ELIXIR : the distributed infrastructure for life-science information

Anne-Françoise Adam-Blondon

ELIXIR: the distributed infrastructure for life-science information

A-F Adam-Blondon, INRAE, ELIXIR-FR Head of node

@ELIXIREurope

/company/elixir-europe

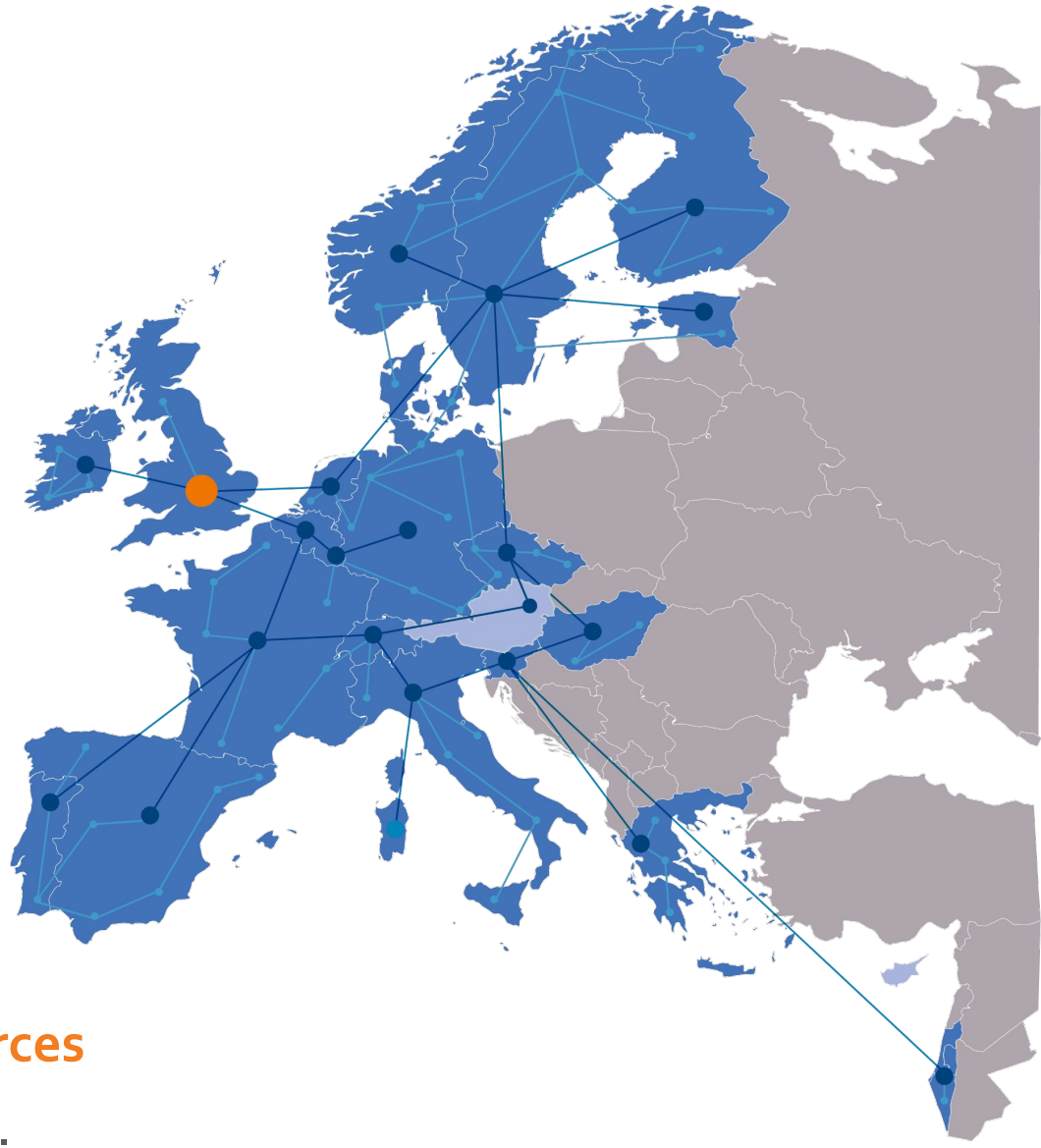
LERU Open Science Ambassadors meeting,
Université Paris-Saclay, France, 2023 Nov 9th

www.elixir-europe.org

ELIXIR Europe

ELIXIR is an intergovernmental organisation that brings together **life science** resources such as

- databases
- software tools
- training resources
- interoperability resources
- compute resources
- data management support



The goal of ELIXIR is to **coordinate bioinformatics resources from across Europe so they form a single infrastructure**.

