elixir
ITALY
DATA FOR LIFE

03 *Introduction*

04 *ELIXIR*

06 *ELIXIR-IT*

08 *Organization*

10 *Organigram*

13 *Impacts*

14 *Industry Strategy*

16 *Data Platform*

19 *Tools Platform*

20 *Interoperability Platform*

23 *Training Platform*

24 *Compute Platform*

26 *Team*

INTRODUCTION

Life sciences research is becoming significantly more and more impactful on many aspects of our daily lives, including healthcare, food quality and safety standards, agriculture and environment preservation.

Several groundbreaking research avenues, mostly involving scientists working in multidisciplinary teams from many nations, are generating an extraordinary amount of data, particularly from large-scale DNA sequencing projects. These datasets then need to be properly analyzed, stored, and shared among researchers, in order to foster further research in all aspects of life sciences. The potential of big data is one of the major drivers of innovation in all fields of life sciences.

Accessing and analyzing biomolecular data is becoming progressively easier and more cost-effective with the recent impressive growth of cloud computing. Also, Europe's bioindustries are becoming much more streamlined in their R&D processes through intensive use of biomolecular data and bioinformatics analysis. Indeed, untangling the interplay between genes and environment and elucidating at the molecular level all biological processes is crucial for new advances in personalized medicine, healthy ageing, earth ecology and to obtain significant breakthroughs in biotechnological applications.

None of this would be possible without the dedication and passion of thousands of scientists; all of whom depend on databases, software tools, cloud resources and supercomputers, interoperability standards, and specialized training. All these tools and systems require a common and coordinated effort at national and international level.

The goal of **ELIXIR** (www.elixir-europe.org) is to address this challenge and establish a **European Research Infrastructure for life sciences data** in order to make it easier for scientists to find and share data, exchange expertise, and agree on best practices.

ELIXIR-IT is the Italian Node of ELIXIR; its mission is to drive the development of bioinformatics resources in Italy and integrate them into ELIXIR's pan European service portfolio.

ELIXIR

ELIXIR is an European Research Infrastructure that brings together life sciences resources from across Europe. These resources include **databases, software tools, best practices, training materials, cloud storage and supercomputers.**

ELIXIR coordinates these resources so that they form a single infrastructure, which makes it easier for scientists from Academia, Research Institutions and commerce to find and share data, exchange expertise and agree on best practices.

ELIXIR adopts a "Hub and Nodes" organization model: The **ELIXIR Hub** (Wellcome Genome Campus, Cambridge, UK) operates as the headquarters that coordinates the activities across ELIXIR.

Each of the participating member states of ELIXIR establishes a national 'Node', which in turn is a network of organizations that work within a member state.

ELIXIR has 22 Members and 1 Observer: 22 European countries and EMBL-EBI as international organisation.

ELIXIR-IT

The Italian Node of ELIXIR (**ELIXIR-IT**) has been formally established as a **Joint Research Unit** (JRU) named ELIXIR - Infrastruttura Italiana di Bioinformatica (ELIXIR-IIB), which is in charge of the coordination and delivery of existing bioinformatics services at a national level, all the while ensuring they integrate perfectly in the overall ELIXIR infrastructure.

ELIXIR-IT is led by the National Research Council (CNR) and currently involves 23 partners including Universities and Research Institutions / Facilities of national relevance.

As every ELIXIR member, Italy participates the ELIXIR Board, the highest decision-making body in ELIXIR, with one scientific (Prof. Rita Casadio) and one admin representatives. *(Dott. Salvo La Rosa)*

Head of Node

Graziano Pesole
(CNR)

"The aim of ELIXIR-IT is to bring together and coordinate the best expertise and infrastructural resources for life sciences data management and processing across our country, and to make them fully accessible to academic and industrial researchers".

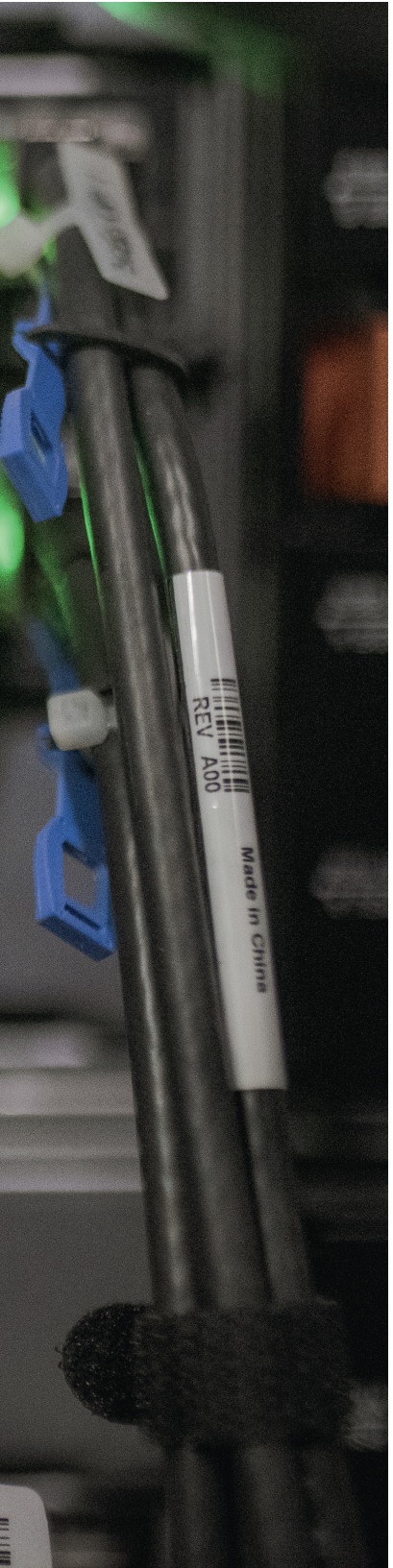
ELIXIR Board Member

Rita Casadio
(UNIBO)

"Our Italian Node (ELIXIR-IT) contributes to the ELIXIR infrastructure for bioinformatics resources across Europe with unique expertises, shared data bases, artificial intelligence based methods for big data analysis and end-to-end modelling of biological complex systems".

