



Annual Report

2021





Annual Report

2021



Contents

6 Foreword

7 Preface

10

Our EU grants

10 BY-COVID

11 ELIXIR-CONVERGE

12

**Our progress against
strategic objectives**

14

ELIXIR and COVID-19

16

Our Services

16 ELIXIR Services

17 ELIXIR
Commissioned
Services

18

Our Core Structure

19 ELIXIR Nodes

31 ELIXIR Platforms

41 ELIXIR Communities

48

Our flagship events

- 48 All Hands Meeting - connecting all ELIXIR members
- 49 BioHackathon Europe - ELIXIR's foray in hybrid hacking

50

Our collaboration, outreach and industry support efforts

- 50 Strengthening global collaboration
- 51 Enabling and supporting industry collaboration
- 52 Demonstrating ELIXIR's impact
- 53 Stakeholder engagement
- 54 Node capacity building

55

Our people

- 55 ELIXIR Hub staff
- 56 Governance

60

Financial Data

ELIXIR Glossary



Foreword



As a member of the ELIXIR Scientific Advisory Board (SAB) since April 2016, I have had the opportunity to see the great work performed by ELIXIR across its Nodes, Communities and Platforms. This past year was no exception. The mid-term review was a showcase of ELIXIR's excellent work and skills in enabling stakeholders to consolidate and expand their expertise to advance Open Science in the rich European and international landscape.

The SAB has been impressed with the successful efforts of ELIXIR's Tools and Compute Platforms to improve the reproducibility of data analysis and in enabling federated environments to deliver large-scale workflow analysis across international boundaries with the GA4GH initiative. Many other large-scale programmes are to be commended, including ELIXIR's coordination of COVID-19 initiatives, the Beyond One Million Genomes (B1MG) project and efforts to support transnational access to genomics data in Europe.

As an advisory body, the SAB never hesitates to make recommendations for improvements, this year we were more than impressed with not just how the suggestions were received, but how many were implemented and others challenged. We have reached a state in the relationship where we help each other whilst learning together, making us even better ambassadors of this important initiative.

In passing the mid-point of the 2019-2023 ELIXIR Scientific Programme, the agility and skills of the ELIXIR scientific community remains evident. The team has demonstrated the ability to turn unforeseen challenges into opportunities to improve Open Science and build an interactive community. All of this whilst keeping an eye on the goal: bioinformatics in 2023 – the opportunity of digital life science!

Like the former chair of the SAB, Robert Gentleman, I believe the future of ELIXIR is bright, and that it will continue to harness the potential of life-science data to maximise impact in research, Open Science, innovation and society at large.

BF Francis Ouellette

Chair of ELIXIR Scientific
Advisory Board (2018 - 2022)

Montréal, Canada

March 2022

A handwritten signature in blue ink, reading "Francis Ouellette".

Preface



2021, the second year of COVID-19, began with lockdowns in European countries and many of us grappling with the impact of the pandemic on our health and personal lives. The year also brought good news with efficacious vaccines offering hope. While much remains to be done before the world is safe, I see 2021 as a year where science showed the value it brings to society, and international collaborations showed their power to tackle global challenges.

Despite the challenges of the pandemic and the disruption it brought to our work, I remained delighted that ELIXIR once again showed its value and impact, including, but not limited to, work coordinating infectious disease data-related activities across Europe. The EU-funded BY-COVID project was awarded to ELIXIR in 2021, complimenting the award of new funding uplifts to the ELIXIR-CONVERGE project to support the coordination of infectious disease data across borders. These successes show the importance of ELIXIR and of European collaboration, in bringing together experts, their databases and tools, and coordinating activities so scientists, policy-makers and health experts have an effective infrastructure to re-use and analyse essential data.

We responded to these challenges by developing new ways of working to allow the smooth integration of new joiners and the incorporation of new ELIXIR Communities into our activities. This would not have been possible without the established collaborations and collegiate spirit of all those involved in ELIXIR across all of our Member States and associated projects. The pandemic has also demonstrated the critical role of Open Science in international collaborations: without an Open Science foundation, ELIXIR would not have succeeded in quickly bringing together teams across organisations, countries and projects to tackle the pandemic.

2021 was also the mid-way point of the current ELIXIR Scientific Programme (2019-2023), presenting the opportunity for us

to reflect on activities so far and look to future plans. A mid-term review carried out by our Scientific Advisory Board (SAB) in June delivered expert, independent advice on our progress. I am pleased to see the positive feedback received on all aspects of our scientific programme and would like to thank everyone involved in ELIXIR for their efforts and hard work, both in preparing for the review and in developing our plans in response to the many constructive suggestions from the SAB. The summer also saw the launch of a new tranche of Commissioned Services - projects funded through ELIXIR's core budget, which are used to connect, coordinate and integrate bioinformatics resources across the ELIXIR Nodes. We started 12 Commissioned Services, including important work to lay foundations for federated access and compute functions via container services, and to expand Node capabilities in project monitoring and impact assessments.

As we look ahead into 2022 we will build on the learnings from the pandemic: the value of strong collaborations, European networks and Open Science to mount rapid responses. ELIXIR is a distributed infrastructure - our services, expertise and support to scientists is delivered through our Nodes - and these values are at the core of our work. We are a truly pan-European federated data infrastructure for the life sciences, built upon strong national capabilities, all connected through solid collaborations and high-quality services.

Niklas Blomberg

ELIXIR Director

Cambridge, UK

March 2022

ELIXIR in 2021



DRIVING PROJECTS FORWARD

13

ONGOING
EU-FUNDED
PROJECTS

In 2021 the ELIXIR Hub project portfolio contained 13 ongoing EU-funded projects, 4 in a coordination role (starred):

*B₁MG

BiCIKL

CINECA

*ELIXIR-CONVERGE

EOSC Enhance

EOSC-Future

*EOSC-Life

eTRANSafe

European Joint Programme on Rare Diseases

EU-STANDS₄PM

*FAIRplus

HealthyCloud

RtrainPlus

Has started coordinating one new project:

BY-COVID

€12M

to provide the infrastructure for making data from COVID-19 and other infectious diseases open and accessible to everyone

And has received a funding uplift:

ELIXIR-CONVERGE

€2M

to mobilise SARS-CoV-2 variant surveillance data tracking services and tools to feed the European COVID-19 Data Platform

Funded 12 new Commissioned Services with a total investment of

€3.3M



Held the ELIXIR Scientific Programme 2019-2023 mid-term review

40 people

15 for 15 hours

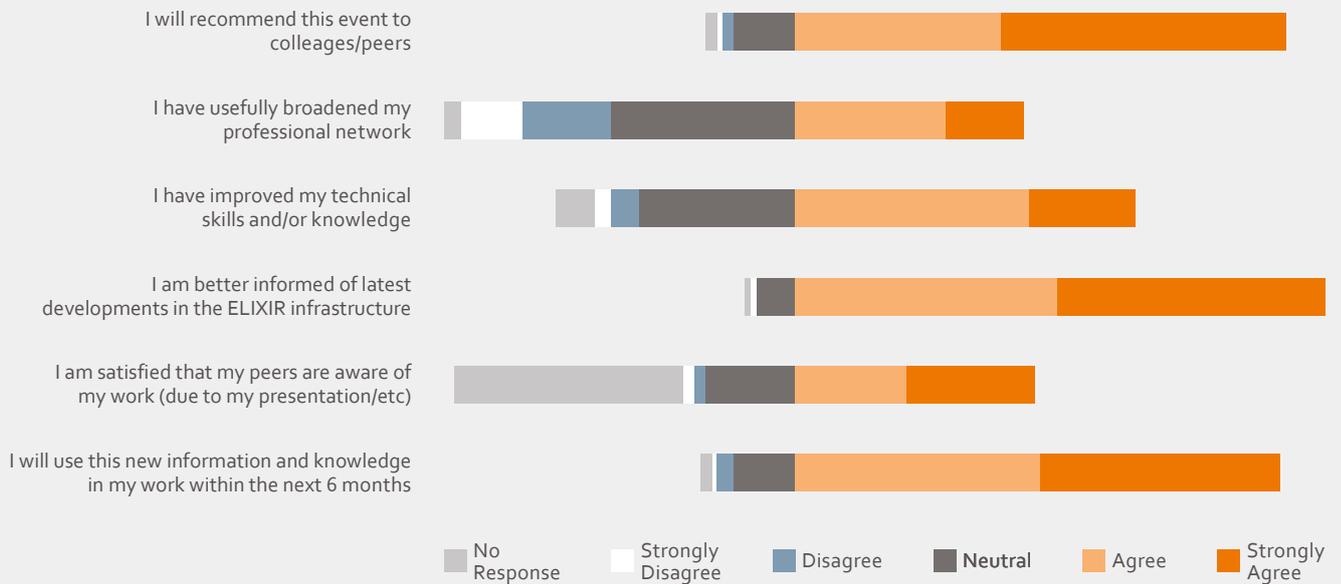
3 over 3 days

"We congratulate ELIXIR on its success and continued progress since its establishment in 2014 and throughout the current 2019-23 ELIXIR Programme. In this time, ELIXIR has built an effective infrastructure, usable applications and a strong community." ELIXIR Scientific Advisory Board, Scientific Programme 2019-2023 mid-term review