ELIXIR – who we are

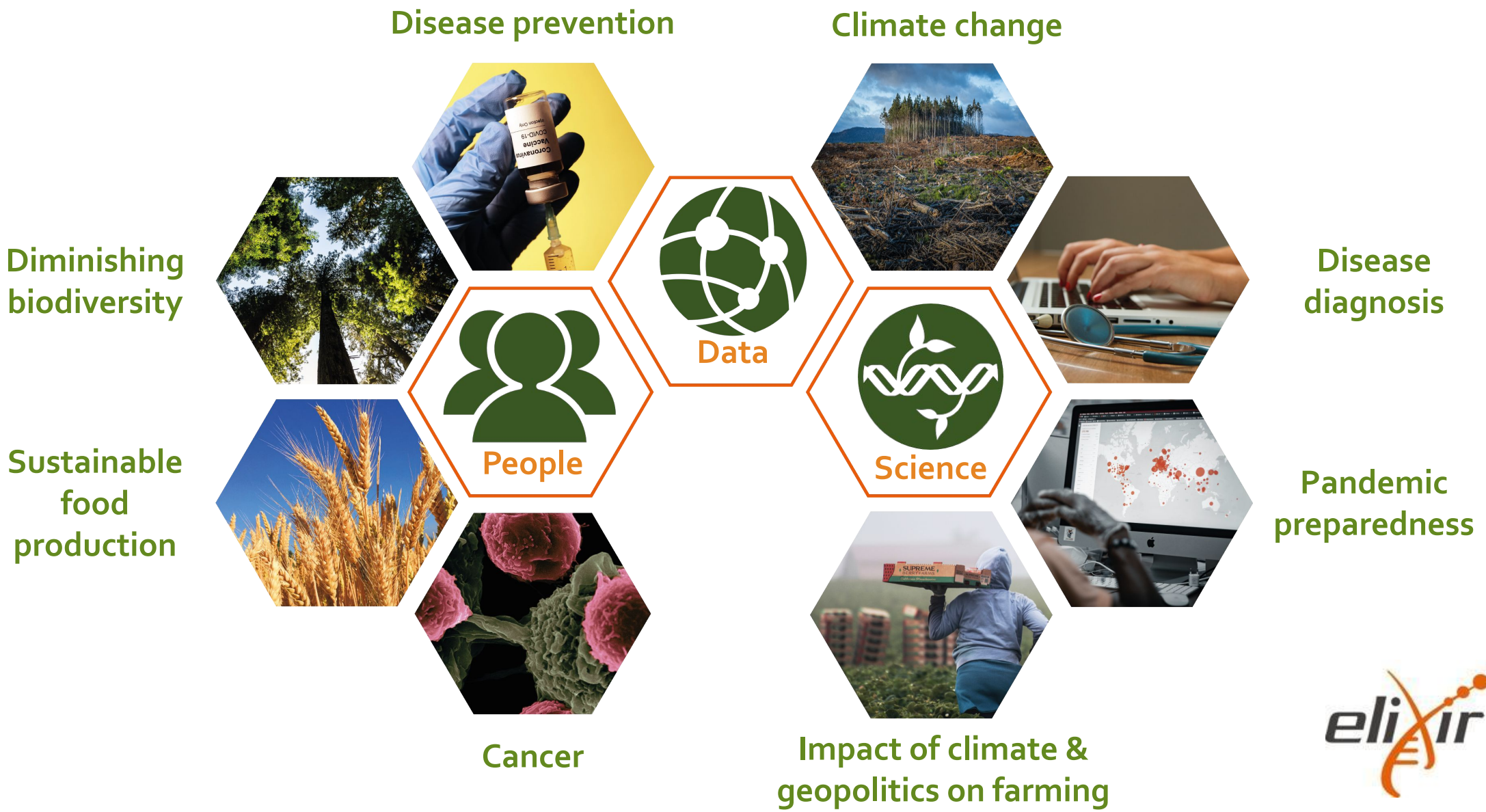
24 Nodes

- ELIXIR Nodes form national networks of excellence, containing a total of **245** institutes/universities
- **EMBL-EBI** is one of the nodes

1 Hub

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

Enable data-driven solutions to societal challenges



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

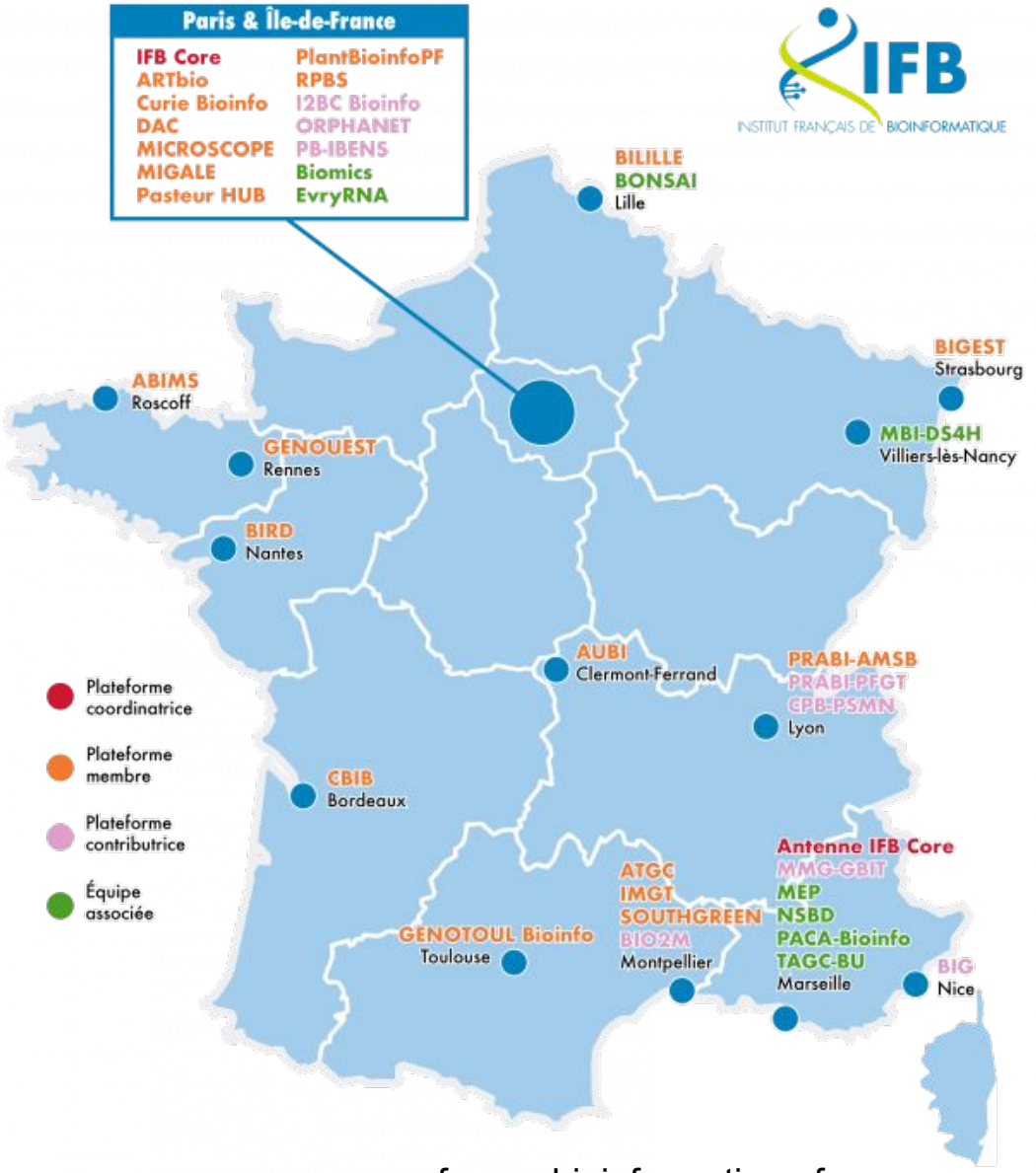
- Multi-research organizations : CNRS, Inserm, INRAE, CEA