CNR
ENEA
CINECA
Istituto Nazionale Fisica Nucleare
CRS4
GARR
Stazione Zoologica Anton Dohrn
Fondazione Edmund Mach
Istituto Superiore di Sanità
TIGEM
Università di Torino
Università di Napoli
Università di Salerno
Università di Parma
Università di Firenze
Università di Bari
Università della Tuscia
Università di Padova
Università di Bologna
Università di Milano
Università di Milano, Bicocca
Università Roma, Tor Vergata
Università Sapienza, Roma

ELIXIR-IT
Organization



ELIXIR is organized in **operational Platforms** and **thematic communities** all of which help in connecting experts sharing a strategic interest into a specific domain.

The national organisation of ELIXIR Italy mirrors the organisation of ELIXIR at the European level. ELIXIR-IT includes five operational Platforms (**Data, Tools, Compute, Interoperability and Training**) and several **thematic communities.**

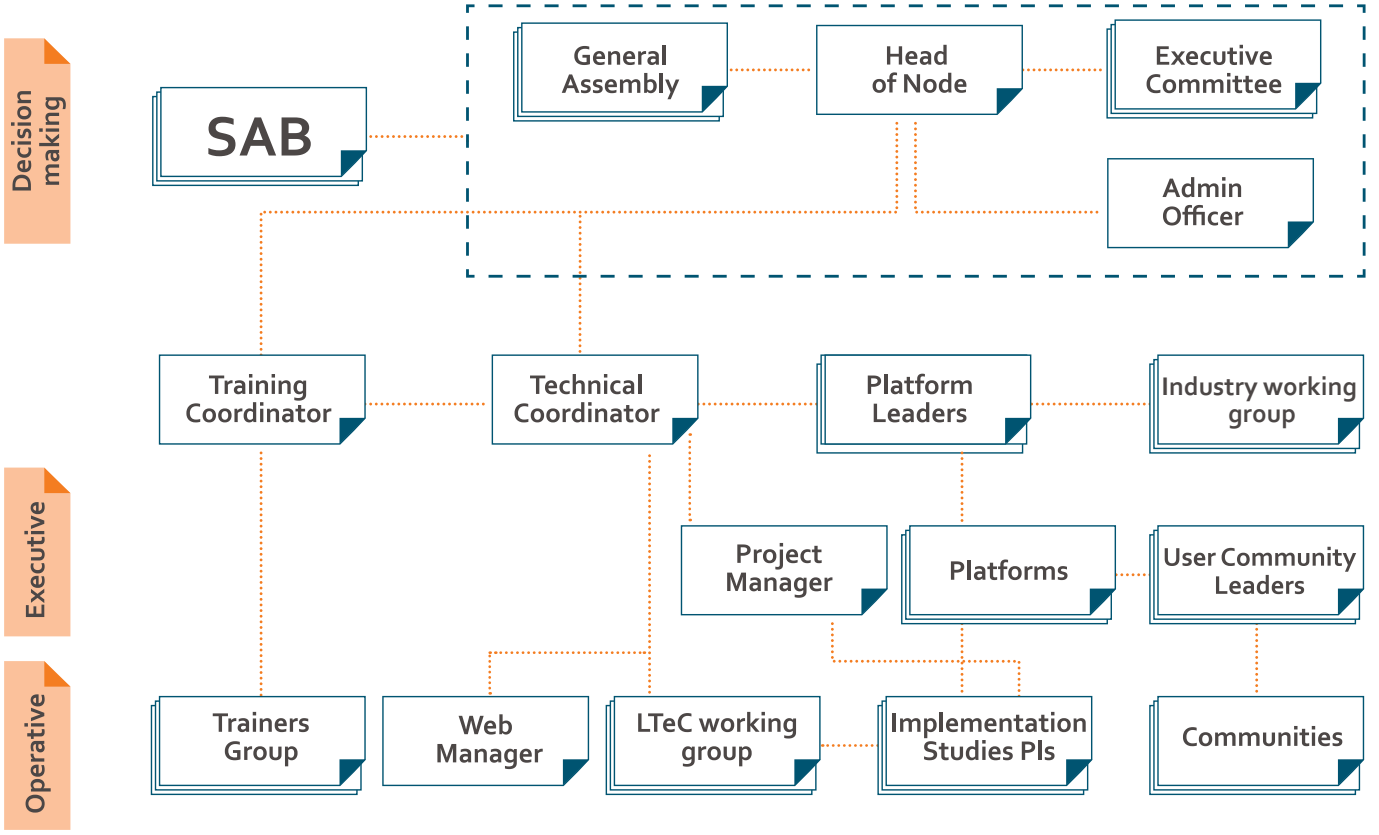
ELIXIR-IT Platforms coordinate the delivery of high-quality computational services for the life sciences and drive the integration of national services within the ELIXIR infrastructure.