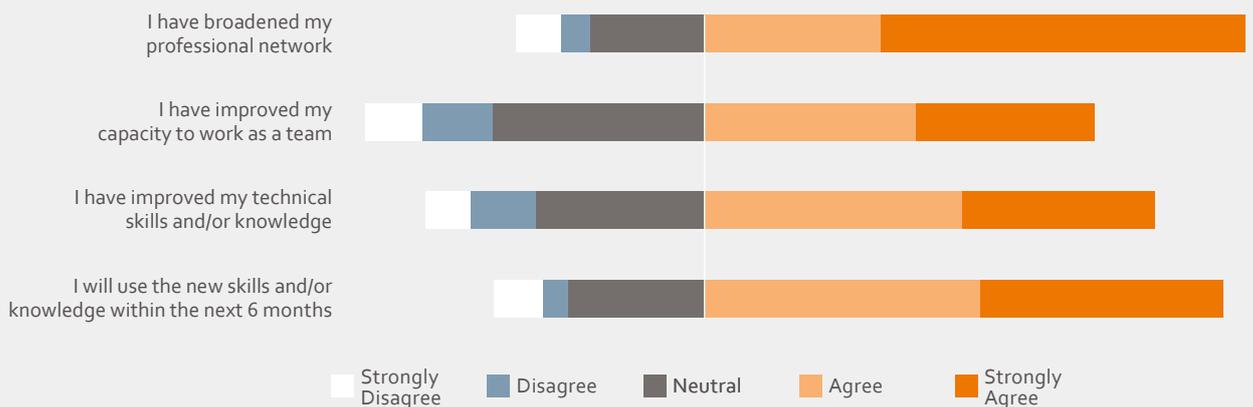


**HOSTED HIGH
IMPACT EVENTS**

All Hands meeting 2021



BioHackathon Europe 2021



Our EU grants

To drive forward our goals in the field of bioinformatics and provide services to the scientific community, ELIXIR participates in several EU-funded projects. These enable ELIXIR to collaborate with key European and global initiatives, industry partners and other research infrastructures.



BY-COVID

In 2021, ELIXIR started one new project as coordinator:

TIMELINE October 2021-September 2024

BUDGET €12M

WEBSITE <https://by-covid.eu>

BY-COVID is an interdisciplinary project uniting life science, medical, policy, social science and public health experts from across Europe. Led by ELIXIR, the project has 53 partners from 19 European countries. The BY-COVID project will run for three years, starting in October 2021, and is part of the European Commission's HERA Incubator plan,

Anticipating together the threat of COVID-19 variants. The project will build and expand upon the successful European COVID-19 Data Platform, a resource initiated in the early stages of the pandemic and led by EMBL-EBI. Rather than focus purely on providing technical solutions to the management of molecular data, BY-COVID will work with partners such as the Versatile Emerging infectious disease Observatory (VEO), the Public Health Information Research Infrastructure (PHIRI) and the Consortium of European Social Science Data Archives (CESSDA) to incorporate data from a broad range of disciplines, including public health and social sciences.



ELIXIR-CONVERGE

In 2021, ELIXIR-CONVERGE received an uplift of €2M:

TIMELINE February 2020-February 2023

BUDGET €7.8M, increased to €9.8M following a second COVID-19 uplift

WEBSITE elixir-europe.org/about-us/how-funded/eu-projects/converge

The ELIXIR-CONVERGE project was launched in February 2020, building on the success of the ELIXIR-EXCELERATE project, which kick-started ELIXIR's work in the field of data management provision across ELIXIR Nodes.

In 2020, in light of the COVID-19 pandemic, the European Commission awarded ELIXIR-CONVERGE an extra €2 million to enhance human and

viral data sharing through the ELIXIR Nodes and support the development of the European COVID-19 Data Platform. Subsequently, in 2021, the ELIXIR-CONVERGE project received a second funding uplift of €2million to focus on two tasks: to facilitate the mobilisation of viral genomes from national sequencing efforts and to offer open data resources for surveillance of COVID-19 variants.

These new tasks will focus on updating the submission and annotation pipelines of the European Nucleotide Archive (ENA), which will help link nascent and established national sequencing efforts to the European COVID-19 Data Platform. In addition, these efforts will ensure the commitment to the open sharing of variant data through the ENA.

2021 achievements:

- Established the Data Manager Network and Data Managers

Coordinator Group (over 120 international experts actively participating)

- Developed and delivered 36 data management and data stewardship training events to 1,121 participants by 13 ELIXIR Nodes
- The RDMkit community contributed domain pages in plant sciences¹, marine metagenomics², human data³, biomolecular simulation data⁴, intrinsically disordered proteins⁵, microbial biotechnology⁶, epitranscriptome data⁷, proteomics⁸, toxicology data⁹ and bioimaging data¹⁰
- Tool Assemblies¹¹ provided exemplar toolkit configurations cross-linked to domain pages, focusing on human data (clinical and translational data, COVID-19 data), bioimaging, plant sciences, marine metagenomics and general data management

1 https://rdmkit.elixir-europe.org/plant_sciences.html
2 https://rdmkit.elixir-europe.org/marine_metagenomics.html
3 https://rdmkit.elixir-europe.org/human_data.html
4 https://rdmkit.elixir-europe.org/biomolecular_simulation_data.html
5 https://rdmkit.elixir-europe.org/intrinsically_disordered_proteins.html
6 https://rdmkit.elixir-europe.org/microbial_biotechnology.html
7 https://rdmkit.elixir-europe.org/epitranscriptome_data.html
8 <https://rdmkit.elixir-europe.org/proteomics.html>
9 https://rdmkit.elixir-europe.org/toxicology_data.html
10 https://rdmkit.elixir-europe.org/bioimaging_data.html
11 https://rdmkit.elixir-europe.org/tool_assembly.html

Our progress against strategic objectives

As defined in the Scientific Programme for 2019-2023, ELIXIR has five strategic ELIXIR-wide objectives. These cascade down to the more specific objectives in ELIXIR Platforms, Communities, and in scientific and technical collaboration with ELIXIR partners.

1

ELIXIR will operate a portfolio of integrated services that meet the data needs of life scientists at a European scale

DATA
PLATFORM

P. **31**

INTEROPERABILITY
PLATFORM

P. **33**

TOOLS
PLATFORM

P. **34**

COMPUTE
PLATFORM

P. **37**

TRAINING
PLATFORM

P. **39**

2

ELIXIR Communities will drive service uptake, support standards development, and connect ELIXIR's experts in life science disciplines

ELIXIR
COMMUNITIES

P. **41**

3

ELIXIR Core Data Resources will be the global standard for bioinformatics resource management and the foundation for an international funding and life cycle management strategy that secures the long-term sustainability of those resources

DATA PLATFORM

P. 31

GLOBAL BIODATA COALITION COLLABORATION

P. 50

4

ELIXIR will be the recognised and trusted life science foundation of the European Open Science Cloud

DATA PLATFORM

P. 31

INTEROPERABILITY PLATFORM

P. 33

TOOLS PLATFORM

P. 34

COMPUTE PLATFORM

P. 37

5

All ELIXIR Nodes will connect life science users in academia and industry to our open, federated service network

ELIXIR NODES

P. 19

BIOHACKATHON EUROPE

P. 49

ENABLING AND SUPPORTING INDUSTRY COLLABORATION

P. 51

DEMONSTRATING ELIXIR'S IMPACT

P. 52

ELIXIR and COVID-19

ELIXIR HAS

RECEIVED

€12M

funding for a new project, BeYond-COVID (BY-COVID), to provide open data on SARS-CoV-2 and other infectious diseases across scientific, medical, public health and policy domains



AND RECEIVED A SECOND FUNDING UPLIFT OF

€2M

for ELIXIR-CONVERGE, to further expand the rapid collection of COVID-19 data to improve research efficiency and pandemic preparedness



EOSC-LIFE HAS

RECEIVED

€2.6M

funding for further development of the COVID-19 Data Portal and the COVID-19 Trial Repository



ELIXIR NODES HAVE

LAUNCHED

9

COVID-19 NATIONAL DATA PORTALS

PUBLISHED

13

COVID-19 PAPERS

Facilitated data deposition to ELIXIR's Deposition Databases, in particular the European Nucleotide Archive (ENA)

Enabled access to SARS-CoV-2 data through curated tools

Increased harmonisation of SARS-CoV-2 data through ELIXIR's Recommended Interoperability Resources

Provided software and workflows to analyse SARS-CoV-2 data via the ELIXIR tools registry

Run compute resources to help analyse SARS-CoV-2 datasets

7.5M

Records added to
European COVID-19 Data
Platform in 2021

 New National
COVID-19 Data
Portals



Our Services

ELIXIR Services

ELIXIR coordinates the provision of life science services, developed and managed across Europe by ELIXIR Nodes. These services are available to researchers around the world and support efficient manipulation, analysis, storage and exchange of life science data.

The portfolio of ELIXIR services is organised into five technical Platforms: Data, Tools, Compute, Interoperability and Training. Additionally, selected services are part of three key services collections: ELIXIR Core Data Resources, ELIXIR Deposition Databases and ELIXIR Recommended Interoperability Resources.

Selected ELIXIR Services:

The content and usage of a number of key ELIXIR services and resources in 2021

ELIXIR AAI

- Over 6200 logins per month
- Enabled members from over 1100 institutions to use ELIXIR AAI
- Integrated 132 production services in total and an additional 110 in testing
- Had 6400 users

FAIRsharing

- 1846 databases
- 152 policies
- 1568 standards

ELIXIR TeSS

- 1591 training materials
- 22 training workflows
- 83 content providers

BioContainers

- 9k tools
- 69.2k containers and packages
- 18TB in total

BioTools

- 23,000 entries
- 411,000 annotations
- 4,300 users
- 35,000 monthly visits

The number of services run by ELIXIR Nodes continued to expand in 2021

430

SERVICES IN TOTAL

21

COMPUTE

128

DATA RESOURCES

47

INTEROPERABILITY

282

TOOLS

27

TRAINING

ELIXIR Commissioned Services

ELIXIR Commissioned Services are funded through the ELIXIR budget to drive the integration of services operated by the ELIXIR Nodes. They are proposed and managed by a particular Platform or Community, agreed with the ELIXIR Heads of Nodes Committee, and approved by the ELIXIR Board.