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)



www.france-bioinformatique.fr

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

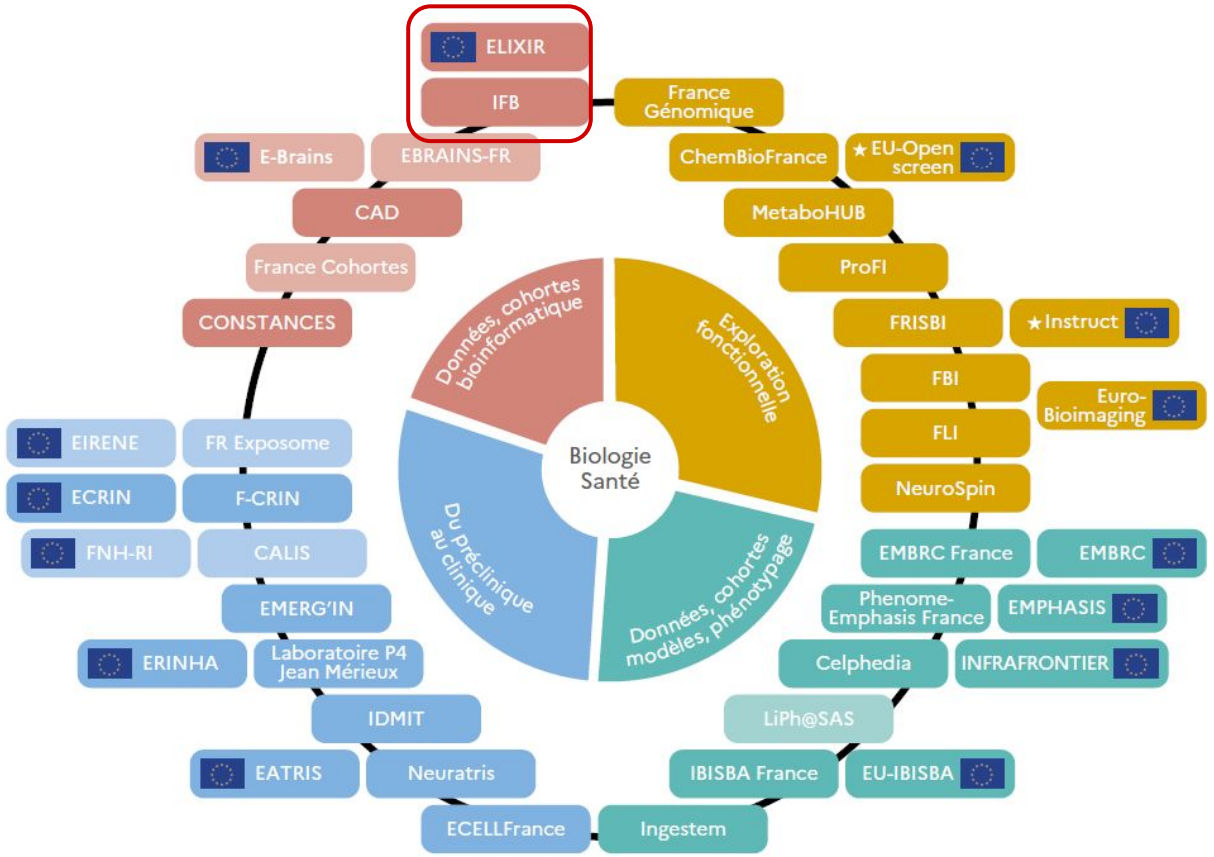
Mobilise and integrate molecular data to support transnational research programmes in biodiversity, food security and pathogens

The promise of genomics research to improve health and disease outcomes

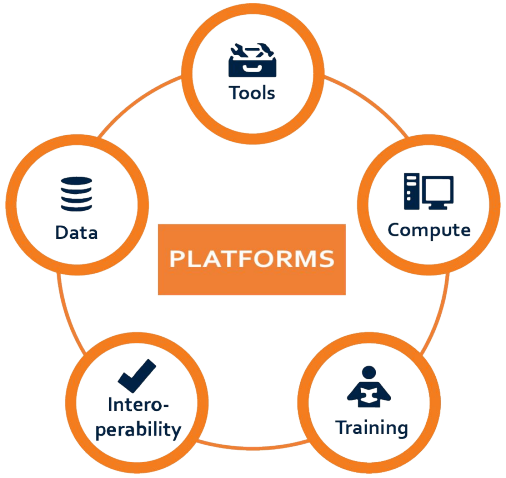
Provide the infrastructure to support the discovery, access, sharing and analysis of human genomics data and linked data on a massive scale

IFB is part of the French network of national Research Infrastructures in Life Sciences

- **Sharing expertise and working groups on subjects of common interest**
- **Partnership in large national projects**
- **Enhance partnerships in ELIXIR activities and in European projects**



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

LS LOGIN

Data

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

elixir Core Data Resource

Tools

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

elixir bio.tools WorkflowHub OPENEBENCH Galaxy

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**

elixir Recommended Interoperability Resources

Training

Strengthens **national training** programmes; grows bioinformatics and research data management **training capacity and competence across Europe**; empowers researchers to use ELIXIR's resources