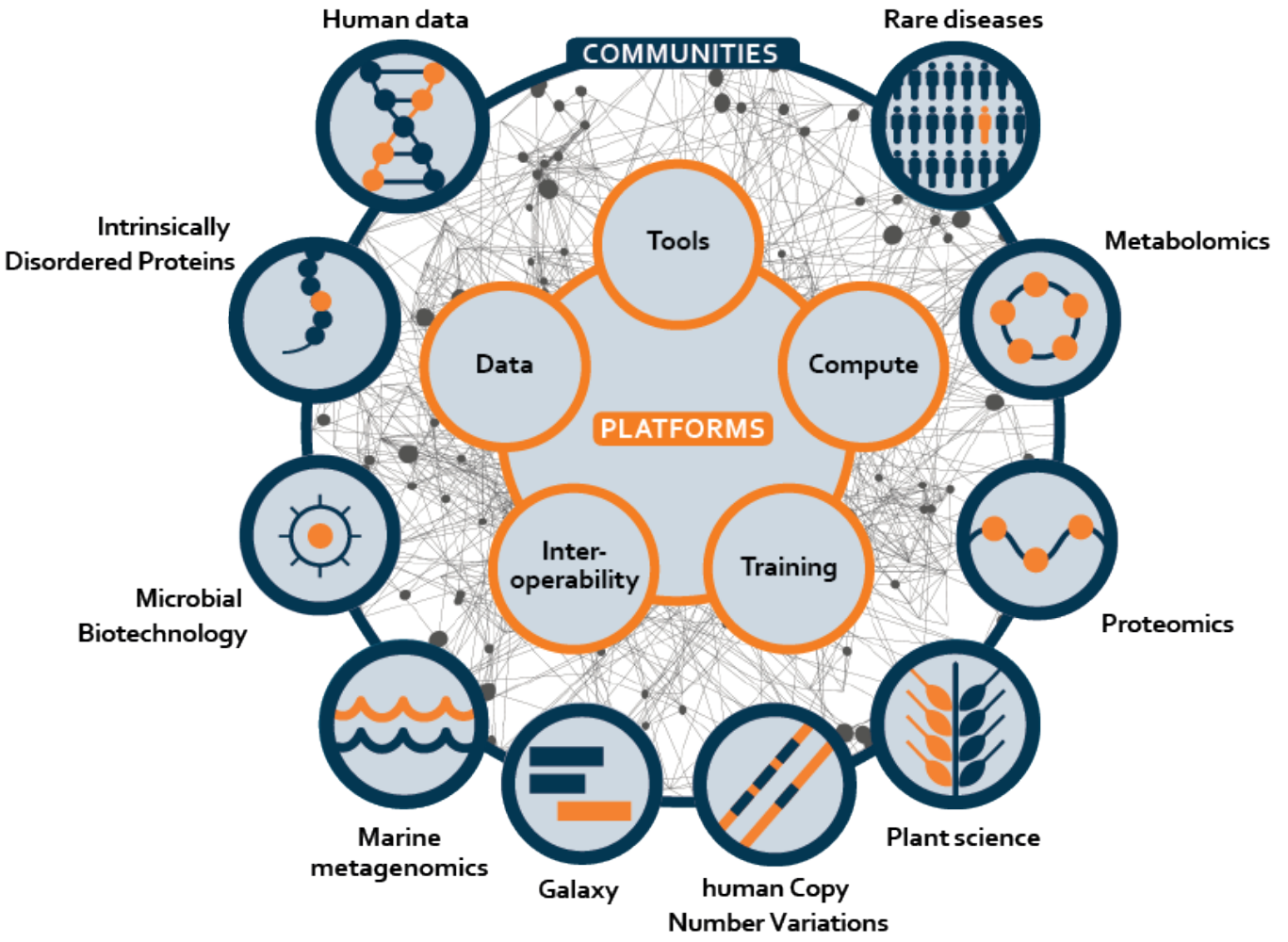
The **JRU** is managed by the Head of Node, abiding to the policies drawn by the JRU Board (Assemblea Generale) and the Management and Coordination Committee (Comitato di Gestione e Coordinamento).

The implementation of strategic decisions and the day to day operations are carried out by the Local Technical Coordinators group (LTeCG), led by the Technical and Training Coordinators, and by the Platform Coordinators and their deputies.

An international Scientific Advisory Board (SAB) acts as a consultative body to provide recommendations on technical, scientific and strategic matters related to the activities, achievements and long-term goals of the Italian ELIXIR Node.

ELIXIR-IT
Organigram




Human data
COMMUNITIES
Rare diseases
Intrinsically
Disordered Proteins
Tools
Metabolomics
Data
Compute
PLATFORMS
Inter-
operability
Training
Microbial
Biotechnology
Proteomics
Marine
metagenomics
Galaxy
human Copy
Number Variations
Plant science

Covaris
M220

Impacts

Impacts on the national bioinformatics community: there are many excellent researchers and groups in Italy working in bioinformatics and related fields. However, a lack of coordination at the national level has often limited the possibilities for extensive country-wide projects that would gain wide international visibility. Thus, the establishment of the Italian ELIXIR Node, part of a supranational entity like ELIXIR, is expected to strongly influence many aspects of the national bioinformatics community e.g. by fostering collaborations and data sharing, spreading the adoption of best practices and shaping the next generation of bioinformaticians through the ELIXIR training programme.

European added-value of the Node: the Italian ELIXIR Node provides a rich portfolio of unique and valuable services, resources and expertise to the European bioinformatics and user communities. Of particular relevance is the comprehensive training program, the vast majority of which is open and free of charge to European life sciences researchers, and free access to high performance computational resources to European users as demonstrated by services like HPC@Cineca. There are also a wide array of data and tools resources that are maintained by skilled and passionate scientists. Notable also is the scientific transversality underlying the Italian ELIXIR Node, for example the National Institute for Nuclear Physics (INFN) is part of the Node too. What sets this apart is that it fosters the sharing of different expertise, points of view and solutions to the challenges that bioinformatics has in common with other scientific fields, such as big data handling and cloud computing.

Economic and societal impacts of the Node: the Italian ELIXIR Node will drive the adoption of shared standards and best practices, the training of skilled professionals and the development and maintenance of state-of-the-art data and tools resources. In turn this shall improve the scientific efficiency of Italian research in life sciences and related fields like translational medicine, pharmaceutical and agri-food industries, environment protection, and so on, with a relevant positive social and economic fall-out.

