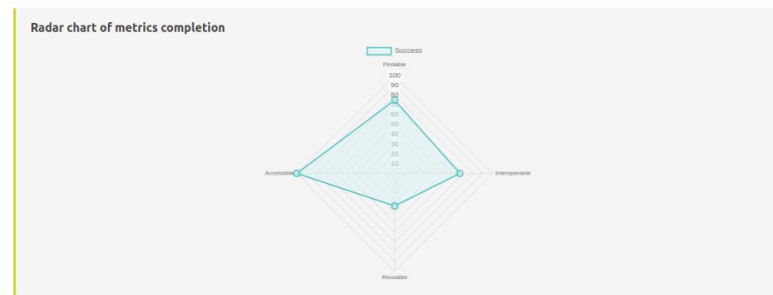
Developers can explore and inspect exposed metadata.



Demo at EOSC Symposium 2023

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

The screenshot shows the input interface of the FAIR checker. It features a text field for entering a resource identifier (URL/DOI), which contains the URL `https://search.datacite.org/works/10.7892/boris.108387`. A "Test all metrics" button is located to the right of the input field. Below the input field, a progress bar indicates the status of the check, and a "Check results" button is visible. At the bottom, there are tabs for "Dataset Database", "Workflow", "Publication Database", "Dataset", and "Tool".



List of metrics with details and results

Principle	Name	Description	Comment	Recommendation	Score	Result	Test	Details
F1A	Unique IDs				2	Success	Check	
F1B	Persistent IDs				0	Failure	Check	
F2A	Structured metadata				2	Success	Check	
F2B	Shared vocabularies for metadata				2	Success	Check	

- Online catalogue
- FAIR principles
- DMP principle
- Integrative Bioinformatics
- Connected to the EU ELIXIR TeSS catalogue

IFB INSTITUT FRANÇAIS DE BIOINFORMATIQUE

Contact Connexion

À propos Services Formation Innovation ELIXIR-FR Catalogue de ressources

Search here..

L'IFB recrute

Catalogue des formations

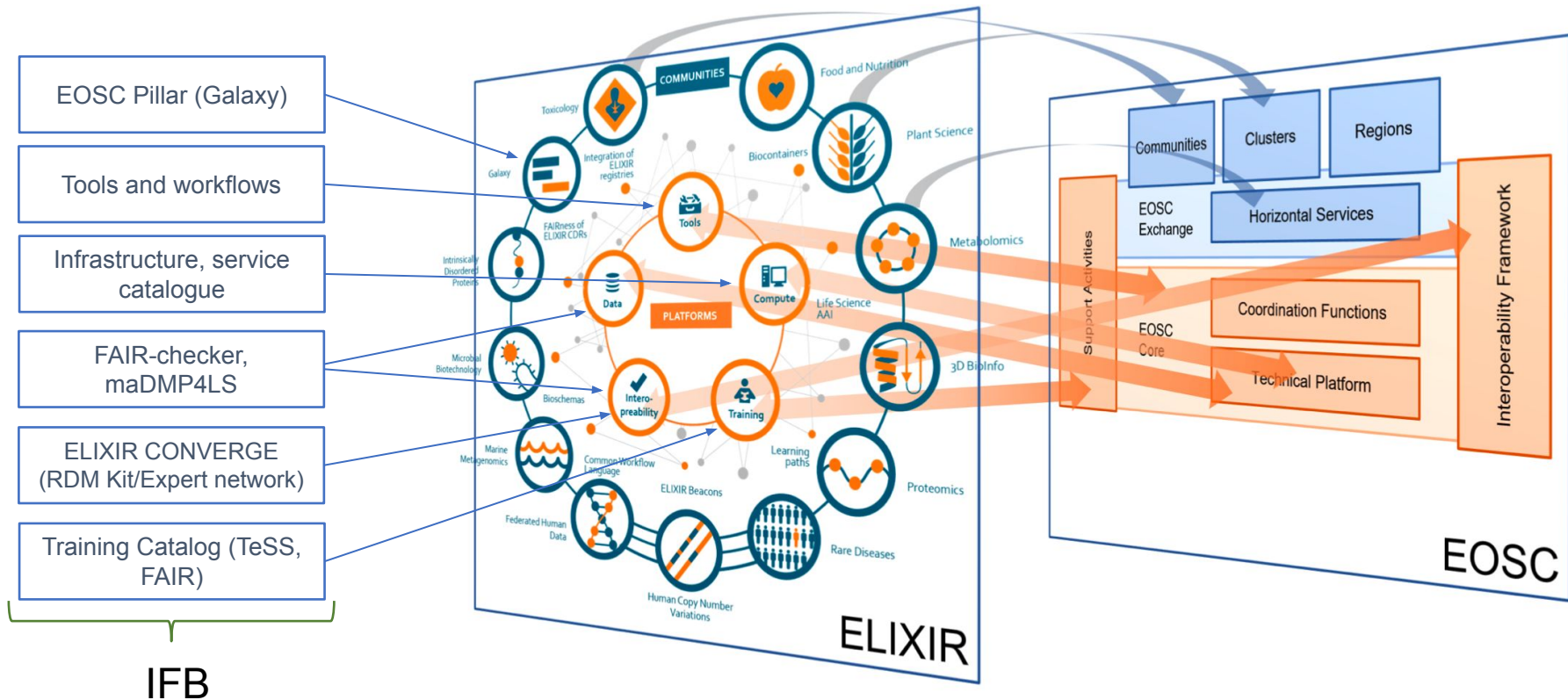
Number of data-items: 415 , show on 21 pages.