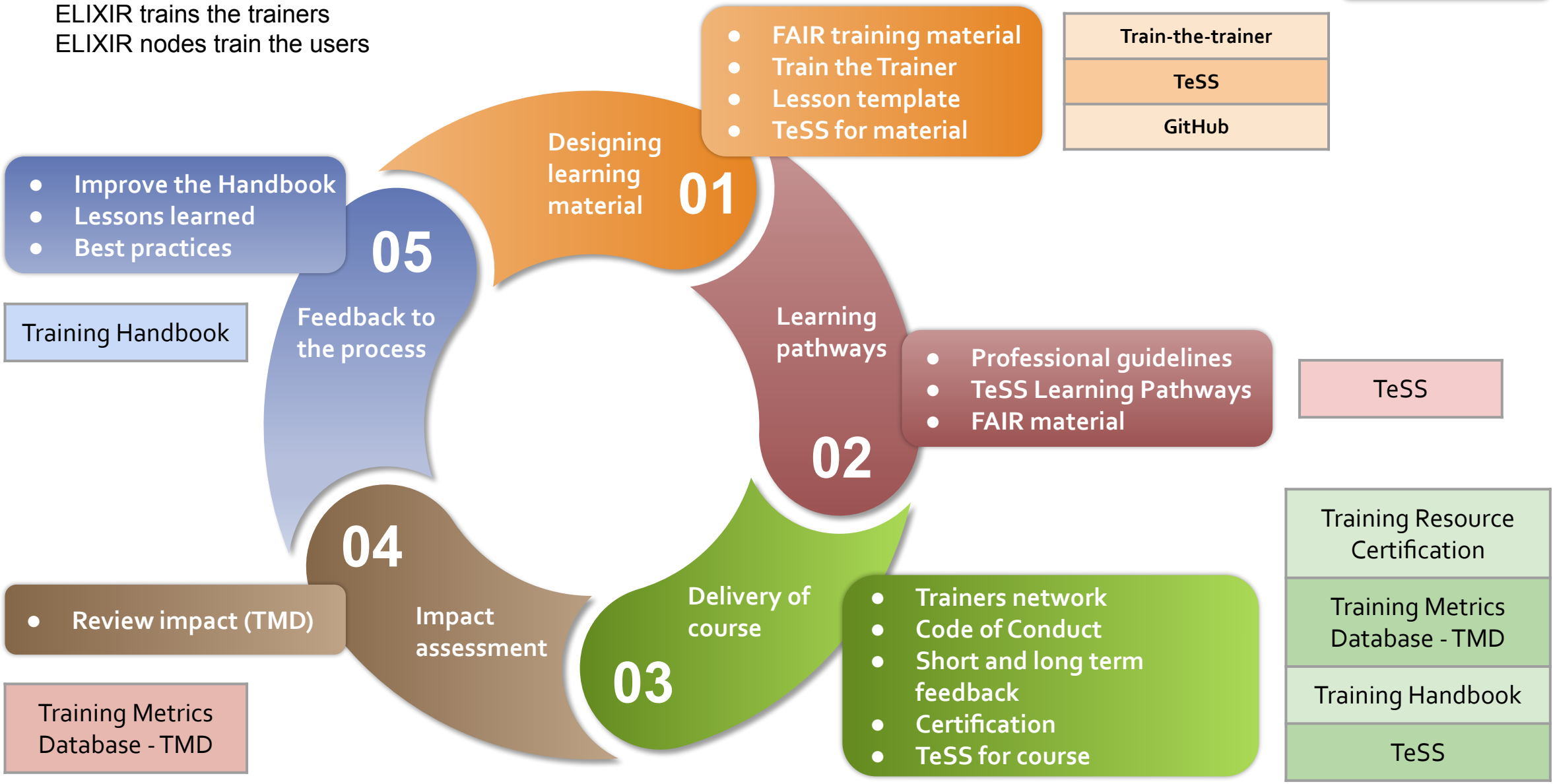
elixir TeSS

ELIXIR Training Platform

ELIXIR trains the trainers
ELIXIR nodes train the users

Service/Product

- Activity



How we work - 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-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 =>
- Contributing to European projects: 9 since 2020

Science

Communities

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

Human Data Communities

- Federated Human Data
- Human Copy Nu Variation
- Rare diseases
- Cancer Data (focus group)

Technology

Platforms

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

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.

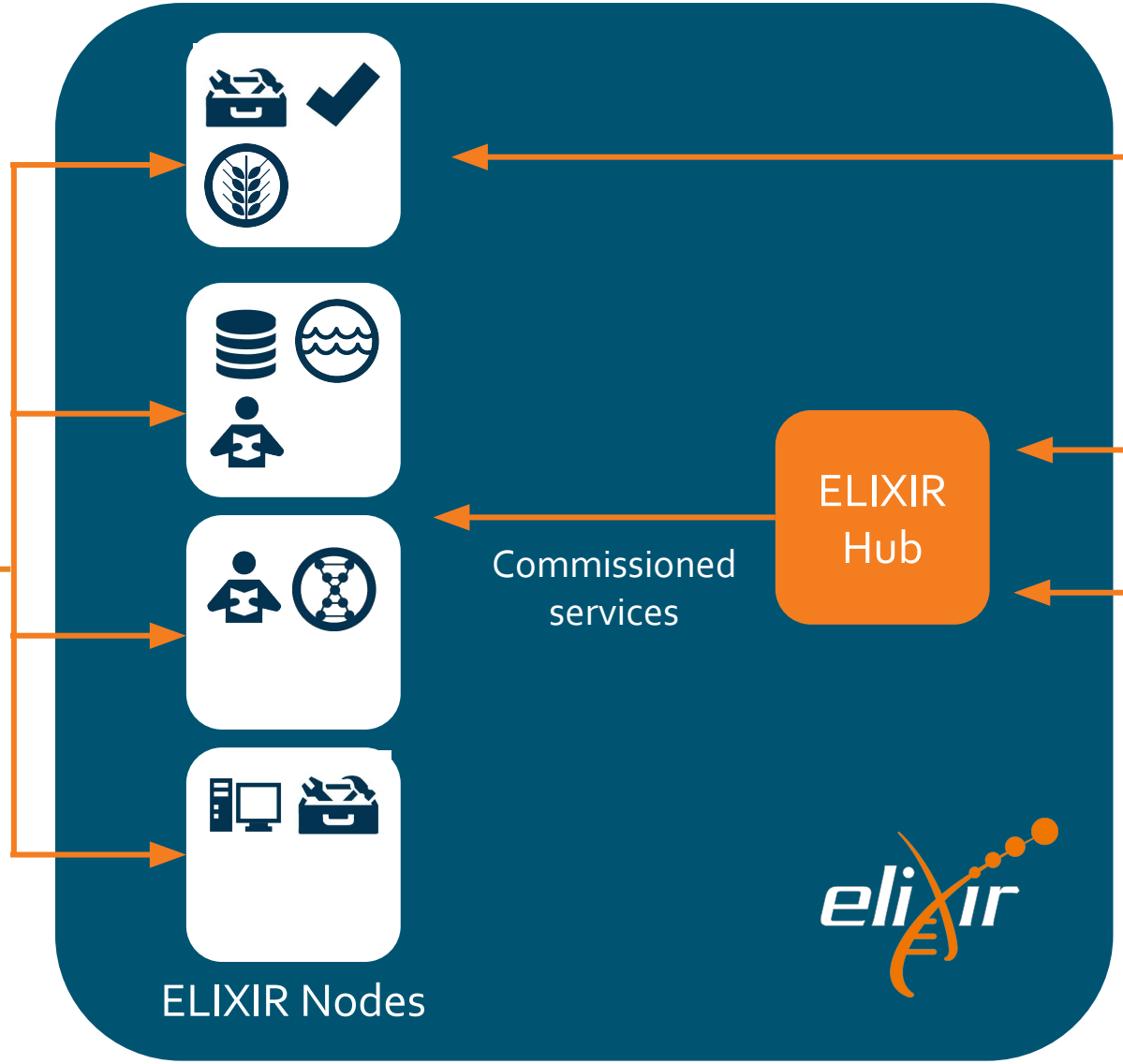
Interoperability Platform

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.