ELIXIR-IT
Industry Strategy

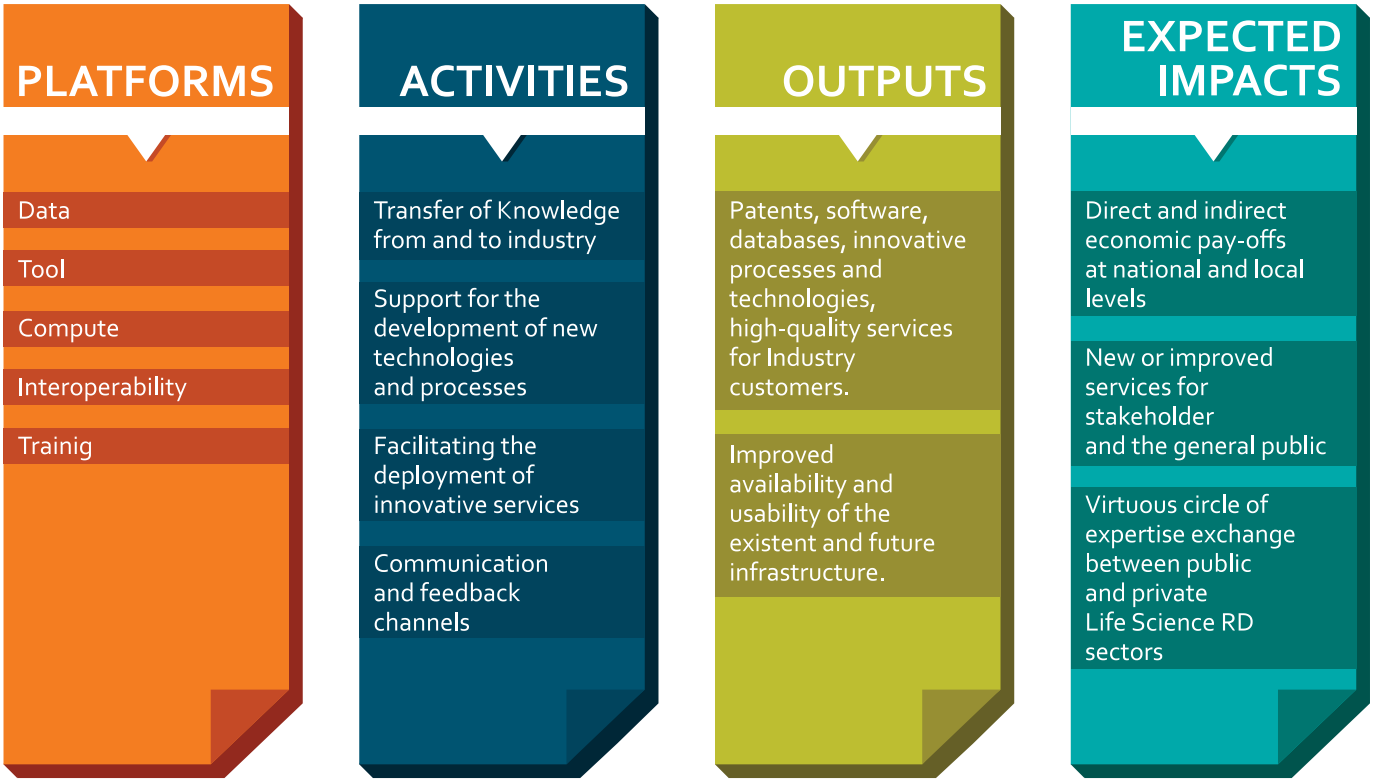
Industry is a major user of Europe's key bioinformatics resources. Industrial users range from large multinational companies to micro-SMEs, covering fields such as pharma, biotech, food, agriculture and aquaculture.

The overall objectives of ELIXIR's industry Strategy are to:

- Increase industry usage of ELIXIR resources and ensure that its name is synonymous with quality.
- Enable open innovation by Europe's SMEs.
- Build effective partnerships with key industry stakeholders and initiatives.
- Ensure effective communication between industry and ELIXIR.
- Support the bioinformatics training needs of industry.

The range of industry sectors that ELIXIR has the potential to support is also broad -from publishers to SMEs to HPC and cloud providers. In Italy, ELIXIR-IT engages an increasingly large share of researchers, expertise and services, while Industries and SMEs are beginning to show more interest in bioinformatics as a means to widen their business and improve their products. ELIXIR-IT aims to intercept this trend and to extend its support beyond public scientific research in the next years: the long-term strategy is gearing up for the delivery of on-demand services targeted to the private sector.

The ELIXIR-IT industry Group is working towards this goal and its first objective is to enhance existing partnerships and to create new ones with industry and SMEs through the establishment of effective communication channels between the private sector and ELIXIR-IT.



ELIXIR Innovation and SME Forum

The Italian Node contributes to the ELIXIR's Innovation and SME Forum, a series of thematic events for companies hosted by ELIXIR Nodes across Europe. They highlight examples of research-intensive companies already making use of big data in their respective fields by presenting the free tools, data resources and services that are available through ELIXIR. The topics of the Innovation and SME events are intertwined with ELIXIR's resources and services across various life-science domains.

ELIXIR-IT
Data Platform

The goal of the **ELIXIR Data Platform** is to drive the use, re-use and value of life sciences data. What it aims for is to provide its users with robust, long-term sustainable data resources within a coordinated, scalable and connected data ecosystem.

Bioinformaticians and life sciences researchers in both academic and industrial settings depend on solid confidence in the sound governance, life cycle management, and long-term sustainability of those data resources.

ELIXIR Core Data Resources are a collection of scientific top-quality resources reckoned to be fundamentally important for the long-term conservation of biological data.

The **MINT** and **MobiDB** databases are part of ELIXIR-IT data resources, which have been identified as a '*Core Data Resource*' by ELIXIR, indicating their fundamental importance to the life-science community. Furthermore, the portfolio of ELIXIR-IT data resources currently includes 17 unique, high-quality, proteomics, interactomics, transcriptomics and genomics databases that satisfy ELIXIR-IT's strict quality management policy for data resources.

The **ELIXIR-IT Data Platform** is committed to working towards a rational organization and integration of **ELIXIR-IT** data resources and to pursue possible collaborations and joint opportunities with related resources in the public or private research sectors.