The Commissioned Services are the main instrument to achieve the strategic objectives of the Scientific Programme for 2019-2023. Commissioned Services are divided into seven different categories: Platform Tasks, Community Implementation Studies, Community-led Implementation Studies, Strategic Implementation Studies, Infrastructure Services, Staff Exchange Programme and ELIXIR Industry Engagement.

In 2021, a total of 44 Commissioned Services were active, engaging activity from across all ELIXIR Nodes.

Twelve of these projects were new and launched in 2021:

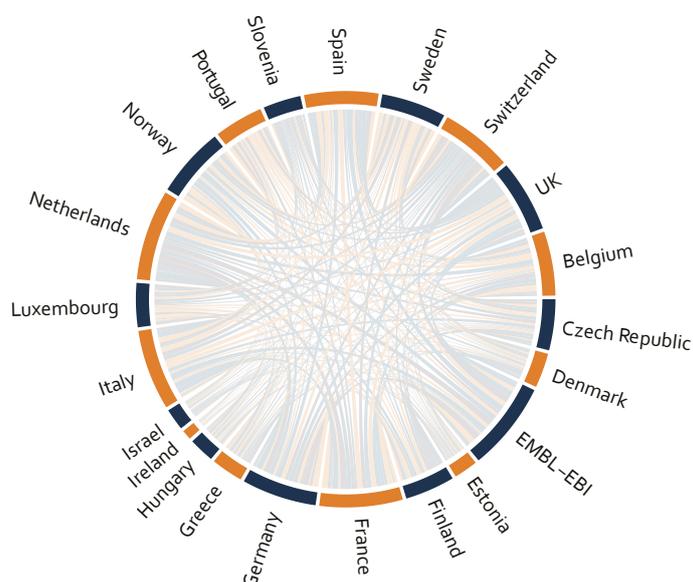
- ELIXIR AAI Service Plan 2021-2023 Infrastructure Service
- Standardising Intrinsically Disordered Proteins (IDPs) data
- Tools Platform Ecosystem
- Impact evaluation at Node-level - getting it done
- Increasing plant data findability and reuse beyond ELIXIR
- Strengthen data management in Galaxy
- Building on PDBe-KB to chart and characterise the conformational landscape of native proteins
- Reference hCNV datasets, use-case workflows and benchmarking
- Beacon and beyond - implementation-driven standards and protocols for CNV discovery and data exchange
- Increasing the translational value of public proteomics datasets: automatic metadata-driven reanalysis in cloud infrastructures
- Improving IDP tools interoperability and integration into ELIXIR
- Making container services integratable, sustainable and widely used.

Each Commissioned Service is led by a team of experts from one or more ELIXIR Nodes, drawing on their national scientific priorities and expertise. Their collaborative nature also enables newly established Nodes to take part in ELIXIR work and quickly integrate their national communities into ELIXIR.

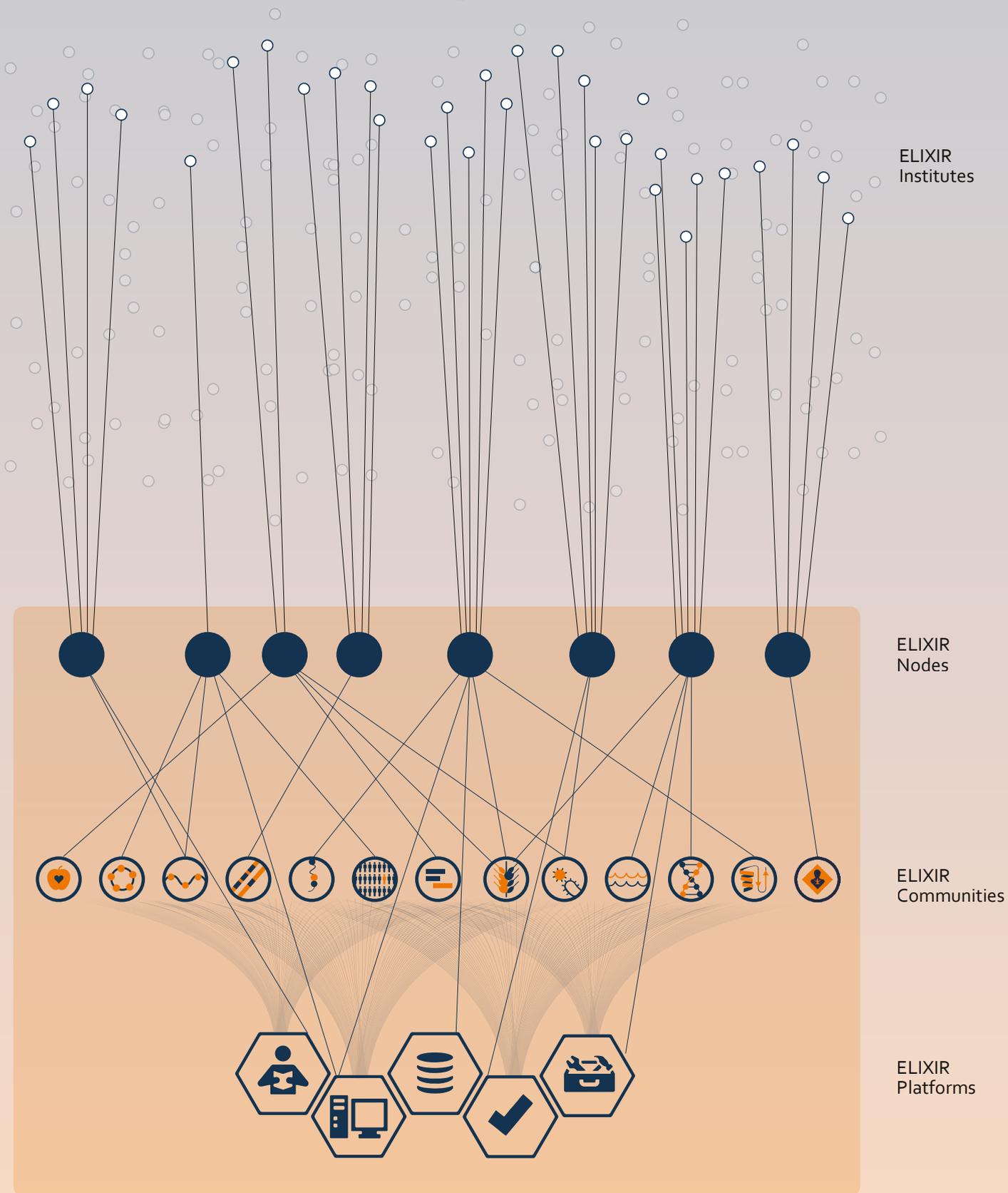
44

COMMISSIONED
SERVICES RAN DURING 2021

Connection between ELIXIR Nodes facilitated by ELIXIR Commissioned Services in 2021