ELIXIR's funding model



Node funding:

- National roadmap funding
- Competitive research grants
- EU Structural Funds or Recovery and Resilience Facility (RRF)
- Trusts and foundations
- Industry collaboration

https://elixir-europe.org/about-us/how-funded/sustainability-plan

Commissioned services

ELIXIR Hub

External funding:

(EC funded projects: ELIXIR represented by ELIXIR Hub or by an ELIXIR Node)

- Horizon Europe
- Digital Europe
- EU4Health

Internal funding:

Annual contribution by member states

ELIXIR's key stakeholders

ELIXIR Nodes

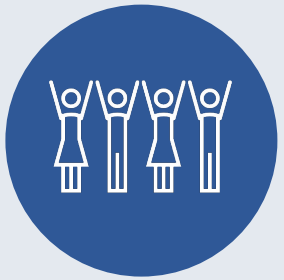
24 ELIXIR Nodes
245+ institutes
850+ scientists

Users

Bioinformaticians
500,000+ life science researchers
Users in industry

Funders & decision makers

European Union
National funding agencies
ESFRI delegates
ELIXIR Board members



Collaborators

ESFRI RIs
Global initiatives (GA4GH, Galaxy)
National initiatives (Australian BioCommons, NIH)

Collaborator example – the ELIXIR EOSC strategic vision

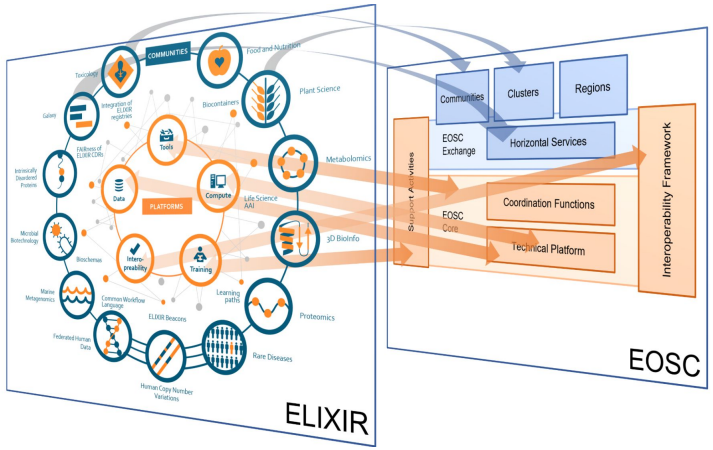
ELIXIR maps to EOSC at all levels

- Throughout the EOSC Core and Interoperability frameworks
- Bidirectional relationships, with best of breed solutions being adopted by both sides
- ELIXIR Communities will over time adopt EOSC Services and become driving user communities for both organisations

ELIXIR and the EOSC Association Task Forces

ELIXIR on Advisory Groups

- ✔ The Implementation of EOSC
- ✔ Metadata and Data Quality
- ✔ Research Careers and Curricula
- ✔ Sustaining EOSC
- ✔ Technical Challenges on EOS



ELIXIR involvement in EOSC projects

EOSC-Pilot

EOSC-Enhance

EOSC-Life

EOSC-Future

ELIXIR flagship events

BioHackathon Europe

- Bringing together bioinformaticians for **five days of hacking**
- **Five successful annual events** since 2018
- Projects to advance **open source infrastructure** for data integration to accelerate scientific innovation
- Supporting operations across ELIXIR Platforms, Communities and Focus Groups through technology implementations (e.g. **FAIR, identifiers, metadata standards, ontologies and metadata catalogues**)

Innovation and SME Forums

- Approximately **two one-day** events per year
- Providing **Small to Medium Sized Enterprises** (SMEs) the opportunity to present their innovative ideas
- Enabling companies to learn more about **current and emerging ELIXIR services**
- Forging strong links with the **local ELIXIR Node representatives** running ELIXIR services

ELIXIR Bioinformatics Industry Forum (EBIF)

- **One-day** annual event
- Discussions with industry experts around **visionary ideas, bottlenecks and solutions** to major challenges in the data-driven life science sector
- A forum for **knowledge exchange and collaboration** in the pre-competitive space
- **Networking** opportunities with bioinformatics opinion leaders, academic experts in ELIXIR and the commercial sector

ELIXIR's Code of Conduct – the principles

The Code of Conduct pertains to ELIXIR organised or funded events

We value
each other's perspectives

We adopt
a zero-tolerance approach to harassment and discrimination

We maintain
high ethical standards

We'll apply
honesty and integrity in the dealing of any transgressions against the Code.

We're committed
to making ELIXIR events collaborative, supportive and enjoyable

We'll ensure
a respectful and inclusive environment

What we offer

Guidelines

Guidelines and best practices to help you manage life science data, run training courses, develop software and more.

Web portals

Find the right software, training courses, standards and more in our interlinked portals to life science resources.

Services

Find compute services, databases, and the full list of resources ELIXIR coordinates.

Partnerships with Industry and SMEs

Join events and projects that bring the private and public sectors together.

Opportunities to work together

Join a scientific group in ELIXIR or partner with us to apply for EC funding.

For ELIXIR members

If you work at an institute that is part of ELIXIR, then remember to take advantage of the benefits ELIXIR offers.

ELIXIR coordinates the provision of life science services and resources, developed and managed by ELIXIR Nodes

The services and resources are freely available to researchers around the world

The services and resources support efficient manipulation, analysis, storage and exchange of life science data

https://elixir-europe.org/what-we-offer

Examples of ELIXIR high impact services

LS LOGIN	**Life Science Login**	Authentication service	13K	Users
FAIRsharing.org	**FAIRsharing**	Data and metadata resource including standards, databases and policies	1.5K	Standards
elixir TeSS	**ELIXIR TeSS**	ELIXIR's training portal	1.9K	Training materials
</>	**BioContainers**	Software standardisation resource	70K	Containers & packages
elixir bio.tools	**BioTools**	Registry of tools, databases and services	35K	Monthly visits

Metrics from 2021

elixir

ELIXIR research data management resources

Guidance for data stewards, project managers and researchers

Overview of good data management practices

RDMkit

The Research Data Management Kit (RDMkit) guides you through the whole data management life cycle and includes advice specific to your domain, your role and your country.

Step-by-step instructions

FAIR cookbook

The FAIR Cookbook contains step-by-step recipes to accomplish specific data management tasks and to make your data FAIR (Findable, Accessible, Interoperable, Reusable).