The Platform provides a periodical evaluation of the uniqueness and competitiveness of our resources in order to drive our strategy within the data landscape.

MiSeq
illumina

ELIXIR-IT
Tools Platform

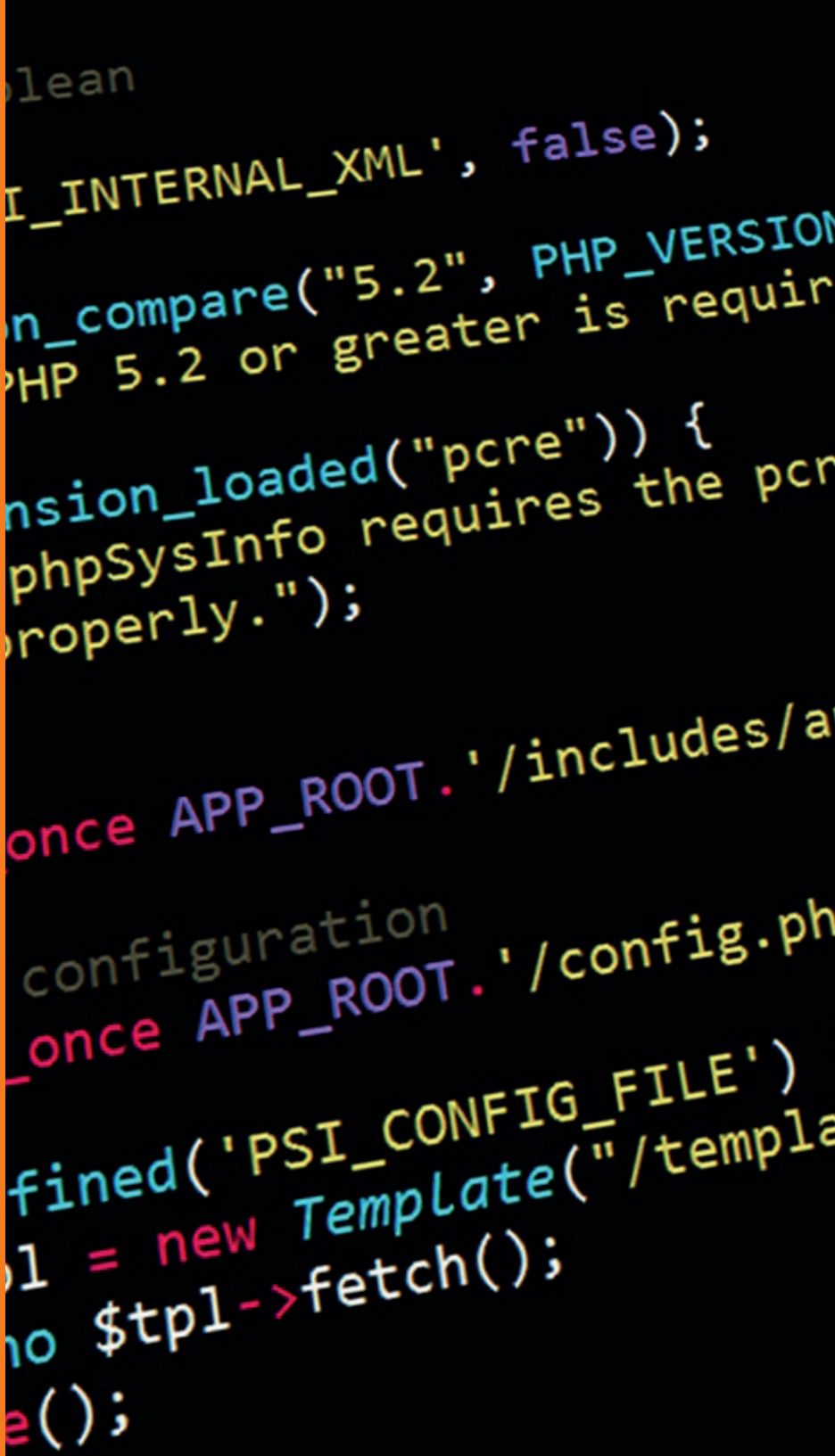
Scientists need software tools to gain greater insights from life sciences data. In order to be used and integrated into robust computational workflows, bioinformatics tools have to be released with a specific license alongside complete and detailed documentation. In this context, the main objective of the ELIXIR Tools Platform is to improve the discoverability, quality and sustainability of bioinformatics software resources.

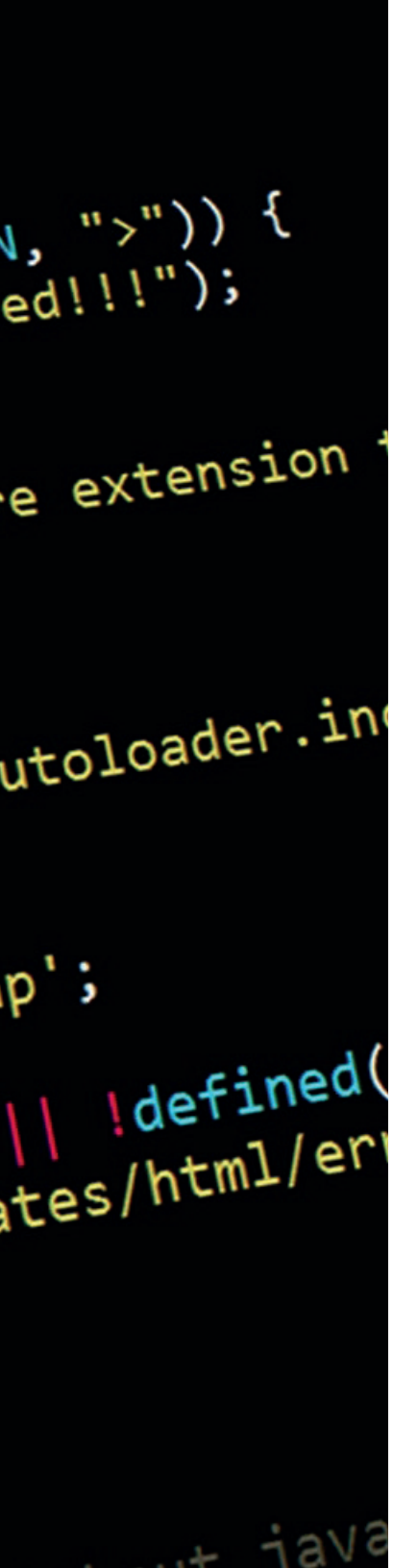
The Italian bioinformatics community has a long-standing and well-established tradition in the development of novel tools for the computational and statistical analysis of different types of biological data, several of which are seen as state of the art in their respective domain of application.