Our Core Structure

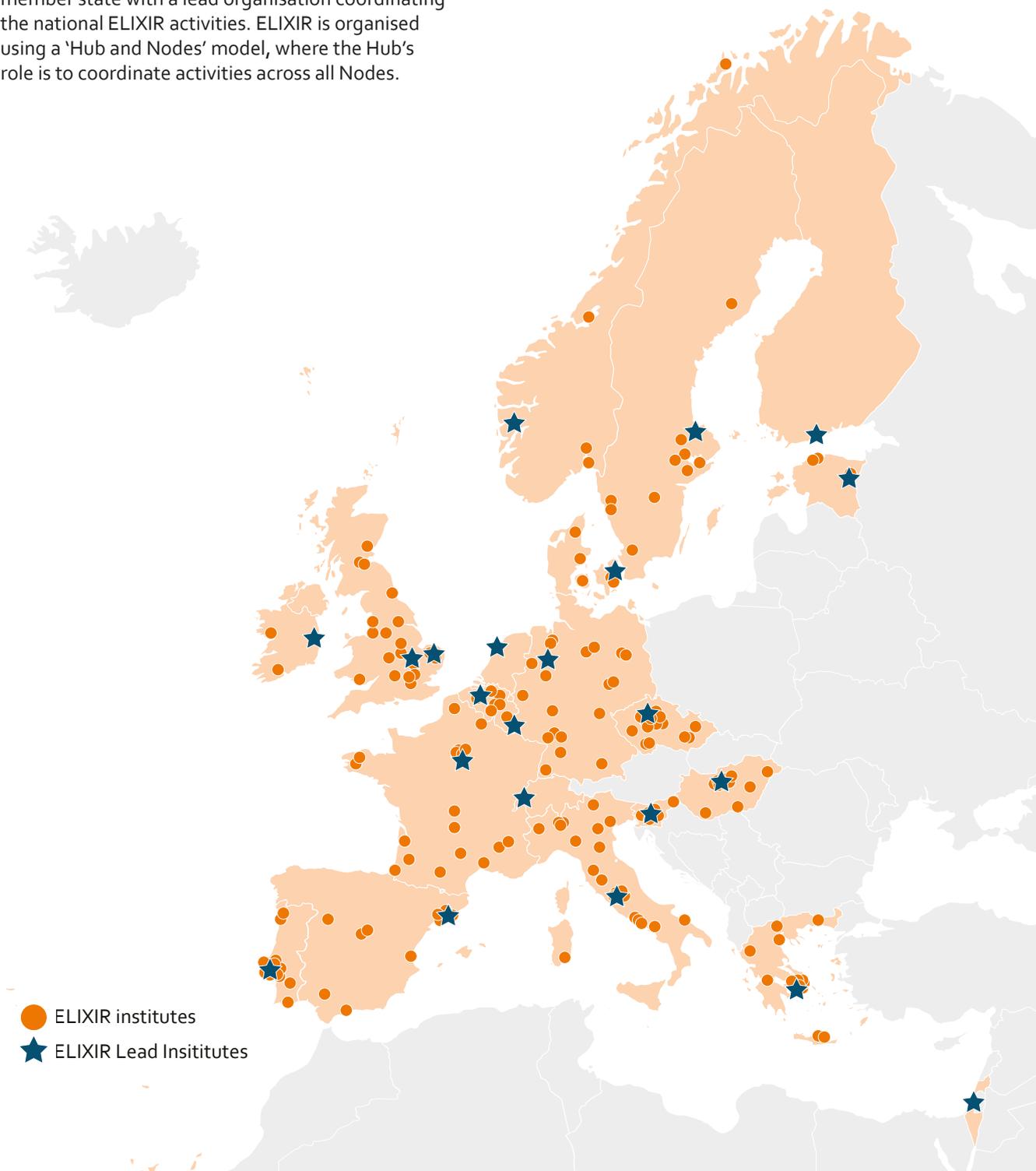


ELIXIR Nodes

Each member state of ELIXIR establishes a Node. A Node is a network of organisations that work within a member state with a lead organisation coordinating the national ELIXIR activities. ELIXIR is organised using a 'Hub and Nodes' model, where the Hub's role is to coordinate activities across all Nodes.

Strategic Objective

5





ELIXIR Belgium

LEAD INSTITUTE VIB

WEBSITE www.elixir-belgium.org

- Contributed to shaping policies around human data infrastructure and establishing collaborations with relevant stakeholders in the Belgian landscape
- Contributed to the development of a Flemish Data Architecture plan as part of the Flemish Open Science Board activities towards EOSC
- Increased Node involvement in ELIXIR Human Data activities through participation in BY-COVID
- Expanded the core team to three members, enlarging communication efforts, activities in data management and developments in useGalaxy.be
- Established DataHub and further extended contributions to the FAIRDOM-SEEK development, in close collaboration with UK and Norwegian Node
- Co-developed WorkflowHub for sharing and publishing of computational workflows
- Organised 22 training courses for a total of 685 participants
- Started three Strategic and six Community-led ELIXIR Implementation Studies
- Co-organised the ELIXIR Innovation and SME event for the Agritech sector and the International Galaxy Community Conference
- Collaborated with the Plant Community in Implementation Studies on BioSchemas and FAIRification of data
- Further developed the COVID-19 data submission toolbox, used by, for example, the Estonian Node.



ELIXIR Czech Republic

LEAD INSTITUTE Institute of Organic Chemistry and Biochemistry of the CAS

WEBSITE www.elixir-czech.cz

- Initiated implementation of the ELIXIR CZ Strategy, leading to an increased number of scientists engaged in ELIXIR CZ activities, the introduction of new services and improvement of existing tools and services
- Released four new ELIXIR CZ services
- Published 50 articles
- Received 931 citations
- Organised 25 events (training and workshops) including training on Galaxy, AI in Life Sciences, and a workshop on Advanced In Silico Drug Design
- Organised 31 webinars and tutorials focused on provided services, recordings are hosted on the ELIXIR CZ Youtube channel¹²
- Participated in the Horizon 2020-funded ELIXIR-CONVERGE project
- Carried out nine Implementation Studies
- Took part in an ELIXIR Staff Exchange.

12 <https://www.youtube.com/channel/UCtoSKet24szBGjN-V1d4EKw/videos>



ELIXIR Denmark

LEAD INSTITUTE Technical University of Denmark

WEBSITE www.elixir-denmark.org

- Expanded Node membership to include the Danish agency for disease control (SSI) and the Danish agency for Genome Data (NGC)
- Supported national COVID-19 efforts by providing infrastructure, data analysis and data management expertise
- Participated in several Implementation Studies and the ELIXIR-CONVERGE project
- Collaborated with European Joint Programme on Rare Diseases (EJP RD)
- Organised new training and courses in precision medicine and computational biology in partnership with University partners
- Hosted the 6th Annual Danish Bioinformatics Conference 2021 in Aalborg attracting more than 260 participants
- Participated in the EOSC-Nordic project
- Updated entries in bio.tools and reorganised online access to Danish tools
- Creating a national SANDBOX for de-identified health data to train students and researchers in the handling of sensitive data.



EMBL-EBI

WEBSITE www.ebi.ac.uk

- Hosted the European COVID-19 Data Platform
- Contributed to the FAIR Cookbook and to a report defining metrics to assess the FAIRness of data produced by IMI projects
- Deposited regular EBI plant data into the FAIR Data-finder for Agronomic REsearch (FAIRDARE)
- Participated actively in the ELIXIR Plant Community
- Collaborated in efforts to establish a new ELIXIR Single Cell Omics Community
- Led the development of the MAGE-TAB-Proteomics data standard, enabling the annotation of the experimental design in proteomics experiments
- Authored the white paper *Data Management of Sensitive Human Proteomics Data: Current Practices, Recommendations, and Perspectives for the Future*
- Co-organised the 3D-BioInfo Community meeting with more than 300 active participants
- Participated in the development of an RDMkit for archiving structure data including experimental and predicted models
- Participated in the Containers Strategic Implementation Study, implementing bioconda and biocontainers best practices for the Single Cell Expression Atlas downstream analysis pipeline and defining the use case for the RNA-Seq Single Cell.



ELIXIR Estonia

LEAD INSTITUTE University of Tartu

WEBSITE elixir.ut.ee

- Expanded Node membership to include Tallinn University and Estonian University of Life Sciences
- Awarded Core Research Infrastructures of National Importance funds from the Estonian Research Council
- Developed new courses to enhance data management best practises
- Held nine training events
- Participated in the new Horizon Europe consortia (BY-COVID and HealthyCloud)
- Developed and maintained several public COVID-19 web tools¹³ and launched the Estonian COVID-19 Data Portal¹⁴
- Supported a national COVID-19 prevalence study and SARS-CoV-2 sequencing project with data management, analysis and visualisation expertise.



ELIXIR Finland

LEAD INSTITUTE CSC - IT Center for Science

WEBSITE www.elixir-finland.org

- Benefited from the contributions of 80 experts (totalling more than 25 FTE)
- Renewed Node funding (2022-2026), expanding main objectives to support biobank, national sequencing and imaging centre use cases
- Coordinated four pillars: artificial intelligence support, sensitive data in HPC, health data availability, identity and data access technology
- Led on EU 1+ Million Genomes technical infrastructure development with ELIXIR Sweden, and supported the B1MG project to produce the Proof of Concept key outcome
- Participated in ELIXIR Compute Platform leadership (ExCo and task leadership)
- Published main outcomes of the GA4GH Data Use and Researcher Identity workstream in the Cell Genomics special issue¹⁵
- Coordinated the emerging ELIXIR Single Cell Omics Community
- Participated in the National Genome Centre with the Ministry of Social Affairs and Health, including Sensitive Data Services¹⁶ and RDMKit resources
- Provided training events including ELIXIR-GOBLET events for trainers, training for Data Support Personnel and Sensitive Data Services for Research tutorials
- Collaborated with GA4GH Passport and DUO standards
- Worked on personalised medicine computing on HPC with ELIXIR Spain (BSC) and LUMI (EuroHPC)
- Submitted the first Finnish SARS-CoV-2 sequence data to ENA through ELIXIR-CONVERGE
- Collaborated on 14 ELIXIR Commissioned Service projects including European Identity and Access management infrastructure with ELIXIR CZ and EOSC-LIFE.

¹³ <https://koroona.ut.ee>

¹⁴ <https://covid19dataportal.ee>

¹⁵ [https://www.cell.com/cell-genomics/issue?pii=S2666-979X\(21\)X0003-1](https://www.cell.com/cell-genomics/issue?pii=S2666-979X(21)X0003-1)

¹⁶ <https://research.csc.fi/sensitive-data-services-for-research>



ELIXIR France

LEAD INSTITUTE CNRS

WEBSITE www.france-bioinformatique.fr

- Provided services including the first Data Management Plan (DMP) template version for bioimaging, DMP OPIDoR, FAIRchecker and EDAM ontology
- Released two new courses on FAIR principles applied to bioinformatics and omics datasets, FAIR-bioinfo¹⁷ and FAIR-data¹⁸
- Initiated new projects including MUDIS₄LS, EMERGEN and ABRomics
- Led two community Implementation Studies (hCNV and Plant Science) and partnered in two others
- Partnered in three Strategic Implementation Studies (Impact, Tool ecosystem, Container services)
- Organised a National Human Data day as a Node contribution to B1MG
- Participated in the Bioschemas steering council
- Organised the online course Introduction to Machine Learning using R with ELIXIR-GR and ELIXIR-CH instructors and assistants
- Co-authored the Biodiversity Focus Group vision paper¹⁹
- Published the Institut Français de Bioinformatique (IFB) 2019-2020 activity report²⁰.



