



ELIXIR Infrastructure: *Advancing Management, Sharing, and Analysis of Life Sciences Data across Europe*

X [@elixir_it](https://twitter.com/elixir_it)

in [/company/elixir-iib/](https://www.linkedin.com/company/elixir-iib/)

Pier Luigi Martelli
Bologna Biocomputing Group

www.elixir-italy.org/



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ELIXIR – who we are

ELIXIR is an intergovernmental organisation, formally established in **2016** as a **Landmark European Research Infrastructure (ESFRI Roadmap)**. The Roadmap has been updated in **2021** and consists of **22 ESFRI projects** and **41 ESFRI Landmarks**.

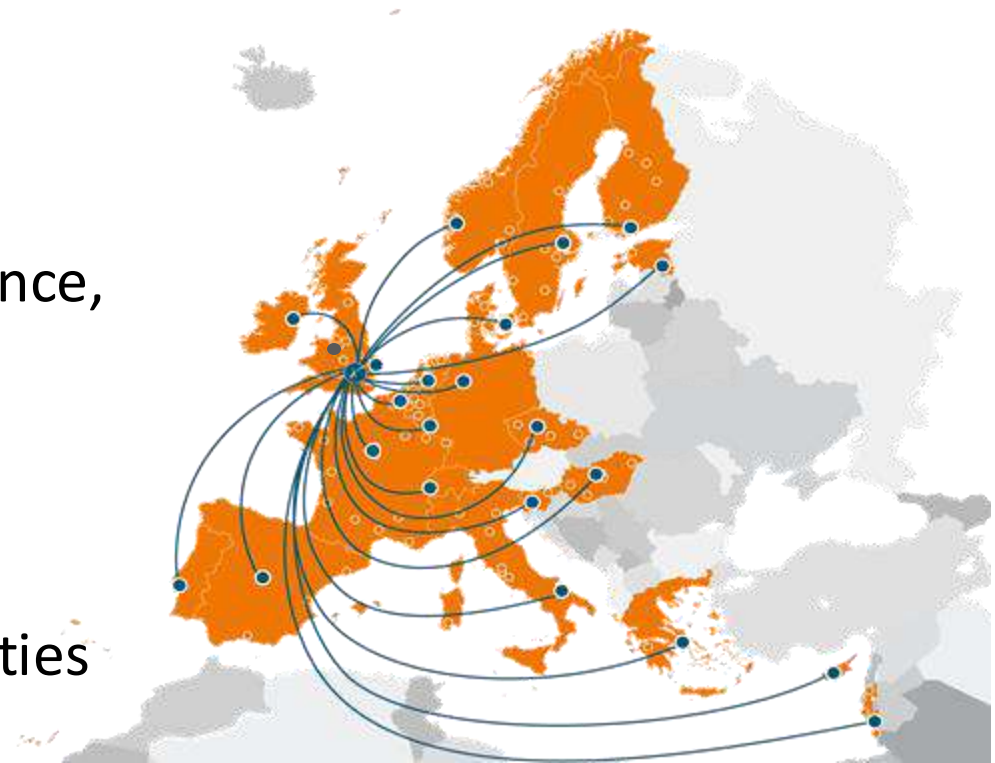
ELIXIR, brings together “bioinformatic resources” for life sciences from across Europe, including **databases, software tools, training materials, best practices, cloud storage and supercomputers.**

23 Nodes

ELIXIR Nodes form national networks of excellence, containing a total of **245** institutes/universities

1 Hub

The Hub, located at EMBL-EBI, Cambridge, UK, provides the secretariat and coordinating activities for ELIXIR



ELIXIR-IT

The **Italian Node of ELIXIR (ELIXIR-IT)** is formally established as as a **Joint Research Unit (JRU)** through a specific agreement among the coordinating organization (National Research Council, CNR) and the other 30 members. Apart from the CNR, the largest public research institution in Italy, the other 30 members are composed of 16 Universities, and 14 scientific and technological institutes that, in most cases are leaders in the country in their respective fields of expertise (e.g. CINECA and INFN in supercomputing, ISS in public health, GARR in IT infrastructure for research, TIGEM for rare diseases).

The **leading Institution (CNR)** has been designed by the Ministry of University and Research.

The **organisational layout** of the node has three layers: 1) policy and oversight; 2) coordination; and 3) operational. The General Assembly (GA), advised by an international SAB, is the main body of the Node, leaded by the HoN (designated by the President of CNR) and including delegates designated by the legal representatives of participating institutions. The coordination and operational layer, including the staff involved in all ELIXIR-IT activities includes both in-kind and full-time staff (about 20 FTE in 2023).