The ELIXIR-IT Tools Platform currently includes 27 "core tools", selected on key parameters like their user base, the novelty and uniqueness of the approach, their development expectations and their current/expected "market share". These tools cover virtually every domain of modern bioinformatics research:

- Protein analysis, modeling and annotation
- Genomic Sequencing, Assembly and Annotation
- Transcriptomics and Epigenomics
- Integrative omics analysis
- RNA bioinformatics

ELIXIR-IT
Interoperability Platform





The situation today is that life-science research frequently produces data in a wide variety of formats. This diversity of formats means that scientists cannot easily find and compare datasets from different sources, so their ability to make new discoveries is hampered. The goal of **the ELIXIR Interoperability Platform** is to help people and machines to discover, access, integrate and analyse biological data. It encourages the life sciences community to adopt standardized file formats, metadata, vocabularies and identifiers. The final goal is to make bioinformatic data and software tools Findable, Accessible, Interoperable and Reusable (FAIR).

ELIXIR-IT contributed to define data formats and standards in different fields such as the protein-protein interaction one, in collaboration with the Human Proteome Organization and the International Molecular Exchange Consortium. Moreover, it defined the *de facto* standard ontology for Intrinsically Disordered Proteins and the ontology for Nutritional Studies. ELIXIR-IT contributes to the BioSchema specification, a standard schema for life sciences data allowing dataset and tools to be indexed by internet search engines like Google or Bing and easily findable worldwide.

ELIXIR-IT is committed to promoting and encouraging the adoption of Interoperability best practices by all the public and private subjects operating in life sciences and offers support and training for accessing the ELIXIR Interoperability services.

Electropherogram - 2A_3
Assay Properties
Chip Summary
Gel
Electropherogram
Result Flagging
Log Book
Size [bp]
Conc. [ng/µl]
Molarity [nmol/l]
Observations
424.2 Lower Marker
2.1 Upper Marker

ELIXIR-IT
Training Platform

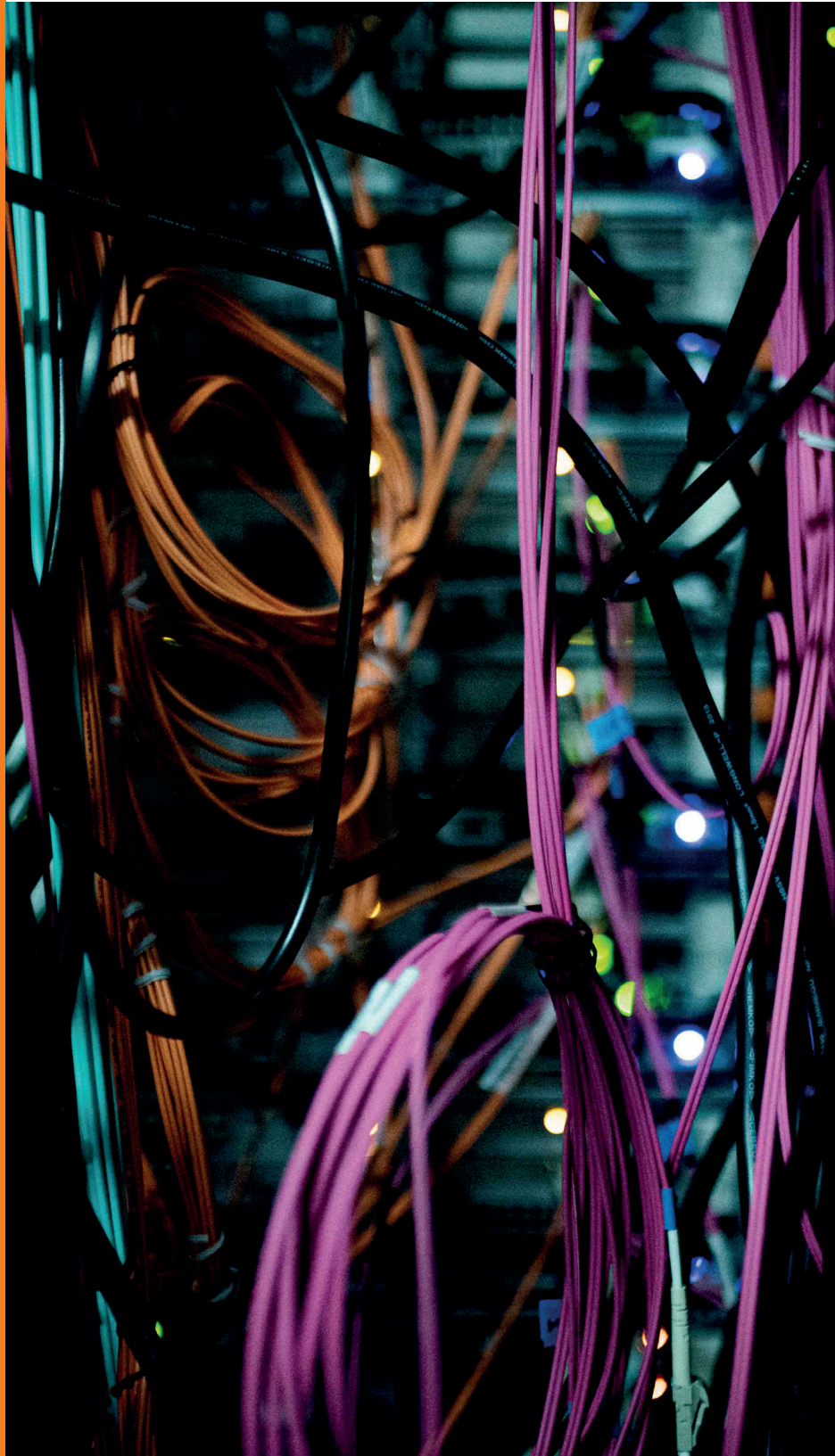
The **ELIXIR Training Platform (TP)** was established to develop a training community spanning all ELIXIR Nodes. It aims to strengthen national training programmes, grow bioinformatics capacity and competence across Europe, and empower researchers to use ELIXIR's services and tools.