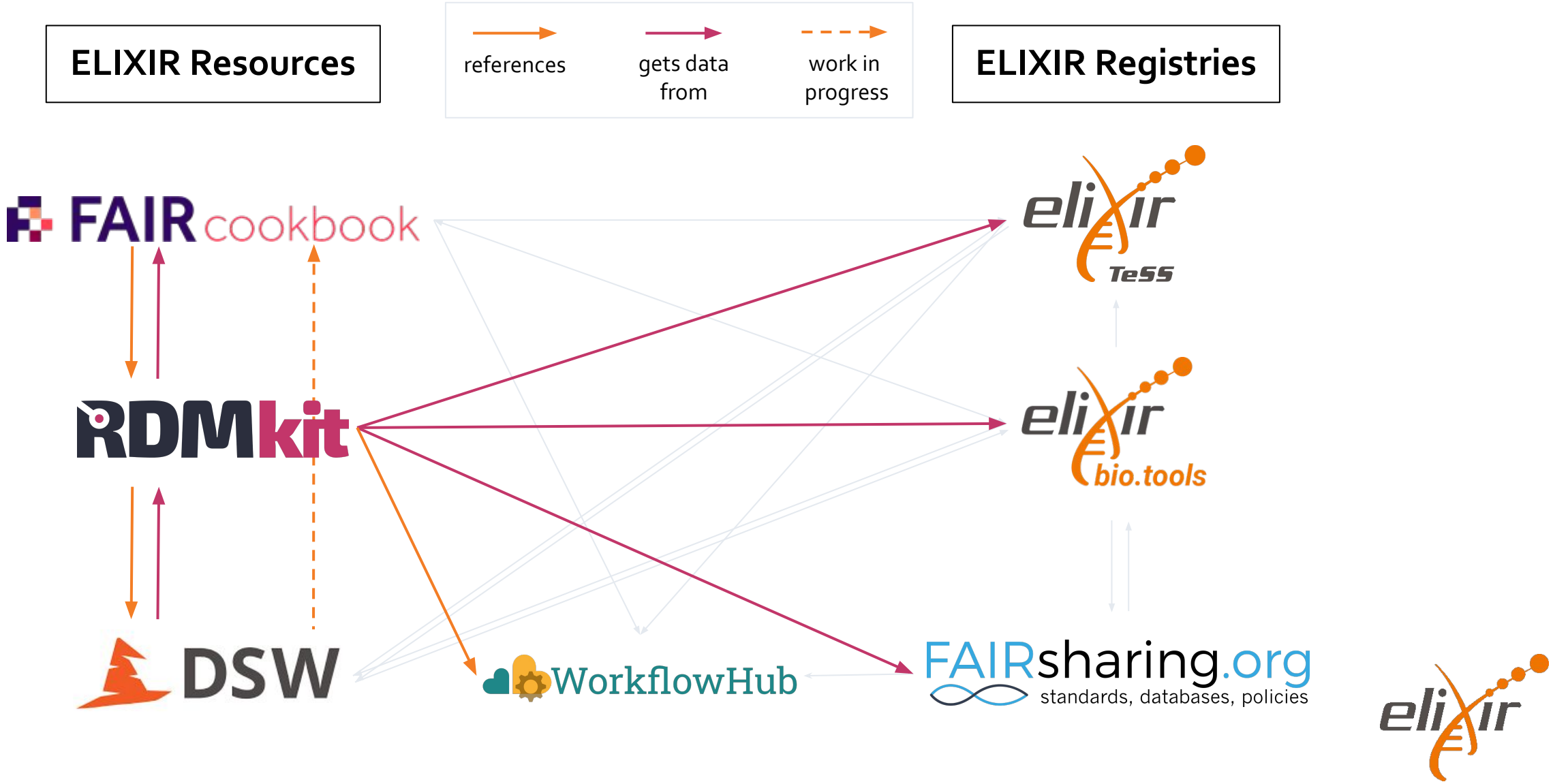
Data management plan wizard

DSW

The Data Stewardship Wizard (DSW) is an online tool that guides researchers and data stewards through their data management planning.

elixir

Research Data Management Toolkit



RDMkit

https://rdmkit.elixir-europe.org

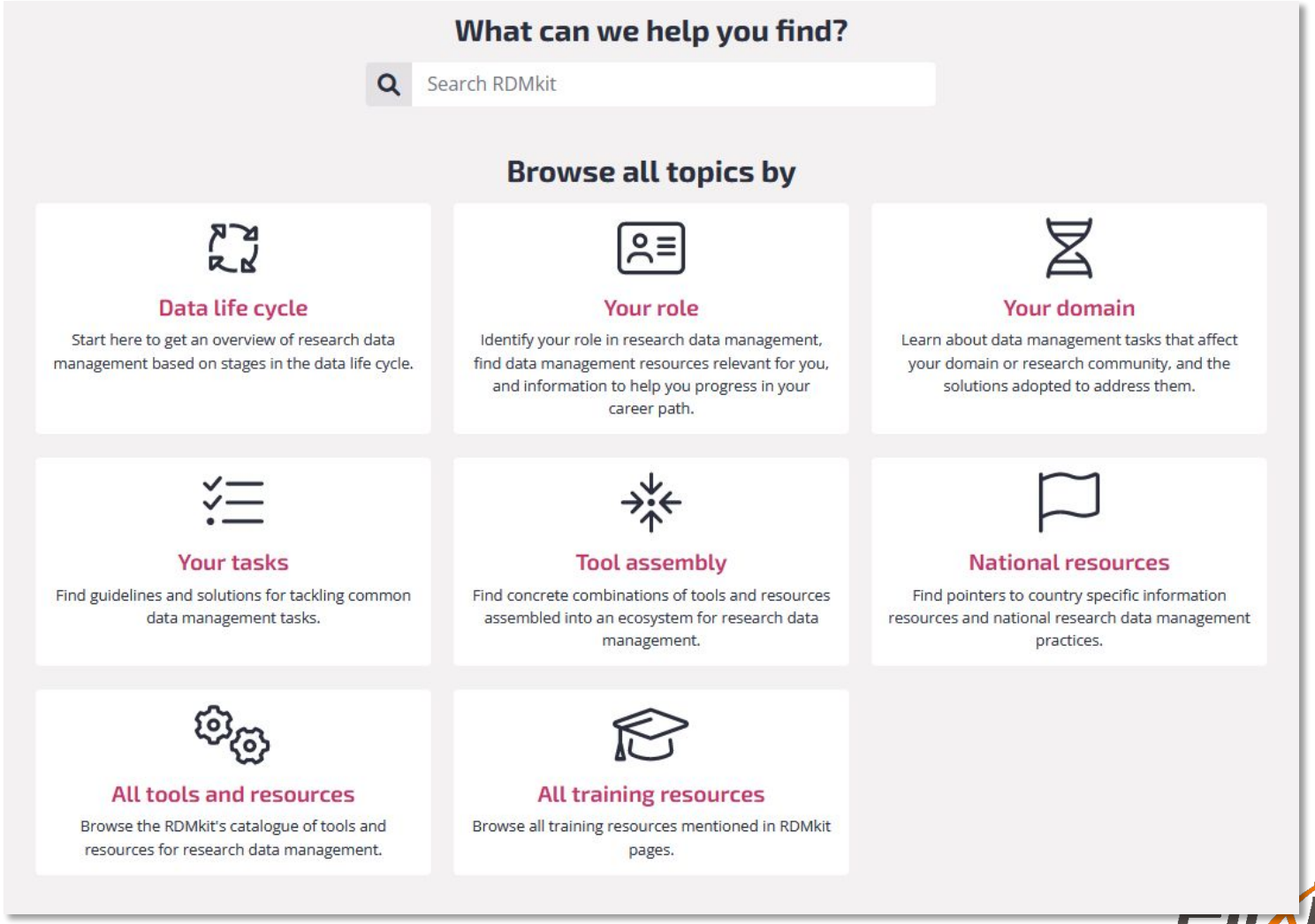
A web-based toolkit for the bioscience community written by the bioscience community