31 Members



ELIXIR-IT – EPTRI-IT agreement

An operational agreement has been recently signed between EPTRI-IT and ELIXIR-IT

Goals

- To foster the development of innovative solutions in paediatric data management and analysis
- To promote the adoption of FAIR and Open Science principles in paediatrics
- To promote the integration of paediatric research in national and European infrastructures
- To drive digital innovation in health in paediatrics
- To promote the joint application to funding calls



ELIXIR - How we work

Platforms



The five ELIXIR Platforms develop and coordinate cross-domain services across ELIXIR.

Communities



ELIXIR Communities work in a particular life science domain and give feedback on the Platform services.

Focus Groups



Focus Groups are informal groups that look at emerging areas of interest in life science. Most are open to non-ELIXIR members.

Internal projects



Internal projects fund the Platforms and Communities to develop resources, and encourage collaboration across the Nodes.

Coordinators



Technical and Training Coordinators advise the Platforms and promote collaboration between the Platforms and the Nodes.



EU Projects

ELIXIR both collaborates in and coordinates EU projects.



Industry and innovation

ELIXIR organises events and projects that bring the private and public sectors together.



International engagement

ELIXIR collaborates extensively within and beyond Europe, and works with key research infrastructures.

<https://elixir-europe.org/how-we-work>



ELIXIR Platforms

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

Compute

Builds and integrates **cloud, compute, storage and access services** for the life-science research community

Data

Drives the use, re-use and value of life science data by providing **robust, long-term sustainable data resources** within a coordinated, scalable and connected data ecosystem

Tools

Helps communities **find, register and benchmark** software tools; maintains information standards and produces, adopts and promotes **best practices for tool development**

Interoperability

Helps people and machines to **discover, access, integrate and analyse biological data**; encourages the life science community to adopt **standardised file formats, metadata & vocabularies**

Training

Strengthens **national training** programmes; grows bioinformatics **training capacity and competence across Europe**; empowers researchers to use ELIXIR's services and tools

Omics (Italy only)

Provides services for **genomics, transcriptomics, metabolomics** and **proteomics** data production



ELIXIR Communities –

connecting infrastructure & life science experts



Formed around domain experts in ELIXIR Nodes
(including non-ELIXIR partners)



Provide a mechanism for long-term collaborations with
other ESFRIs and large-scale initiatives



Drive service developments in the ELIXIR Platforms



Provide a framework to develop and maintain
community standards



The [ELIXIR Communities Handbook](#) tells you what a
Community is, who can join, what the benefits are, and
how Communities are structured.



ELIXIR Infrastructure Landscape

Communities & Focus groups



3D-BioInfo

Helps to understand the 3D structure of macromolecules like proteins and DNA.



Food and Nutrition

Aims to help us understand the effect of food choices on human health.



Metabolomics

Provides the resources, analysis tools and infrastructure to help metabolite identification.



Microbiome

Promotes standards, training, and other resources for microbiome research across all biomes.



Proteomics

Develops and maintains sustainable proteomics tools and data resources.



Research Data Management

Brings together RDM experts to coordinate and promote the use of RDM resources.



Systems Biology

Aims to make systems biology modelling a central pillar of research in biology.



Biodiversity

Aims to help us understand biodiversity, and how it responds to changing environmental pressures.



Galaxy

Fosters a Galaxy community in Europe, together with Galaxy resources and training.



Microbial Biotechnology

Helps the development of tailor-made microbes and biological systems.



Plant Sciences

Develops an infrastructure to facilitate genotype-phenotype analysis for crop and tree species.



Intrinsically Disordered Proteins

Develops standards, tools and resources to help identify and characterise IDPs.



Single-Cell Omics

Identifies and addresses challenges in single-cell and spatial omics.



Toxicology

Supports the integration of standards, tools and resources to aid toxicology research projects.

Human Data Communities



Federated Human Data

Develops long-term strategies for managing and accessing sensitive human data.



Human Copy Number Variation

Aims to make it easier to detect, annotate and interpret human Copy Number Variations (hCNVs).



Cancer Data

Enhancing data for precision medicine by improving data management, interoperability and accessibility.



Rare Diseases

Supports the development of new therapies for rare diseases.