ELIXIR Germany

LEAD INSTITUTE Bielefeld University

WEBSITE www.denbi.de/elixir-de

- Benefited from the contributions of 25 principal investigators at 21 institutes
- Extended the ELIXIR Germany Service Delivery Plan to 104 bioinformatics services
- Participated in eight ELIXIR Communities with a leadership role in three of them: Galaxy, Proteomics and Plant Sciences
- Participated in six Community-led Implementation Studies
- Participated in 17 ELIXIR Implementation Studies
- Contributed to the Data Management Network, training activities and RDMkit as part of the ELIXIR-CONVERGE project
- Supported the COVID-19 disease map consortium through the FAIRDOMHub service
- Delivered extensive online training with 53 courses attended by more than 2200 participants
- Mobilised cloud resources for the virtual COVID-19 hackathon and supported COVID-19 activities using the European Galaxy server
- Organised the second metabolomics hackathon as an online workshop
- Organised two joint Bayer AG and de.NBI/ ELIXIR Germany online conferences on the theme of *Women in Data Science - Perspectives in Industry and Academia*
- Contributed to four ELIXIR publications and published an ELIXIR-DE brochure outlining Node tasks and developments²¹.

¹⁷ <https://ifb-elixirfr.github.io/IFB-FAIR-bioinfo-training>

¹⁸ <https://ifb-elixirfr.github.io/IFB-FAIR-data-training>

¹⁹ Waterhouse, Robert M., et al. Recommendations for connecting molecular sequence and biodiversity research infrastructures through ELIXIR. *F1000Research* 10.1238 (2021).

²⁰ <https://www.france-bioinformatique.fr/en/activity-reports/activity-reports-2019-2020>

²¹ https://www.denbi.de/images/Downloads/ELIXIR-DE_brochure_2021.pdf



ELIXIR Greece

LEAD INSTITUTE Biomedical Sciences Research Centre (BSRC) 'Alexander Fleming'

WEBSITE www.elixir-greece.org

- Received ratification of the Consortium Agreement by the Greek Parliament
- Received a second evaluation by the Node's Scientific Advisory Committee
- Rolled out HYPATIA, ELIXIR-GR's cloud infrastructure
- Organised and delivered three training events including the first ELIXIR Fluxomics School
- Co-led the DOME Recommendations publication²²
- Participated in the Systems Biology white paper
- Participated in ELIXIR-CONVERGE, including RDMkit development (via the BioHackathon)
- Participated in ELIXIR Communities (Metabolomics, Proteomics, IDP, Microbial Biotech, Biodiversity)
- Initiated a dedicated Biodiversity-GR Task Force
- Participated in Community-led Implementation Studies (ELIXIR-Impact, IDP)
- Launched the COVID-19 Data Portal-Greece and facilitated deposition of Greek SARS-CoV2 sequence data to ENA
- Published extensively in COVID-19 research and genomic surveillance
- Joined the Core Data Resource Task Force for the streamlining of submission and evaluation of CDR efforts.



ELIXIR Hungary

LEAD INSTITUTE Institute of Enzymology, Research Centre of Natural Sciences, Eötvös Loránd Research Network

WEBSITE elixir-hungary.org

- Selected as one of the top research infrastructures of Hungary by the National Research, Development and Innovation Office
- Participated in ELIXIR-CONVERGE (TTK (Budapest), PTE (Pécs) and NAIK (Gödöllő)) and BY-COVID (ELTE (Budapest))
- Benefited from 57 PhD students attached to ELIXIR Hungary research groups
- Launched the first open Galaxy server at Semmelweis University, Budapest
- Expanded Node membership to include Szeged University, Szeged and Pazmany Peter Catholic University, Budapest
- Organised joint online courses with the ELIXIR Hub
- Organised 24 training courses (by Semmelweis (Budapest), ELTE (Budapest), MATE (Gödöllő), DE (Debrecen) and PPKE (Budapest) Node members)
- Welcomed the National Genomic Laboratory (Pécs) as a bioinformatics partner.

²² Walsh, Ian, et al. DOME: recommendations for supervised machine learning validation in biology. Nature methods 18.10 (2021).



ELIXIR Ireland

LEAD INSTITUTE University College Dublin

WEBSITE elixir-europe.org/about-us/who-we-are/nodes/ireland

- Appointed a Genomics Data Stewardship Manager for training and coordination across multiple institutions and groups
- Participated in the ELIXIR-CONVERGE project
- Participated in developing the ELIXIR Machine Learning Focus Group's published recommendations for supervised machine learning validation in biology (UCD)
- Expanded Trips-Viz as a versatile platform for the interactive exploration of public and user generated ribosome profiling data (UCC)
- Implemented, with NUIG, cloud computing infrastructure for genomics training and research as part of the Science Foundation Ireland Centre for Research Training in Genomics Data Science.



ELIXIR Israel

LEAD INSTITUTE Weizmann Institute of Science

WEBSITE www.weizmann.ac.il

- Participated in the ELIXIR-CONVERGE project
- Participated at the BioHackathon Europe 2021, including in the development of RDMkit
- Participated in a Strategic Implementation Study: *Impact evaluation at Node-level*
- Formed networks with industry and other research institutes
- Reported impact and effectiveness to the Israel Ministry of Science and Technology
- Actively contributed to the 3DBioInfo Community.



ELIXIR Italy

LEAD INSTITUTE CNR Institute of Biomembrane and Bioenergetics

WEBSITE elixir-italy.org

- Prioritised CNR.BiOmics projects, Integrative Omics, computational platforms and human-dedicated resources
- Launched the COVID-19 Data Portal-Italy and collaborated with health authorities for national data-sharing
- Improved Italian COVID-19 data-brokering and infrastructure with ELIXIR-CONVERGE WP9 and BY-COVID
- Organised the ELIXIR Data-Interoperability face-to-face hybrid meeting and the ELIXIR-IT 2021 workshop in Padova
- Co-led the ELIXIR IDP Community and the Machine Learning Focus Group
- Led on the publication of DOME recommendations in Nature Methods
- Participated in 24 Implementation Studies, six of which were newly signed (Data Curation, Interoperability, Benchmarking, Containerisation, Metabolomics, Bioschemas, Beacons and Training)
- Acknowledged by the Italian Ministry for University and Research as a high-priority and global-relevance infrastructure in the National Plan for Research Infrastructures
- Co-organised 15 training events, including five Train-the-Trainer courses, a hackathon on making training material FAIR, two Carpentries workshops and a summer school on Rare Disease Registries and Data-FAIRification
- Contributed to three projects at the BioHackathon Europe 2021.



ELIXIR Luxembourg

LEAD INSTITUTE Luxembourg Centre for Systems Biomedicine

WEBSITE elixir-luxembourg.org

- Provided sustainability solutions to two multi-party international projects: IMI-Pioneer and IMI-PRECISESADS
- Contributed to the establishment of and co-led the ELIXIR Health Data Focus Group
- Participated in BY-COVID, PerMedCoE and HealthyCloud EU Projects
- Co-led three projects at the BioHackathon Europe 2021
- Co-organised the Translational Medicine 2021 Community of Special Interest Group satellite meeting at ISMB 2021
- Continued the co-development of FAIR Cookbook with ELIXIR UK Node
- Organised 10 training events for a total of 287 participants and 124 training hours.



ELIXIR Netherlands

LEAD INSTITUTE Dutch Techcentre for Life Sciences

WEBSITE www.dtls.nl/elixir-nl

- Updated the ELIXIR-NL Service Delivery Plan to include WikiPathways
- Played a key role in the Health-RI national health data infrastructure, bringing expertise on FAIR implementation
- Led the first phase of the national Open Science FAIR implementation programme
- Led training and capacity building activities in CONVERGE, focused on FAIR and metadata
- Organised training on FAIR Data Stewardship and Galaxy
- Contributed to the FAIR Cookbook in the IMI-funded FAIRplus project
- Co-led three new/emerging ELIXIR Communities: Food and Nutrition, Microbial Biotechnology and Toxicology
- Implemented a formal job profile for Data Stewards in the collective Dutch university classification system
- Co-developed Data Stewardship Wizard, recommended by several funders and increasingly connected to RDMkit
- Completed FAIR genomes and defined national standards for common data elements as part of B1MG
- Led the international data steward community via the Data Stewards Interest Group
- Co-led the Interoperability and Training Platforms and the Rare Disease Community
- In the context of 1+MG initiative, co-led the Genome-of-Europe project, data standards working group, Health Technology Assessment work and 2021 Stakeholder Forum
- Initiated the national Digital Life Sciences platform, involving the national Bioinformatics and Systems Biology Community BioSB
- Ran ELIXIR-NL tracks and a Node retreat during the national BioSB conference
- Contributed to the Dutch national mirror group of EJP RD
- Delivered a prototype for FAIR Data Points connecting various distributed COVID patient datasets.



ELIXIR Norway

LEAD INSTITUTE University of Bergen

WEBSITE www.elixir-norway.org

- Facilitated mobilisation of Norwegian SARS-CoV-2 sequences to ENA in collaboration with the Norwegian Institute of Public Health
- Co-led the ELIXIR-CONVERGE uplift to mobilise SARS-CoV-2 variant surveillance data tracking services and tools for the European COVID-19 Data Platform
- Participated in BY-COVID by supporting virological analyses in emerging disease outbreaks and operating the vaccine trial data repository
- Added 14 new services, eight in the service delivery plan: CHOPCHOP, EDAM, FAIRtracks, JASPAR, MirGeneDB, SalmoBase, SARS-CoV-2 DB and TSD
- Joined four new Implementation Studies, co-leading three
- Contributed to the editorial board of the RDMkit
- Received acceptance of FAIRtracks as a Recommended Interoperability Resource
- Organised 20 training and outreach events for more than 400 participants
- Published a report on data management gaps in 10 Norwegian data generating research infrastructures through the BioMedData project
- Established a brokering service to submit data to ENA on behalf of end-users and Norwegian research institutions
- Joined the Norwegian Earth Biogenome Project, aiming to sequence 45,000 Norwegian eukaryotic species
- Secured funding for the third phase of ELIXIR Norway (2022-26), through the Research Council of Norway infrastructure programme.