RDMkit in numbers

151
Contributors
The force behind RDMkit

336
Tools & resources
Explained in the context of real world problems

101
Pages
Helping you with data management



Resources for reproducible environment of analysis

bio.tools

bio.tools helps you find and select bionformatics software and connect it in workflows.

BioContainers

Search a repository of containerised software that you can build into workflows.

WorkflowHub

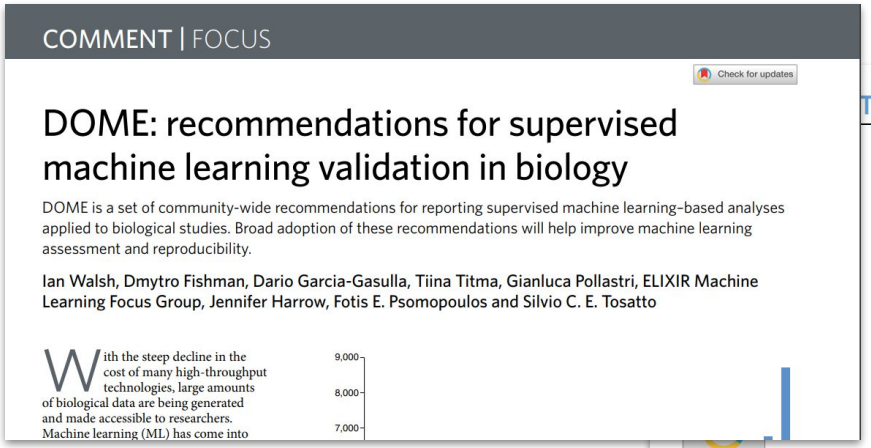
A registry for sharing and publishing scientific computational workflows.

FAIRsharing.org

FAIRsharing.org allows you to search for databases and data policies by aspects such as domain, species and country.

TeSS

Search for training courses, webinars, training materials and workflows in TeSS, ELIXIR's training portal.


COMMENT | FOCUS

Check for updates

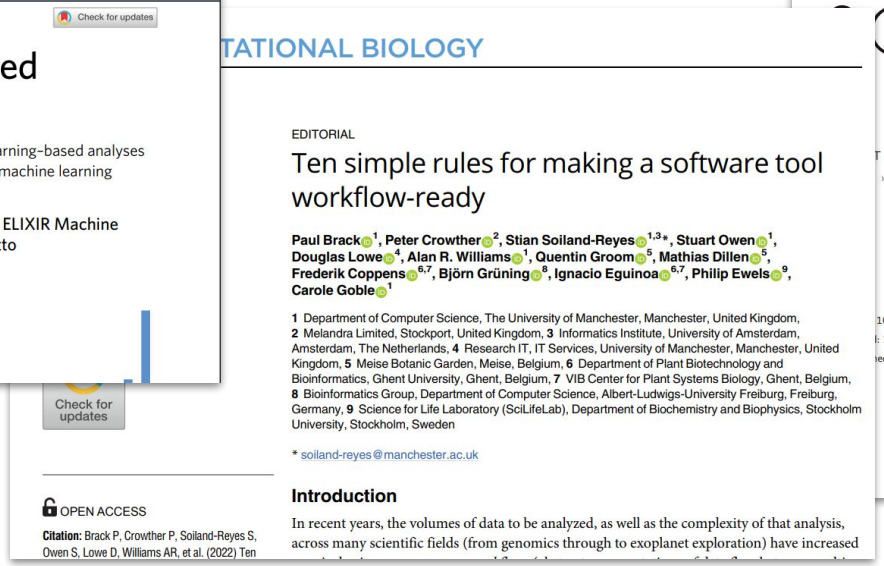
DOME: recommendations for supervised machine learning validation in biology

DOME is a set of community-wide recommendations for reporting supervised machine learning–based analyses applied to biological studies. Broad adoption of these recommendations will help improve machine learning assessment and reproducibility.

Ian Walsh, Dmytro Fishman, Dario Garcia-Gasulla, Tiina Titma, Gianluca Pollastri, ELIXIR Machine Learning Focus Group, Jennifer Harrow, Fotis E. Psomopoulos and Silvio C. E. Tosatto

With the steep decline in the cost of many high-throughput technologies, large amounts of biological data are being generated and made accessible to researchers. Machine learning (ML) has come into

9,000
8,000
7,000


TATIONAL BIOLOGY

EDITORIAL

Ten simple rules for making a software tool workflow-ready

Paul Brack[1], Peter Crowther[2], Stian Soiland-Reyes[1,3]*, Stuart Owen[1], Douglas Lowe[4], Alan R. Williams[1], Quentin Groom[5], Mathias Dillen[5], Frederik Coppens[6,7], Björn Grüning[8], Ignacio Eguinoa[6,7], Philip Ewels[9], Carole Goble[1]