Current Focus Groups

- Biocuration
- Domestic Animals Genome and Phenome
- Environmental Impact
- EOSC
- FAIR Training
- Health Data
- Learning Paths
- Machine Learning (AI)
- Pathogen Data
- Professionalising Careers in Research Infrastructures
- Research Data Alliance (RDA) Activities

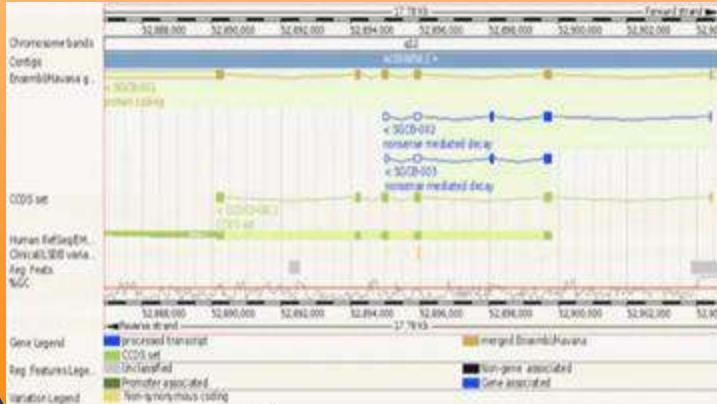
Next presentation

Data Platform

The ELIXIR Data Platform supports:

Core Data Resources

Databases of fundamental importance for life sciences research and involved in the long-term conservation of biological data.



Deposition Databases

Recommended repositories for the deposition of experimental data.

Dataset ID	Description	Technology
EGAD000000000001	WTCCC1 project samples from 1958 British Birth Cohort	Affymetrix 500K
EGAD000000000002	WTCCC1 project samples from UK National Blood Service	Affymetrix 500K
EGAD000000000003	WTCCC1 project Bipolar Disorder (BD) samples	
EGAD000000000004	WTCCC1 project Coronary Artery Disease (CAD) samples	
EGAD000000000005	WTCCC1 project Inflammatory Bowel Disease (IBD) samples	
EGAD000000000006	WTCCC1 project Hypertension (HT) samples	
EGAD000000000007	WTCCC1 project Rheumatoid arthritis (RA) samples	
EGAD000000000008	WTCCC1 project Type 1 Diabetes (T1D) samples	
EGAD000000000009	WTCCC1 project Type 2 Diabetes (T2D) samples	
EGAD000000000010	WTCCC1 project Ankylosing Spondylitis (AS) samples	Illumina 15K
EGAD000000000011	WTCCC1 project Autoimmune Thyroid Disease (ATD) samples	Illumina 15K
EGAD000000000012	WTCCC1 project Multiple Sclerosis (MS) samples	
EGAD000000000013	WTCCC1 project Breast cancer (BC) samples	Illumina 15K

Data Resource Services

List of services provided by different nodes.

Services: Data resources

Search this section:

Name of service	Tag	Related links*
AspicDB	1	Data links
Berlin RNA Toolbox	1 Genes and genomes	
CHOPCHOP	1	Data links
ENZYME	1 Enzymes, interactions and pathways	Data links
Glyco@Expasy	1 Proteins and proteomes	Data links
HCViVdb	1 Molecular and cellular structures	Data links
ITSoneDB	1 Genes and genomes	Data links

ELIXIR Data Platform facilitates the access to vast Omics datasets, adhering to **FAIR** principles and embracing **Open Science**. This approach fosters the development of innovative applications, products, and services