ELIXIR Portugal

LEAD INSTITUTE Instituto de Engenharia de Sistemas e Computadores

WEBSITE elixir-portugal.org

- Constituted as a not-for-profit association
- Participated in eight Implementation Studies, two Staff Exchange Schemes, and one Knowledge Exchange Scheme
- Participated in three Platforms, five Communities and six Focus Groups
- Appointed a new Deputy Technical Coordinator
- Contributed expertise and resources to the publication of 13 articles
- Contributed to the development of the RDMkit and hosted a toolkit webinar
- Organised an industry engagement event to showcase the outcomes of Knowledge Exchange Schemes with the Navigator Company and Phenospex
- Held a monthly virtual webinar, the BioData.pt Talks
- Hosted seven training courses and three workshops for a total of 223 participants.



ELIXIR Slovenia

LEAD INSTITUTE University of Ljubljana

WEBSITE www.elixir-slovenia.org

- Upgraded the Node's research infrastructure, including compute storage, cluster hardware and high-throughput wet lab equipment
- Implemented a single entry point for booking and reporting on Node infrastructure, tools and services
- Collaborated with Slovenian ISBE Node in the emerging Systems Biology Community
- Hosted a national presentation of new ELIXIR-SI infrastructure and eight related training events
- Organised training events in collaboration with other ELIXIR Nodes
- Delivered new data management and high performance computing courses, archiving all training course materials in ELIXIR-SI eLearning Platform
- Released a local Data Stewardship Wizard instance.



ELIXIR Spain

LEAD INSTITUTE Barcelona Supercomputing Center (BSC) as coordinator of the Spanish National Bioinformatics Institute (INB)

WEBSITE inb-elixir.es

- Updated Node services including IntoGen, CancerGenomeInterpreter, PhylomeDB, metaPhORs, 3DBionotes
- Contributed to 13 Implementation Studies, including Community-led, Strategic and Platform-based ones
- Maintained and increased the engagement with five Platforms, nine Communities and six Focus Groups
- Led on the ELIXIR Tools Platform, and the Rare Diseases and Federated Human Data Communities
- Contributed to IMPaCT, the Spanish infrastructure for Personalized Medicine, including the coordination of the data science pillar
- Received acceptance of FAIRtracks as a Recommended Interoperability Resource
- Participated in B1MG/1+MG, including leadership at the European level and at national mirror groups
- Participated in the IMI-funded FAIRplus project and made contributions to the FAIR Cookbook
- Incorporated three additional partners in the second CONVERGE uplift
- Organised the second edition of the EMBO practical course for core facility managers, introducing sensitive data management and ELSI topics
- Organised the *Sex and Gender Bias in Health and Artificial Intelligence* conference series by the Bioinfo4Women programme
- Contributed towards the establishment of the Spanish Society for Bioinformatics and Computational Biology.



ELIXIR Sweden

LEAD INSTITUTE NBIS — National Bioinformatics Infrastructure Sweden

WEBSITE nbis.se

- Expanded Node membership to include the SciLifeLab Bioimage Informatics facility and the AIDA Data Hub
- Handled over 300 project support requests and consultations with 266 principal investigators
- Received additional funding from the Swedish Research Council to finance the federated EGA-SE Node
- Funded and initiated the three-year Nordic project NeIC Heilsa (Tryggve3)
- Awarded a new six-year EU-project BigPicture, creating a European infrastructure for digital pathology
- Organised 18 NBIS courses and collaborated on six additional international training events
- Led the creation of a European data steward expert network within CONVERGE
- Contributed to development of infrastructure for the European 1+ Million Genome project and finalised the first Proof of Concept
- ELIXIR Training ExCo Jessica Lindvall elected chair of a EOSC Task Force
- Co-organised a summer school on Advanced Topics in Single Cell Analysis
- Held the virtual international summer school RaukR (Advanced R for Bioinformatics)
- Participated in four implementation studies, for example, MetaboIGNITER, used by groups in the US, Australia and the World Health Organization's cancer agency
- Engaged in seven Communities and six Focus Groups.



ELIXIR Switzerland

LEAD INSTITUTE SIB Swiss Institute of Bioinformatics

WEBSITE www.sib.swiss

- Provided the Swiss Pathogen Surveillance Platform as a central hub for all Swiss SARS-CoV-2 sequences to serve the federal authorities and boost data sharing
- Launched a federated query system to assess the availability of consented datasets of all university hospitals as part of the Swiss Personalised Health Network
- Joined forces with European public-private research consortium HIPPOCRATES to improve psoriatic arthritis diagnostic and treatment
- Received acceptance of Cellosaurus and Rhea as Core Data Resources
- Co-steered European efforts to foster open sharing of SARS-CoV-2 genomic data as part of ELIXIR-CONVERGE
- Co-led training and capacity-building of COVID-19 Data Hub users and data providers as part of BY-COVID
- Participated and led the h-CNV Community Implementation Study and two follow-up projects
- Participated in the ELIXIR-GA4GH partnership
- Contributed to ELIXIR Beacon by providing a reference resources for the GA4GH approval process (progenetix.org)
- Contributed to the FAIR Cookbook
- Held the [BC]2 Basel Computational Biology Conference including a one-day SME and Innovation Forum
- Received recognition of an Early Career Bioinformatician, a Bioinformatics Graduate Paper and an Innovative Bioinformatics Resource in 11th Bioinformatics Awards.



ELIXIR UK

LEAD INSTITUTE Earlham Institute

WEBSITE elixiruknode.org

- Expanded to 21 members, adding the University of East Anglia and King's College London
- Increased services to 27, adding Flybase, NEOF, FAIR Cookbook, KnetMiner, COPO, VEuPathDB and WorkflowHub
- Won a national award (Hidden REF) for the applications of GtoPdb to research
- Organised the UK Conference on Bioinformatics and Computational Biology
- Received national funding to build a UK Data Stewardship fellowship and seed a national network
- Collaborated with SSI, UK Reproducibility Network, OpenLifeScience and EuroBioImaging
- Ran a major feasibility study with national funders to prepare a case for a national infrastructure and institute for FAIR data stewardship (BioFAIR)
- Co-lead the Health Data Focus Group, ran a UK health data workshop, and collaborated with Health Data Research UK co-leading their technical pillar on federated analytics
- Co-led RDMkit (CONVERGE) and FAIR Cookbook (FAIRplus, Pistoia Alliance)
- Presented FAIRsharing, RDMkit and FAIR Cookbook at NIH meetings and received recommendations by EOSC, the Innovative Health Initiative and the EC
- Co-lead the Galaxy Community, WorkflowHub and FAIR principles for workflows, presenting at EOSC Symposium, FAIR Festival and more than eight international conferences
- Led metadata frameworks RO-Crate and Bioschemas mark-up and piloted EOSC Enhance metadata exchange with OpenAIRE
- ExCo of Interoperability Platform, co-led three Communities and four Focus Groups
- Participated in 36 Commissioned Services, six ELIXIR-related EC projects and six ELIXIR-related Horizon Europe proposals.

ELIXIR Platforms

ELIXIR activities are divided into five Platforms, each focusing on one specific area in bioinformatics service provision: Data, Tools, Interoperability, Compute and Training.

The Platforms bring together experts from within ELIXIR to develop a vision for the development and operation of activities across the ELIXIR Nodes, drawing on their technical expertise and resources.

Each Platform has three Platform Leads appointed by the ELIXIR Heads of Nodes Committee. The Platform Coordinator, based at the ELIXIR Hub, manages the work across the Nodes providing support to Platform Leads, overseeing the implementation projects and liaising with other ELIXIR Platforms and Communities.



Data Platform

In 2021, the ELIXIR Data Platform successfully led the call for Scalable Curation Implementation Study proposals resulting in the selection of five Implementation Studies running during 2022-23. These studies are foundational to understanding Community use cases to enable the Data Platform to create a complete data resource network.

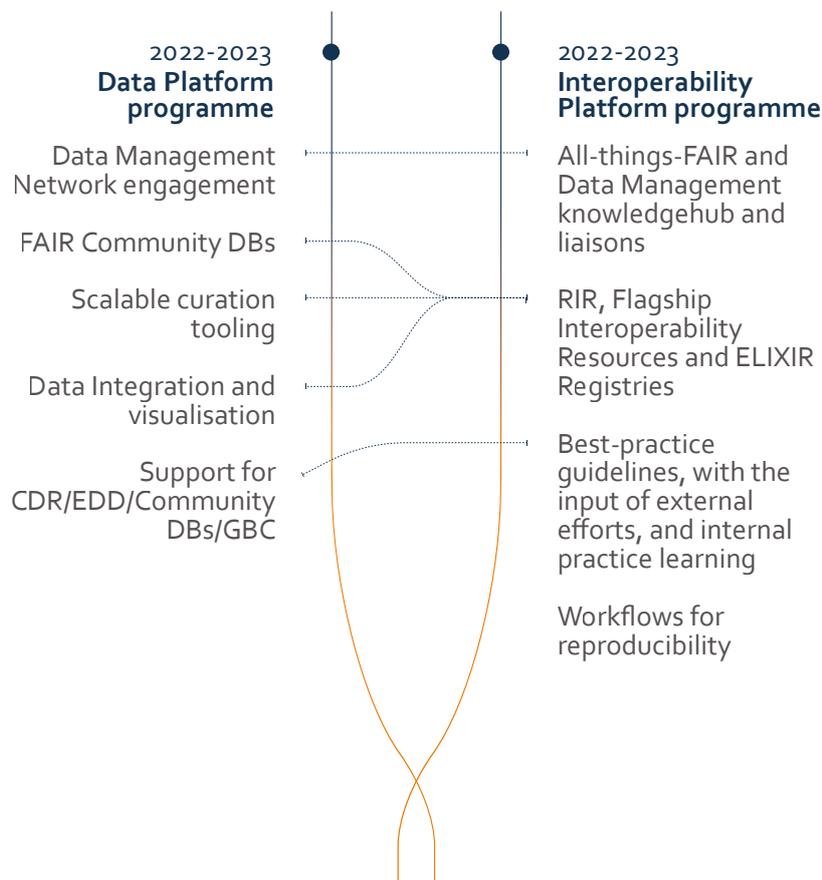
Data Platform 2022-23 work plans have been delivered as part of the on-going Data strategy and will feed into 2024-2028 ELIXIR Scientific Programme. Priorities include:

- Connection across different data resources: reference databases (the Core Data Resource framework), data repositories (the ELIXIR Deposition Database), aggregation databases for specific communities and metadata registries (in partnership with the ELIXIR Interoperability Platform and the Registries Focus Group).
- Support for Communities through scalable curation and community curation.
- Increasing the visibility and impact of ELIXIR through international collaborations.