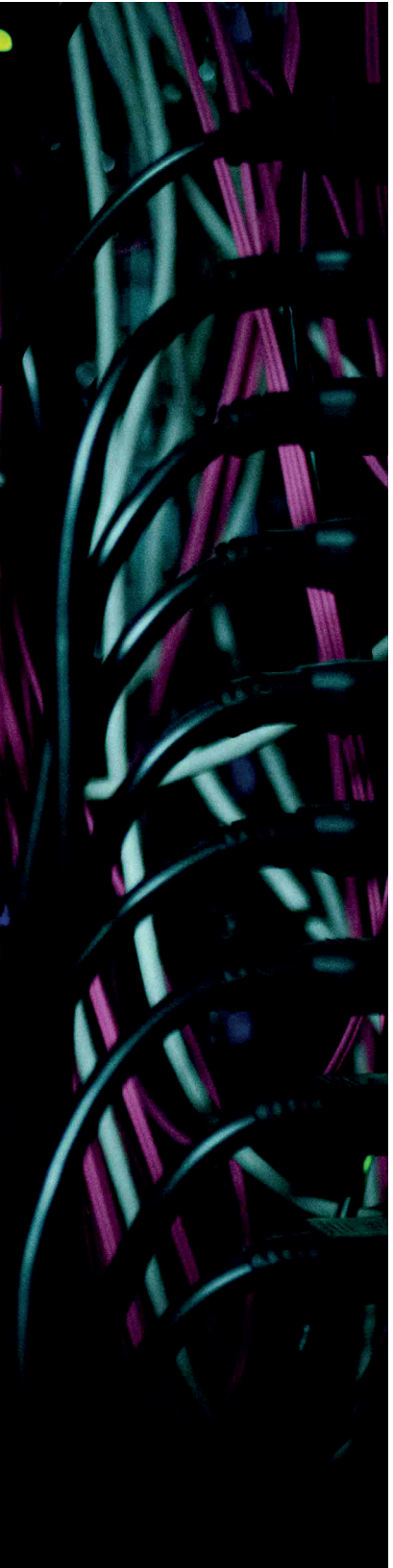
The **ELIXIR-IT TP** supports researchers and professionals in acquiring specialised bioinformatics, computational and data management skills. A community of learners and instructors that actively embrace the idea that quality education is essential to achieve excellence in life sciences research has been setup around the Platform. The ELIXIR-IT TP closely collaborates with other ELIXIR Nodes in Europe on training initiatives and develops and promotes a large variety of national and international training and training-related events, projects and activities:

- Workshops to build capacity in computational and data management skills
- Bioinformatics courses for life sciences researchers
- Targeted initiatives for people conducting and supporting research
- "Train the Trainer" events
- Development of training material
- ELearning activities
- Dissemination of best practices for software development
- Quality and Impacts of training surveys
- Bring Your Own Data events.

The **ELIXIR-IT TP** is also committed to making its courses available as widely as possible. To address stakeholder needs, and as an outcome of a recently approved infrastructural project, live training courses will be converted into eLearning events and made publicly available.

ELIXIR-IT
Compute Platform





Today, thousands of science laboratories across the world generate massive amounts of data that they subsequently make available to collaborators directly or place in public archives for open access. Hence, the traditional method of a researcher downloading and analyzing data locally is no longer viable due to both the data size and scope of the analysis activities. The data needs to be managed as a federation, where data providers work as a single infrastructure providing mechanisms where researchers can bring their analysis to where the data is located.

The ELIXIR Compute Platform infrastructure will allow life scientists to easily access, share and analyse data from different sources across Europe: the main objective is to combine all components of the ELIXIR Compute services into a seamless workflow through the ELIXIR Authorisation and Authentication infrastructure.

The ELIXIR-IT Compute Platform delivers services through the collaboration of its infrastructural partner institutes. CINECA, INFN and ENEA provide state-of-the-art High-throughput (HTC) and High-Performance Computing (HPC) solutions made readily available to life sciences researchers through convenient interfaces and access programs. GARR runs the high-performance and low latency network for research activities that connects to internet with very high bandwidths in bidirectional mode.

CRS4 provides computing and storage facilities integrated with sequencing machines. Finally, the development and deployment of Cloud services for bioinformatics are supported by INFN and GARR data centres.

ELIXIR-IT
Team

HEAD OF NODE

Graziano Pesole (CNR)

DEPUTY HEAD OF NODE

Silvio Tosatto (UNIPD)

ELIXIR - IT BOARD MEMBER

Rita Casadio (UNIBO)

MINISTRY OF UNIVERSITY AND RESEARCH (MIUR) REFERENTS

Gianluigi Consoli
Salvatore La Rosa
Grazia Pavoncello

SCIENTIFIC ADVISORY BOARD

Roderic Guigo (CRG, BARCELONA)
Rodrigo Lopez (EMBL-EBI)
John Moult (IBBR, UNIVERSITY OF MARYLAND)
Christine Orengo (UNIVERSITY COLLEGE, LONDON)
Pedrag Radivojac (NORTHEASTERN UNIVERSITY)
Ben Raphael (PRINCETON UNIVERSITY)