Data Platform – Some of the core resources

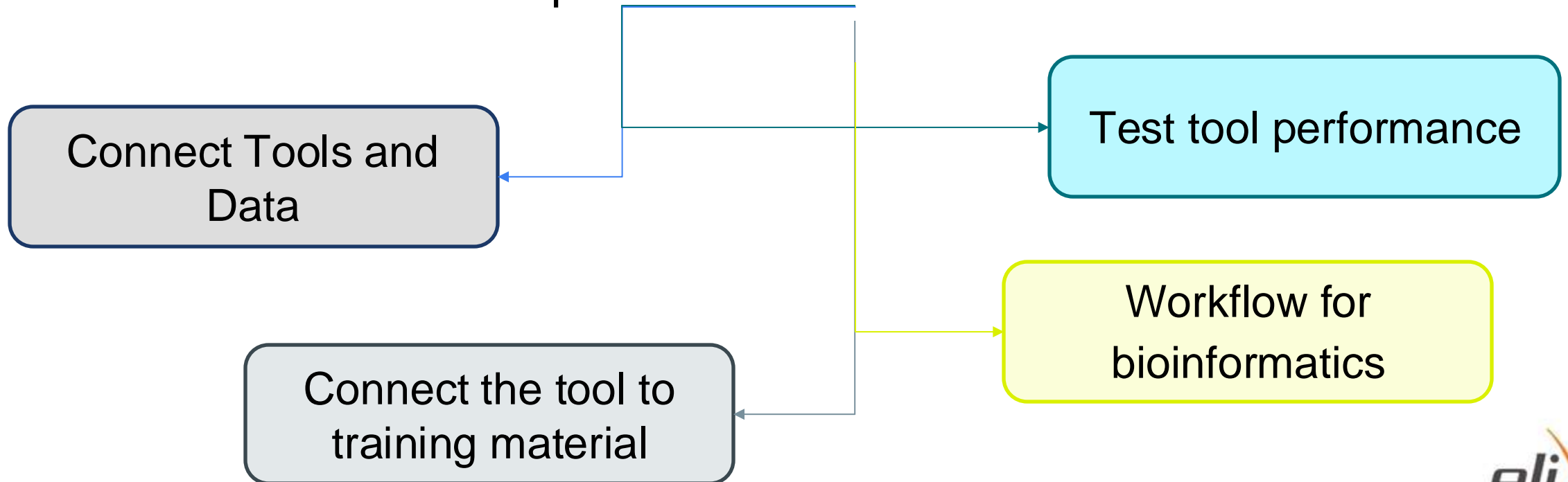


<https://elixir-europe.org/platforms/data>



Tools Platform

- Facilitate access to and use of bioinformatics tools
- Define and promote the use of standards for tool documentation (tool descriptions) and associated metadata
- Define and promote Best Practices for the development and maintenance of developed tools



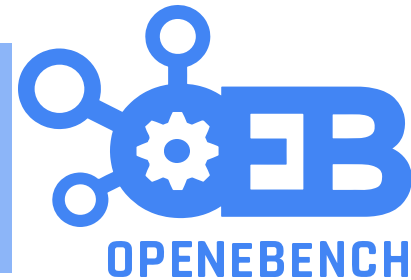
Tools Platform



Bio.tools is a registry of tools and databases (~30k), managed by the community. By using controlled ontologies, it allows users to quickly find the most appropriate tool for their needs.

WorkflowHub is a domain-independent registry for workflows designed based on FAIR principles.

OpenEBench represents a platform for benchmarking tools and FAIRification of reference datasets. It is also designed to support the principles of Open Data, Open Source, and Open Science.



1. **Biocontainer** for gathering individual efforts for tool containerization and platform federation for sustainability.
2. Central repository for managing containers and their distribution.
3. User-driven development to identify the most suitable solutions for users.

Compute Platform

- It is responsible for **providing high-level Computing and Storage services** specialized in supporting the applications of the Elixir community.
- It is composed of several technological partners whose know-how and infrastructure represent the backbone of the node's ICT infrastructure.

Current Technological Landscape

The latest HW and SW technologies are used to implement, through intense **co-design activities with the bioinformatics community**, specialized services that are as closely aligned as possible with the requirements of researchers.

Compute Platform – ELIXIR-IT



CNR Milano:

~ 5000 CPU cores
5 PB storage
(Mirror)

CNR Napoli:

5 PB storage
(Backup)

UniMi

INDACO Update

UniBo

IT infrastructure
update

UniPd

IT infrastructure
update



CNR and University of Bari:
Sequencing facility

RECAS BARI



INFN-Bari:

4.192 CPU cores
2,1 PB storage; 10Gb network

CNR:

12.320 CPU cores, 10 GPUs
7,2 PB storage; 25Gb network



UniBa (Physics Dept.)

4000 CPU cores and GPUs
5.5 PB Cloud Storage 2
PB Posix Storage 20 PB
Tape Library



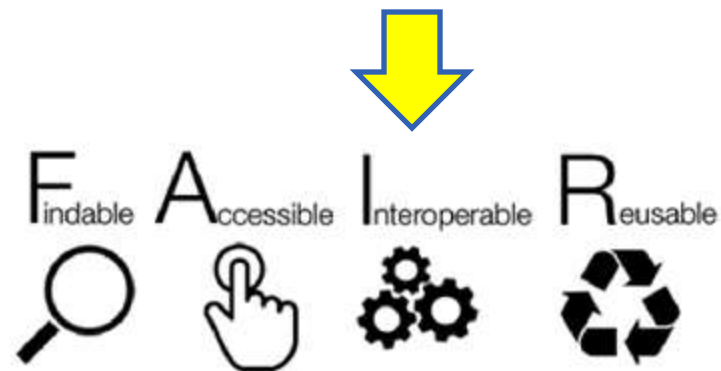
Uniba (Physics Dept.)