1 Department of Computer Science, The University of Manchester, Manchester, United Kingdom, 2 Melandra Limited, Stockport, United Kingdom, 3 Informatics Institute, University of Amsterdam, Amsterdam, The Netherlands, 4 Research IT, IT Services, University of Manchester, Manchester, United Kingdom, 5 Meise Botanic Garden, Meise, Belgium, 6 Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent, Belgium, 7 VIB Center for Plant Systems Biology, Ghent, Belgium, 8 Bioinformatics Group, Department of Computer Science, Albert-Ludwigs-University Freiburg, Freiburg, Germany, 9 Science for Life Laboratory (SciLifeLab), Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden

* soiland-reyes@manchester.ac.uk

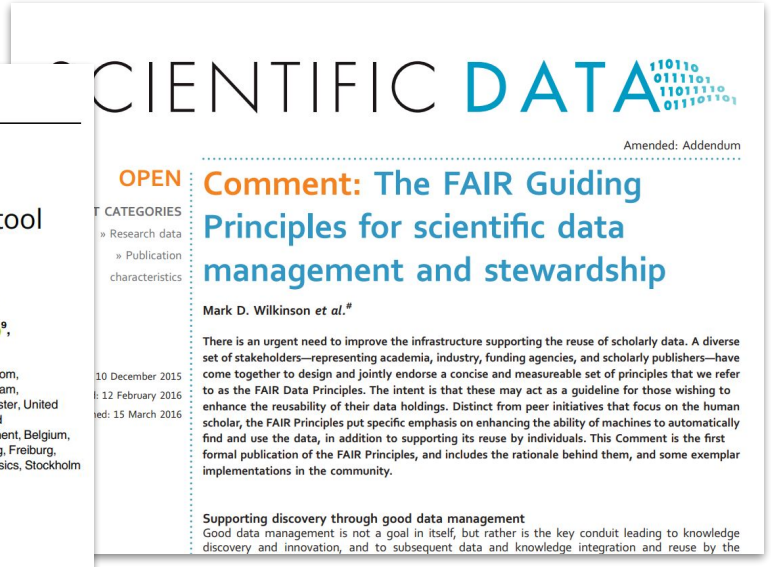
Check for updates

OPEN ACCESS

Citation: Brack P, Crowther P, Soiland-Reyes S, Owen S, Lowe D, Williams AR, et al. (2022) Ten

Introduction

In recent years, the volumes of data to be analyzed, as well as the complexity of that analysis, across many scientific fields (from genomics through to exoplanet exploration) have increased


CIENTIFIC DATA

Amended: Addendum

OPEN

T CATEGORIES
» Research data
» Publication characteristics

10 December 2015
: 12 February 2016
ed: 15 March 2016

Comment: The FAIR Guiding Principles for scientific data management and stewardship

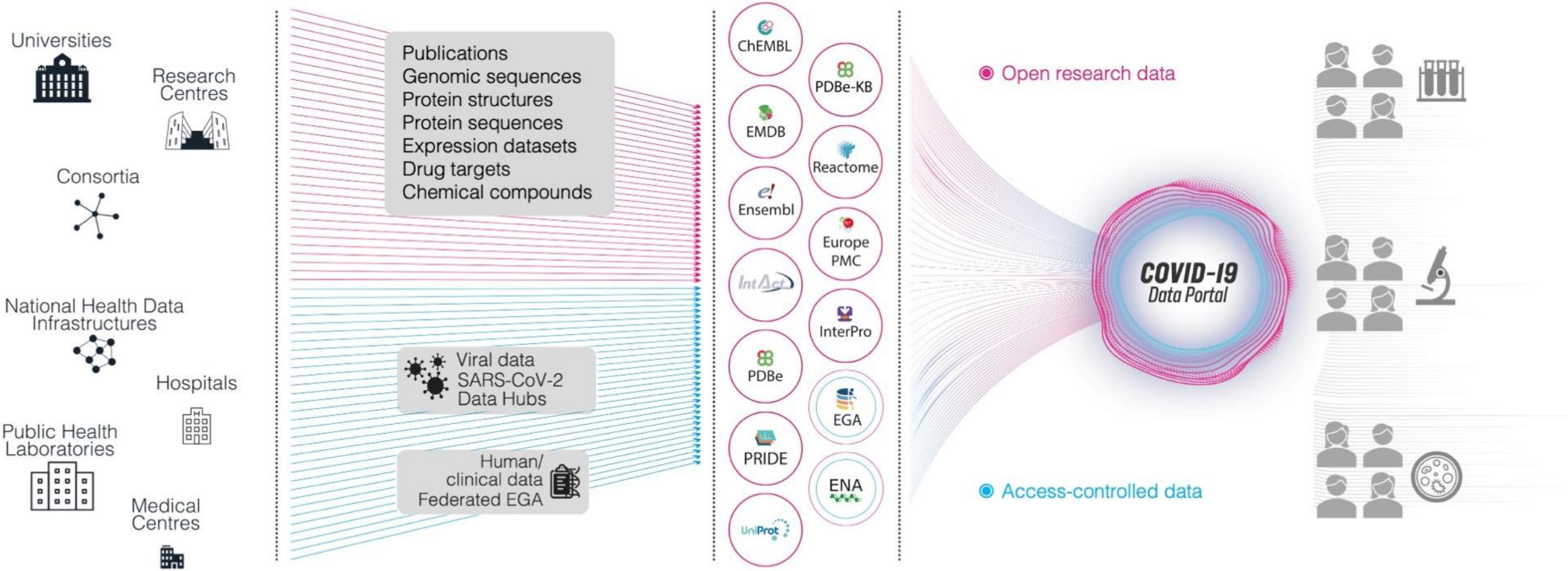
Mark D. Wilkinson *et al.*#

There is an urgent need to improve the infrastructure supporting the reuse of scholarly data. A diverse set of stakeholders—representing academia, industry, funding agencies, and scholarly publishers—have come together to design and jointly endorse a concise and measureable set of principles that we refer to as the FAIR Data Principles. The intent is that these may act as a guideline for those wishing to enhance the reusability of their data holdings. Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals. This Comment is the first formal publication of the FAIR Principles, and includes the rationale behind them, and some exemplar implementations in the community.

Supporting discovery through good data management

Good data management is not a goal in itself, but rather is the key conduit leading to knowledge discovery and innovation, and to subsequent data and knowledge integration and reuse by the

The European COVID-19 Data Platform



https://www.covid19dataportal.org

Guy Perrière
ELIXIR-FR Head of Node

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

Angela Sáenz
ELIXIR-FR Node coordinator

Many thanks to the ELIXIR hub team for their slide deck!

www.elixir-europe.org

@ELIXIREurope

/company/elixir-europe