GENERAL ASSEMBLY

Marcella Attimonelli (UNIBA)
Sanzio Bassini (CINECA)
Raffaele Calogero (UNITO)
Rita Casadio (UNIBO)
Duccio Cavalieri (UNIFI)
Gianni Cesareni (UNIROMA2)
Alessandro Cestaro (FMACH)
Giovanni Chillemi (UNITU)
Maria Luisa Chiusano (STAZOO)
Angelo Facchiano (CNR)
Luigi Frusciante (UNINA)
Giorgio Maggi (INFN)
Giancarlo Mauri (UNIMIB)
Silvio Migliori (ENEA)
Luciano Milanesi (CNR)
Stefano Morabito (ISS)
Francesco Musacchia (TIGEM)
Stefano Pascarella (UNIROMA)
Giulio Pavesi (UNIMI)
Riccardo Percudani (UNIPR)
Graziano Pesole (CNR)
Federico Ruggieri (GARR)
Roberto Tagliaferri (UNISA)
Silvio Tosatto (UNIPD)
Gianluigi Zanetti (CRS4)

EXECUTIVE COMMITTEE

Sanzio Bassini (CINECA)
Rita Casadio (UNIBO)
Alessandro Cestaro (FMACH)
Angelo Facchiano (CNR)
Giancarlo Mauri (UNIMB)
Graziano Pesole (CNR)
Silvio Tosatto (UNIPD)

ADMIN OFFICER

Laura Marra (CNR)

TECHNICAL COORDINATOR AND DEPUTY

Federico Zambelli (UNIMI)
Giacomo Tartari (UNIBO)

TRAINING PLATFORM LEADER AND DEPUTY

Allegra Via (CNR)
Loredana Le Pera (CNR)

INDUSTRY GROUP LEADER AND DEPUTY

Angelo Facchiano (CNR)
Francesca De Leo (CNR)

PROJECT MANAGER

Giacomo Tartari (UNIBO)

WEB MANAGER

Ivan Mičetić (UNIPD)

LOCAL TECHNICAL COORDINATORS

Marco Beccuti (UNITO)
Raffaele Calogero (UNITO)
Davide Carnevali (UNIPR)
Claudio Carta (ISS)
Tiziana Castrignanò (CINECA)
Andrea Cattani (FMACH)
Massimo Celino (ENEA)
Matteo Chiara (UNIMI)
Maria Luisa Chiusano (STAZOO)
Sergio Cocozza (UNINA)
Gianmauro Cuccuru (CRS4)
Nunzio D'Agostino (UNINA)
Gianluca Della Vedova (UNIMIB)
Giacinto Donvito (INFN)
Angelo Facchiano (CNR)
Carlo Ferrari (UNIPD)
Marco Fondi (UNIFI)
Fulvio Galeazzi (GARR)
Arnold Knijn (ISS)
Luana Licata (UNIROMA2)
Flavio Licciulli (CNR)
Anna Marabotti (UNISA)
Ivan Mičetić (UNIPD)
Luciano Milanesi (CNR)
Marco Moretto (FMACH)
Francesco Musacchia (TIGEM)
Vincenzo Nigro (TIGEM)
Luca Parca (UNIROMA2)
Gaetano Perrotta (ENEA)
Ernesto Picardi (UNIBA)
Luca Pireddu (CRS4)
Roberto Prieste (UNIBA)
Matteo Ramazzotti (UNIFI)
David Sasah Staid (UNIROMA)
Bernardina Scafuri (CNR)
Giacomo Tartari (UNIBO)
Sabrina Tommasini (GARR)

ELIXIR-IT
Platforms

DATA PLATFORM LEADER AND DEPUTIES

Silvio Tosatto (UNIPD)
Luana Licata (UNIROMA2)
Tiziana Castrignanò (CINECA)

TRAINING PLATFORM LEADER AND DEPUTY

Allegra Via (CNR)
Loredana Le Pera (CNR)

COMPUTE PLATFORM LEADER AND DEPUTIES

Giacinto Donvito (INFN)
Giovanni Chillemi (UNITU)
Marco Tangaro (CNR)

INTEROPERABILITY PLATFORM LEADER AND DEPUTIES

Pier Luigi Martelli (UNIBO)
Giacomo Tartarti (UNIBO)
Matteo Chiara (UNIMI)

TOOLS PLATFORM LEADER AND DEPUTY

Giulio Pavesi (UNIMI)
Ernesto Picardi (UNIBA)

ELIXIR-IT
User Communities

HUMAN DATA COMMUNITY LEADERS

David Horner (UNIMI)
Matteo Chiara (UNIMI)

RARE DISEASE COMMUNITY LEADER

Rita Casadio (UNIBO)

MARINE METAGENOMICS COMMUNITY LEADER

Monica Santamaria (CNR)

PLANT SCIENCES COMMUNITY LEADERS

Alessandro Cestaro (FMACH)
Giorgio Valle (UNIPD)

GALAXY COMMUNITY LEADERS

Federico Zambelli (UNIMI)
Marco Antonio Tangaro (CNR)

STRUCTURAL BIOINFORMATICS COMMUNITY LEADER

Manuela Helmer Citterich (UNIROMA2)

METABOLOMICS COMMUNITY LEADER

Gianluigi Zanetti (CRS4)

DISORDERED PROTEINS COMMUNITY LEADERS

Silvio Tosatto (UNIPD)
Damiano Piovesan (UNIPD)

CREDITS

Editing:

Francesca De Leo
Graziano Pesole
Federico Zambelli

Contacts:

elixir-italy.org

g.pesole@ibiom.cnr.it

tinyurl.com/elixir-it-ml

elixir_it

Graphic design:

Laboratoriocom.it

CONSIGLIO NAZIONALE DI RICERCHE
ISTITUTO DI BIOMEMBRANE, BIOENERGETICA
E BIOTECNOLOGIE MOLECOLARI

Via Amendola 122 D/O - 70126 Bari
Email: G.pesole@Ibiom.cnr.it