5 PB Storage HPC/HTC
250 CPU cores HPC with GPUs
1.5 PB Storage Cloud

Interoperability Platform

What is needed to efficiently connect the Data, Tools and Computing facilities?

- **Standards:** formats reporting, guidelines, vocabularies, ontologies
- **Metadata services:** ontologies, annotations, validation, indexing, search
- **Registers:** services and datasets
- **Search engines** for tools and datasets
- **Identifiers:** resolution and management
- **Harmonisation and sharing** of tools, workflows and pipelines
- Common **programmable interfaces**
- **Best practices**



Interoperability Platform - ELIXIR-IT Services

<https://elixir-europe.org/platforms/interoperability/rirs>



Catalog of standards and usage and management policies for data and metadata

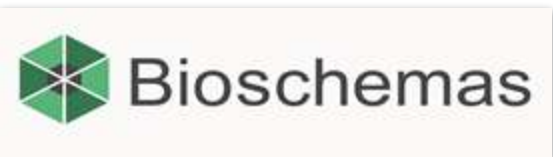
<https://fairsharing.org>



Practices, recipes, and guidelines for making data FAIR

<https://rdmkit.elixir-europe.org/>

<https://faircookbook.elixir-europe.org>



Markup models to improve the findability of resources on the Web (datasets, software, training)

<https://bioschemas.org>



Web resource resolution service that provides consistent and persistent access to data with compact identifiers

<http://identifiers.org/>

Interoperability Platform - ELIXIR-IT Services

<https://elixir-europe.org/platforms/interoperability/rirs>



Benchmarking system for methods, tools, and bioinformatics web services

<https://openebench.bsc.es>



Repository of biomedical ontologies providing a single access point

<https://openebench.bsc.es>



Interoperable open-source platform for scientific data analysis and sharing

<https://usegalaxy.eu/>



Open standard for the description of workflows and analysis tools

<https://www.commonwl.org>



Data warehouse system integrating biological data sources and facilitating data interrogation and analysis

<http://intermine.org>

ELIXIR-IT Training Platform Mission

Training in bioinformatics is essential to achieve excellence in scientific research