Engagement with international stakeholders remains an important focus. The Data Platform has worked closely with the Global Biodata Coalition (GBC) to align the global Core Data Resource (CDR) selection process with the ELIXIR CDR selection. This cooperation ensures visibility for ELIXIR on the international stage with the policy makers and international funders. The Platform has successfully coordinated and launched the ELIXIR Biocuration Focus Group and become the bridge between ELIXIR and the international biocuration stakeholders such as the International Society for Biocuration. The ELIXIR Research Data Alliance (RDA) Focus Group has worked closely with the Data Platform in engagement around biocuration. Active collaboration has been fostered by connecting the Biocuration Focus Group with the RDA Focus Group and across data-centric initiatives.

Strategic Objectives





2024-2028 Programme
Data/Interoperability strategy



The Data Platform has forged closer links with the Interoperability Platform, culminating in a first hybrid face-to-face Data-Interoperability meeting in Padova, Italy. The event was well attended with over 100 participants. Several Data Platform use cases were identified involving a range of stakeholders, including:

- The Health Data Focus Group - using Data Platform tools in scalable curation tasks (for example, Triage).
- The Registries Focus Group (in collaboration with the Interoperability Platform) - defining data resource types to complete the data resource network, and understanding the role of metadata cataloguing and metadata exchange in improving the FAIRness of data.
- The CONVERGE Data Management Expert Network - initiating a knowledge and expertise transfer of the governance structure of the CONVERGE Data Management Expert Network in the context of ELIXIR Technical Activities.
- The FAIR Cookbook - strengthening interactions with closer Data-Interoperability alignment.



Key collaborative advances at joint Data & Interoperability Platform event

The Data and Interoperability Platforms held their first joint face-to-face meeting in Padova, Italy.



Interoperability Platform

The Interoperability Platform has added four additional Recommended Interoperability Resources (RIRs) to the collection of tools recommend for fit-for-purpose FAIRification tasks and actions:

- PLAZA - an integrated resource for functional, evolutionary and comparative genomics in plants developed by the Belgian Node
- FAIRtracks - a JSON schema defining minimal information to consolidate genomic track metadata hosted by the Norwegian Node
- OmicsDI - a knowledge discovery framework for metadata exchange of heterogeneous omics data led by the EBI-EMBL Node
- BILL (Bioimaging Informatics Index) - a bioimaging resource of

software tools, training materials and imaging datasets led by the French Node.

Together, the RIRs fulfil the FAIRification needs identified by the Interoperability Platform FAIR Service Framework based on ELIXIR Community user requirements.

The Interoperability Platform has prioritised engagement with international stakeholders in 2021:

- The RDA ELIXIR Bridging Force Interest Group, under the coordination of the Interoperability Platform, has evolved to become the RDA Life Science Infrastructure Interest Group and include international stakeholders such as the Australia BioCommons, the

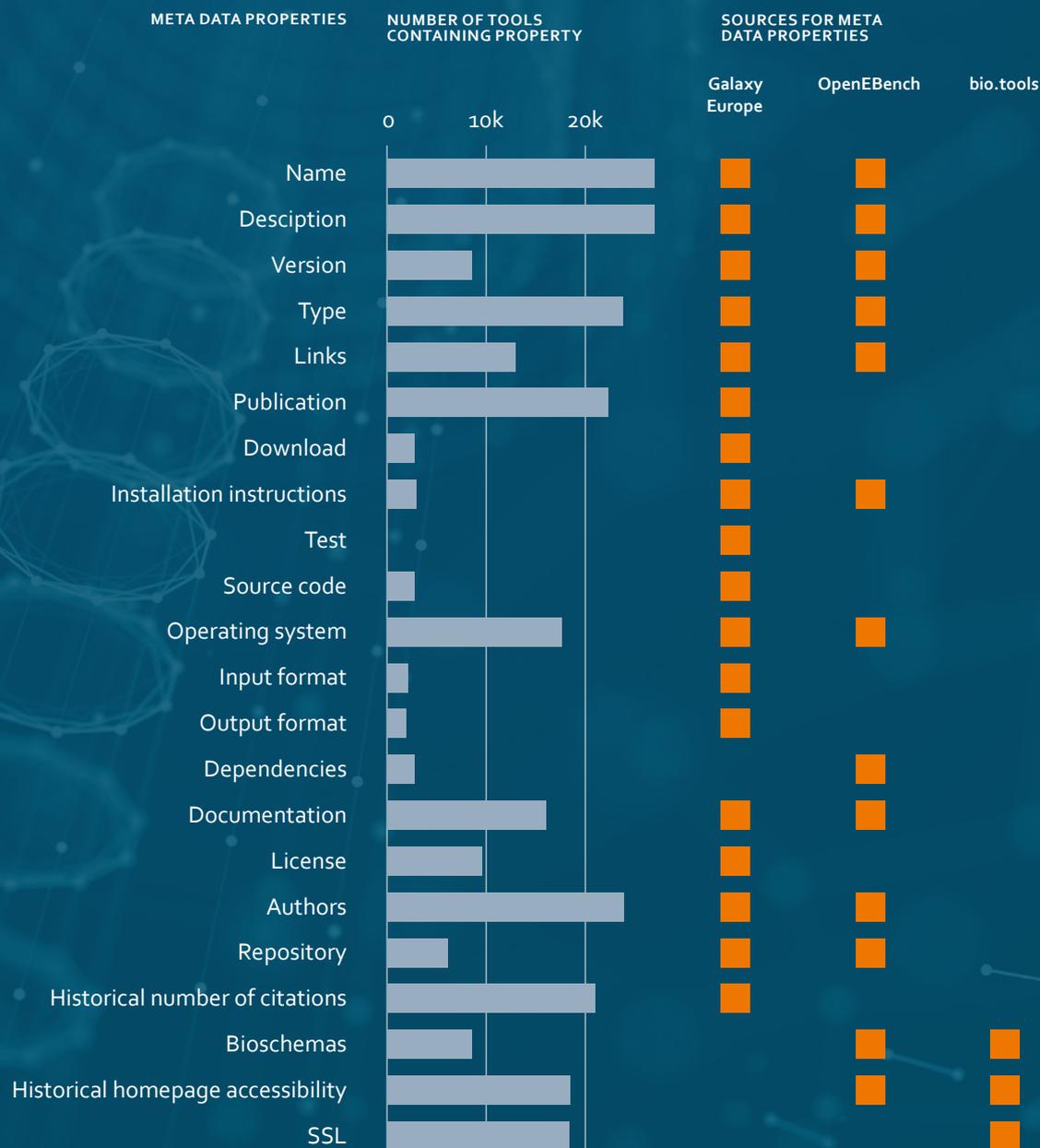
- US NIH Office of Data Science Strategy and H3ABioNet in Africa
- The RDA FAIRsharing Working Group has started a new activity with funders and publishers in RDA Interest Groups to harmonise data policies
- The FAIR Cookbook, with hundreds of authors in ELIXIR Nodes and the pharmaceutical industry, has been endorsed in IMI guidelines
- FAIRsharing and the RDMkit have been named and endorsed in the Horizon Europe guidelines, which also include a reference to the RIRs and several other Interoperability Platform services
- RO-Crate and FAIRsharing have been included as *de facto* elements of the EOSC ecosystem.

Strategic Objectives

1 4



ELIXIR Tools ecosystem metadata origins and properties. Source of metadata properties (right) and number of entries containing the property (left).



Galaxy Europe Galaxy instance hosted by University of Freiburg
 OpenEBench ELIXIR benchmarking & technical monitoring platform
 bio.tools ELIXIR registry of software tools

Tools Platform tasks and engagement beyond ELIXIR



The Tools Platform brings together experts to work on activities to help scientific communities find, register and benchmark software tools and workflows. The activities also help improve software sustainability and quality through best practices. Research software helps researchers access, analyse and integrate biological data to drive scientific discovery across the life sciences.

In 2021, the Tools Platform focused on improving container production, widening community adoption of OpenEbench and bio.tools and increasing adoption of the ELIXIR Software Management Plan in life sciences. The EDAM ontology service

was added to the Norway and French Node delivery plans and WorkflowHub to the UK Node delivery plan.