Training name

Perform a filter using the legend below

Open for registration Upcoming training Past training

- 02 Mar 2023** NGS data analysis on the command line - Session 5
2023-03-02 | 2023-03-06
Strasbourg France
- 16 Jan 2023** Ecole Thématique de Bioinformatique Intégrative - session 2023 / Integrative Bioinformatics training school - 2023 session
2023-01-16 | 2023-01-20
Fréjus France
- 28 Nov 2022** FAIR_bioinfo_@_AuBi
2022-11-28 | 2022-12-02
CLERMONT-FERRAND France



Bring together experts to **define the strategy** and **provide services** in a particular area.



Interoperability Platform

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



Contact : *M-D Desvignes; A. Gaignard*

Activities 2021:

- The EDAM ontology is on the FR&NO Service Delivery Plan (SDP)
- FR in the steering committee of Bioschemas
- Release of FAIRchecker



Biolmage Informatics Index labelled RIR



Data Platform

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



Contact : *P. Lieby; S. Kossida*

Activities 2021: *Call for projects 2021 To be strategically re-analyzed*



Tools Platform

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



Contact : *H Ménager (ExCo); O. Sallou*

Activities 2021:
-the EDAM ontology is on the FR and NO service delivery plans
-IS Tools Ecosystem, IS Container Services



Compute Platform

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



Contact: *C. Blanchet, J-F Guillaume, J. Pansanel, J. Seiler*

Activities 2021:
- IS Container Services



Training Platform

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



Contact: *O. Sand, H. Chiapello*

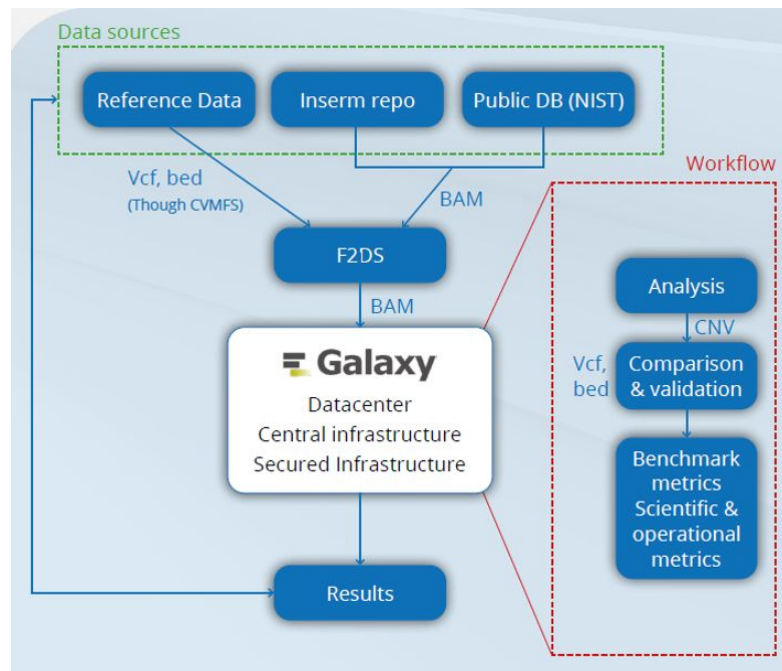
Activities 2021: "Introduction to Machine Learning using R" organised online with ELIXIR-GR and ELIXIR-CH

T6.6: Exploring reference data through existing computing services for the bioinformatics community

- Design the possible interoperations between already available Galaxy computing services and data repositories, serving the Elixir user community as a whole
- Allow access to reference data from different Galaxy deployments for data analysis
- Facilitate the deployment of Galaxy instances close to the data

IFB contribution

- Galaxy-as-a-service in Biosphere cloud catalogue (RAINBio)
- Extending existing recipe with tools and public reference datasets
- hCNV use case



With an infrastructure that is **internationally aligned**, **scalable** to 500 000 life scientists, long-term **sustainable**, **FAIR** data at its heart

Expert Network

- Best practices guidelines
 - Models for brokering data to ELIXIR Deposition Databases
 - Understanding Data Management (DM) business models
 - Sustainable and scalable operating model for harmonised DM in European projects
- *A Data Management Coordinator Network*
- *Beyond 2023: integrate one of ELIXIR platforms*



ELIXIR-Converge has received fundings from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 871075

A portal: RDMkit – <https://rdmkit.elixir-europe.org/>

RDMkit Data management About Contribute GitHub

The Research Data Management toolkit for Life Sciences
Best practices and guidelines to help you make your data FAIR (Findable, Accessible, Interoperable and Reusable)

What can we help you find?
Search RDMkit

Browse all topics by

- Data life cycle**
Start here to get an overview of research data management based on stages in the data life cycle.
- Your role**
Identify your role in research data management, find data management resources relevant for you, and information to help you progress in your career path.
- Your domain**
Learn about data management tasks that affect your domain or research community, and the solutions adopted to address them.
- Your tasks**
Find guidelines and solutions for tackling common data management tasks.
- Tool assembly**
Find concrete combinations of tools and resources assembled into an ecosystem for research data management.
- National resources**
Find pointers to country specific information resources and national research data management practices.
- All tools and resources**
Browse the RDMkit's catalogue of tools and resources for research data management.
- All training resources**
Browse all training resources mentioned in RDMkit pages.

- Guidance: linked guidelines, tools, training
- Best practices
- Examples

Open contribution and editorial process: 156 contributors, 345 tools/resources, 105 pages



INSTITUT FRANÇAIS DE BIOINFORMATIQUE



france-bioinformatique.fr



elixir-europe.org



@IFB_Bioinfo



@InstitutFrançaisDeBioinformatique



contact@france-bioinformatique.fr