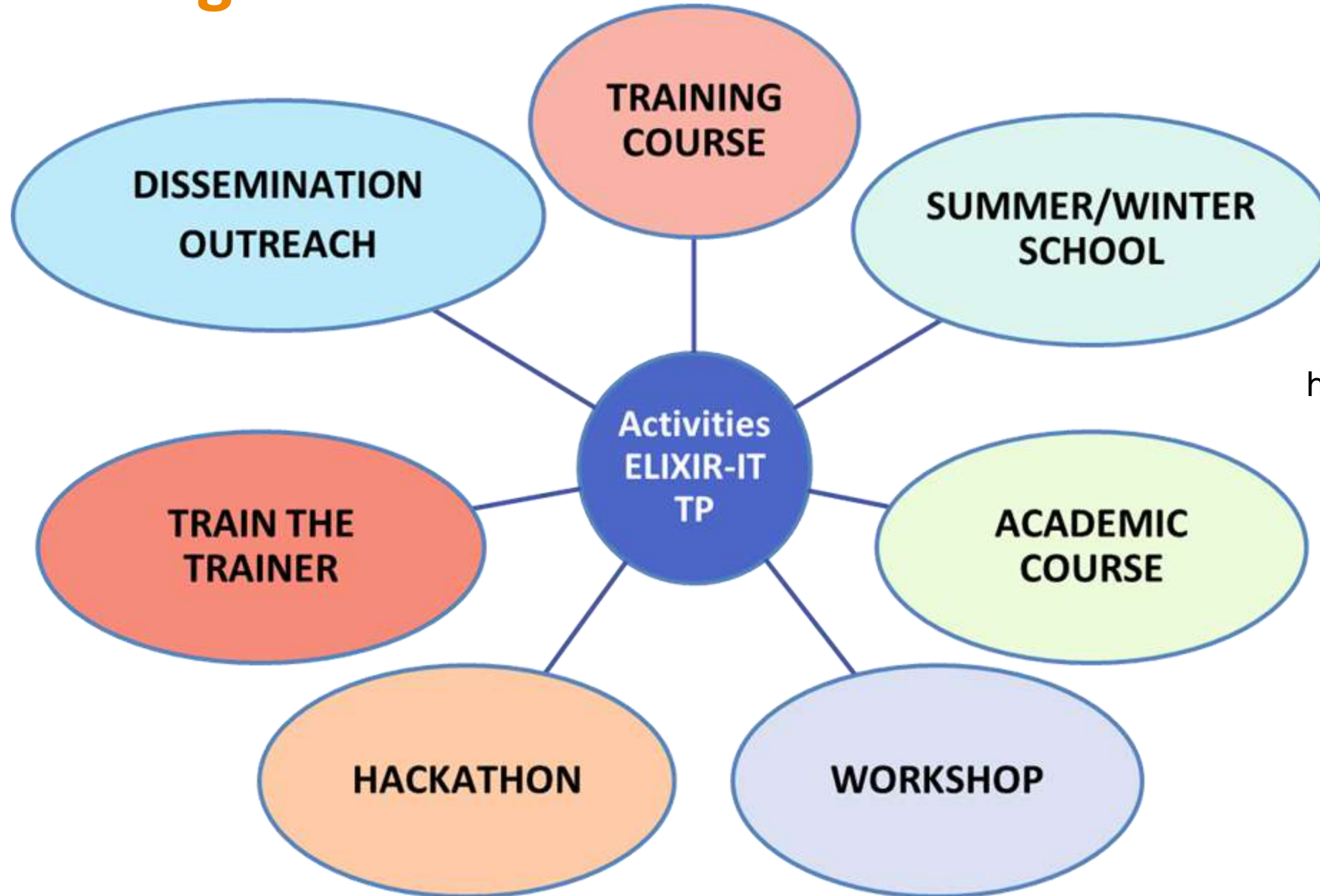
Support researchers and professionals in acquiring bioinformatics skills

Promote effective teaching and learning practices

Establish national and international collaborations

Create a community of instructors and educators

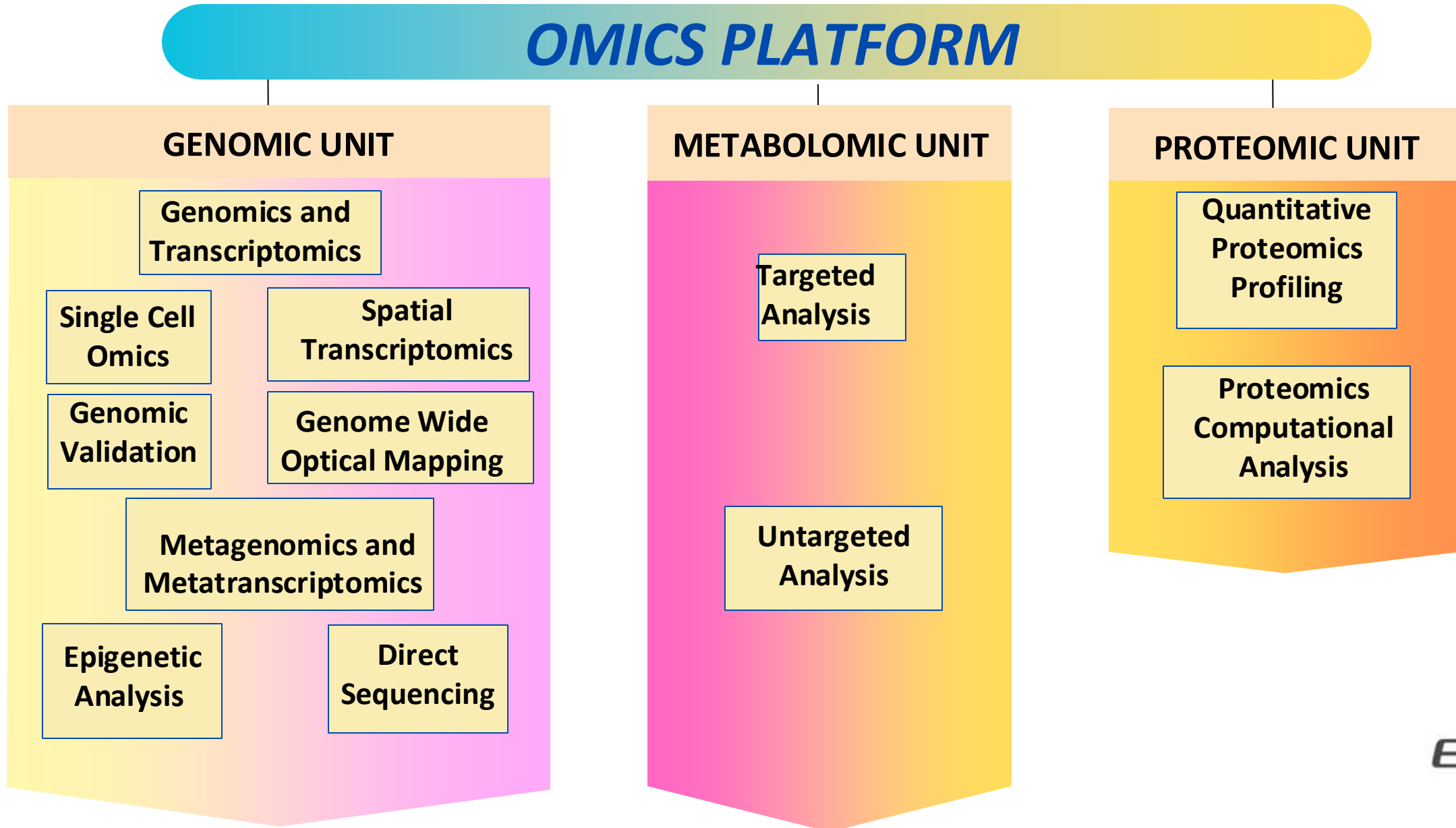
Training offer



<https://tess.elixir-europe.org/>




The Omics Platform



ELIXIR Services

The Italian node of ELIXIR offers a vast portfolio of high quality bioinformatic services including data, tools, compute, standards, compute and training provision as well as “Omics” data production.



ABOUT US ▾ SERVICES ▾ HOW WE WORK ▾ EVENTS ▾ NEWS INTRANET LOGOUT

Home » Services »

SERVICES

Service list

ELIXIR support to COVID-19 research

Scientific domain
<Any>

Key service collection
<Any>

Type of service
<Any>

ELIXIR Node
ELIXIR Italy

Search all

Search

Name of service	Description	Related links*
Argot	Annotation Retrieval of Gene Ontology Terms using semantic similarity in Gene Ontology.	bio.tools
AspicDB	Alternative splicing database of human genes and functional annotation of their isoforms using the ASPic algorithm.	bio.tools
BaCellLo	Predicts the subcellular localization of proteins in eukaryotes.	bio.tools
Cloud Services	<ul style="list-style-type: none">Cineca Cloud farmCloud@ReCaS-BariCloud VenetoCRS4 Orione Cloud farmGARR Cloud	



UNIVERSITÀ DI BOLOGNA

BaCellLo

WEB APPLICATION

Predictor for the subcellular localization of proteins in eukaryotes.

Model organisms, Plants, Protein targeting and localisation



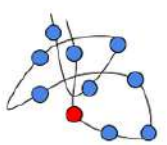
UNIVERSITÀ DI BOLOGNA

SNPs And GO

WEB APPLICATION

A server for the prediction of single point protein mutations likely to be involved in...

Data mining, Genetic variation, Protein sites, features and motifs, Rare diseases





UNIVERSITÀ DI BOLOGNA

INPS-MD

WEB APPLICATION

Predicting the impact of mutations on protein stability from sequence and structure.

Genetic variation, Protein folding, stability and design, Rare diseases

12 services provided by UniBO



608 services provided by ELIXIR

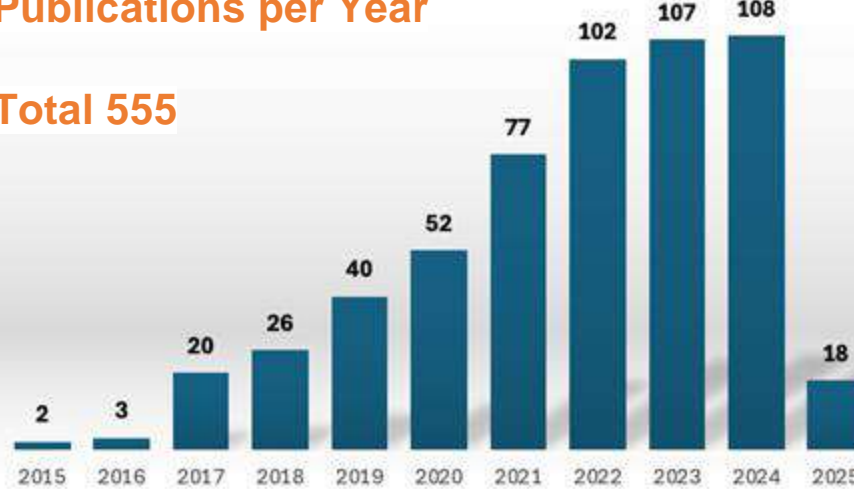
69 services provided by ELIXIR-IT

ELIXIR-IT Impact

Last update February 2025

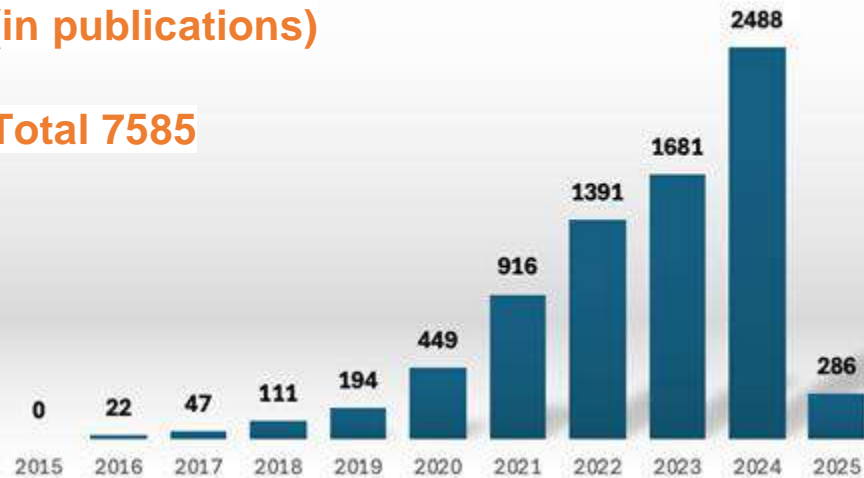
Publications per Year

Total 555



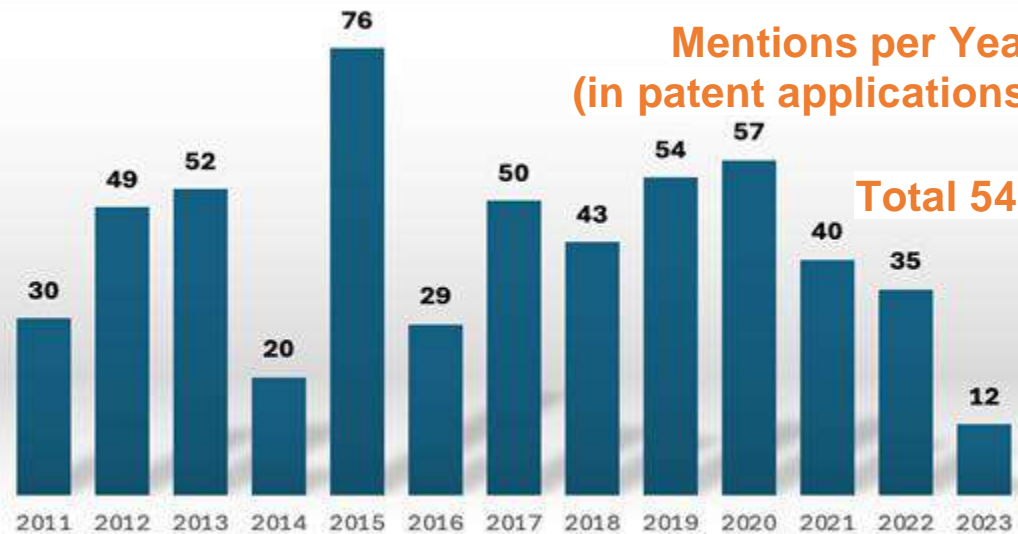
Citations per Year (in publications)

Total 7585



Mentions per Year (in patent applications)

Total 547



National and international projects empowered by ELIXIR-IT



Thanks for your attention



ELIXIR IIB Contacts

Website

<https://elixir-italy.org>

ELIXIR IT Newsletter Subscription



[/company/elixir-iib/](https://company/elixir-iib/)



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



[@elixiritaly3395](https://www.youtube.com/@elixiritaly3395)



elixir.it.iib@gmail.com