The BioContainers project developed tooling and infrastructure to automatically convert software packages into containers. In cooperation with the Bioconda project, BioContainers had created 69,175 containers by the end of 2021, including availability in different flavours. This relates to 18 TB of containers used by thousands of scientists worldwide. Workflow management systems like Galaxy and Nextflow use these containers to orchestrate workflows. Docker and OCI-compatible images and

singularity images are available²³

More than 1300 Conda packages have been annotated with a bio.tools ID and SPDX licence identifiers and *Ten simple rules for making a software tool workflow ready* was published²⁴. Members of the Tools Platform worked on RO-Crate, a packaging format for Research Objects used as a key mechanism for exchanging, archiving, publishing and citing workflows and data. A recent publication from members of the Platform introduced this open, community-driven and lightweight approach to packaging research artefacts along with their metadata in a machine readable manner²⁵.

²³ <https://quay.io>, <https://depot.galaxyproject.org/singularity>

²⁴ Brack, Paul, et al. 10 Simple Rules for making a software tool workflow-ready. PLoS computational biology (2022)

²⁵ Soiland-Reyes, Stian, et al. Packaging research artefacts with RO-Crate. Data Science Preprint (2021): 1-42

The EDAM ontology enhances the findability of bioinformatics resources, including software, and has become a crucial component of multiple services. Two Outreachy interns worked on EDAM, EDAM Browser and EDAM Popovers over the summer of 2021, adding major enhancements.

Work on the bio.tools registry has resulted in three published articles: a description of the biotoolsSchema data model²⁶, a summary of the current state of proteomics software²⁷ in bio.tools and a discussion on automated composition of workflows in life sciences²⁸. Community pages²⁹ have been produced within bio.tools to help navigation and promote the discovery of community content. ELIXIR Communities (for example, Rare Diseases, Intrinsically Disordered Proteins and 3D-BioInfo) are active in contributing to the bio.tools registry. The number of entries increased in 2021 to over 23,000 and the number of monthly visits by users in excess of 35,000. Bio.tools content is also available in JSON-LD

format using the ComputationalTool Bioschemas profile³⁰.

Project leads have welcomed 10 Communities to OpenEBench, assisting each to initiate the benchmarking process. They also took part in the APAeval challenge³¹ to benchmark 18 open-source, poly(A)-site-specific computational tools offering users a basis for choosing the most appropriate tools. OpenEBench has been registered in the EOSC marketplace³², published data into b2share³³ and is a use case in the EOSC-synergy project. There have been substantial updates around the interface design of OpenEBench along with improvements to documentation³⁴.

The Software Best Practices group delivered the first version of the software management plan (SMP), published in BioHackrXiv³⁵, developed during the 2019 and 2020 BioHackathons. Work continues on a machine-actionable version of the SMP, enabling data to be collected

automatically and exchanged with registries and repositories related to research software. The group has also continued engagement with the RDA FAIR for Research Software working group (FAIR4RS)³⁶ and has engaged the research software community by presenting the 4OSS lessons³⁷ and SMP at six different conferences.

Finally, in June 2021, the Tools Platform Ecosystem was funded as an infrastructure service. The kick-off meeting covered the ecosystem governance model, community engagement, ecosystem accessibility beyond API calls, and open and transparent integration with other services and communities within and beyond ELIXIR. A development version of the metadata repository is available and is periodically synchronised with bio.tools, OpenEBench, BioContainers and other community resources. These metadata are now reused by major Galaxy instances³⁸ to enrich tool descriptions and improve annotations using EDAM.

26 Ison, Jon, et al. biotoolsSchema: a formalized schema for bioinformatics software description. *GigaScience* 10.1 (2021)

27 Schwammle, Veit, et al. Proteomics Software in bio.tools: Coverage and Annotations. *Journal of proteome research* 20.4 (2021)

28 Lamprecht, Anna-Lena, et al. Perspectives on automated composition of workflows in the life sciences. *F1000Research* 10 (2021)

29 <https://bio.tools/communities>.

30 <https://bioschemas.org/profiles/ComputationalTool/1.0-RELEASE>.

31 <https://irnacosi.org/2021/07/06/the-apeval-challenge-status-update-june-2021>.

32 <https://marketplace.eosc-portal.eu/services/openebench>.

33 <https://b2share.eudat.eu/communities/OpenEBench>.

34 <https://openebench.readthedocs.io/en/latest>.

35 <https://biohackrxiv.org/k8znb>.

36 <https://www.rd-alliance.org/groups/fair-research-software-fair4rs-wg>.

37 <https://softdev4research.github.io/4OSS-lesson>.

38 <https://usegalaxy.eu>, <https://usegalaxy.org>.



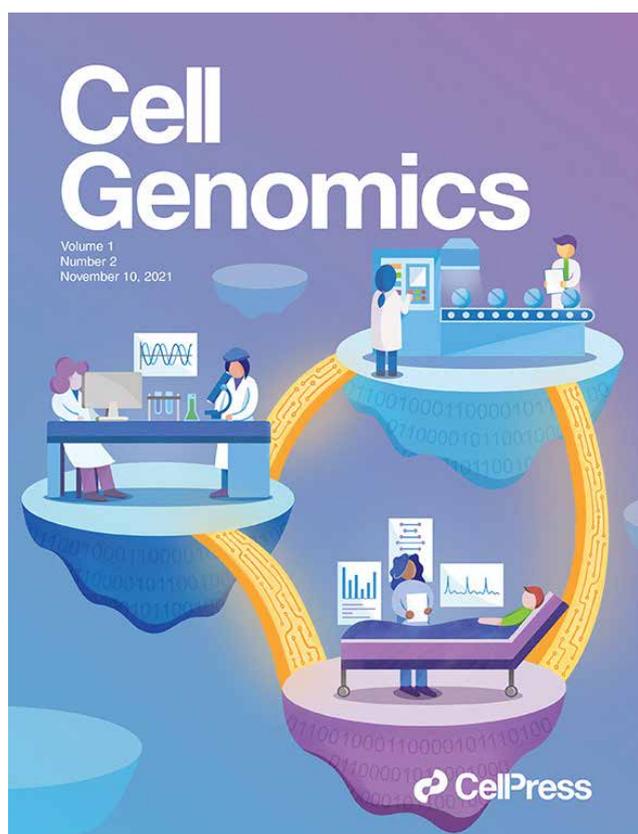
Compute Platform

The ELIXIR AAI Infrastructure Service continued to provide a sustainable AAI service including day-to-day operations and support. It has coordinated closely with the CINECA (human data) and EOSC-Life EU-funded projects in particular. There were 6851 active ELIXIR identities by the end of 2021 (doubled since 2019) and ELIXIR AAI was deployed in 131 Production Relying Services (and 133 in testing). Availability of services was 99.868% with 8530 logins/month, with live statistic monitoring³⁹.

The Compute Platform continued to develop additional multi-factor authentication services and interfaces. Through EOSC-Life and partners, the Platform has co-lead the development and implementation of the Life Science AAI, which will become known to users as Life Science Login⁴⁰, preceding the migration of all ELIXIR AAI users during early 2022. Platform members also presented online training on GA4GH Passports⁴¹ and Relying Services⁴² and was featured in multiple high profile presentations and a Cell Genomics refereed article⁴³. All training materials were made available in ELIXIR's training portal, TeSS.

The Compute Platform addressed the provision of more consistent, standardised and secure dataset access methods. In particular considerable contributions were made to the development of the GA4GH Data Repository Service (DRS) standard, now in version 1.1⁴⁴. The project also trialled and deployed high speed data transfer mechanisms, for example, IBM Aspera is now available at EMBL-EBI for users to

The GA4GH-inspired cover of a Cell Genomics Special Issue (November 2021), including papers by individuals involved in the ELIXIR Compute Platform who contributed to GA4GH, AAI and FASP. Image credit: Stephanie Li, GA4GH Communications.



upload their contributed datasets⁴⁵. Work on data access security issues continues in the GA4GH framework involving technologies such as GA4GH Passports⁴⁶. Work on a DRS API version that incorporates provision for access to sensitive data is ongoing.

Through the Hybrid Cloud Task, the Platform continued to coordinate technical, operational and funding aspects of cloud, data and compute services across Europe. The work

focused on federating the private clouds (EMBL-EBI Embassy) and national community clouds (cPouta at ELIXIR FI, de.NBI at ELIXIR DE or MetaCentrum Cloud at ELIXIR CZ) with the public/commercial clouds (AWS, Google Cloud) into EOSC infrastructure. A metagenomics demonstrator has been implemented within EOSC-Life by Bielefeld University and EMBL-EBI. As a first hybrid cloud prototype the cloud-based

Strategic Objectives

1

4

39 <https://login.elixir-czech.org/statistics>.

40 <https://lifescience-ri.eu/ls-login.html>.

41 <https://docs.google.com/document/d/1oZdW3TsAv2rd4uAJdh6gP3jTbmLuhfLUTCrE4xopGMc/edit>.

42 <https://docs.google.com/document/d/1wodZegZN2KfKrs8GXnitDo7bRWCh7-A-JdJaO6UvVZ8/edit>.

43 Voisin, Craig, et al. GA4GH Passport standard for digital identity and access permissions. Cell Genomics 1.2 (2021)

44 <https://ga4gh.github.io/data-repository-service-schemas/preview/release/drs-1.1.0/docs>.

45 https://www.ebi.ac.uk/arrayexpress/help/ftp_upload.html.

46 <https://www.ga4gh.org/ga4gh-passports/>.



workflow has been moved between different sites of the de.NBI Cloud, demonstrating easy deployment on different infrastructures.

Another problem routinely faced by bioinformaticians is the robust, reproducible analysis of large volumes of data. The Compute Platform has addressed this through the ELIXIR Cloud & AAI initiative⁴⁷, a Driver Project of the Global Alliance for Genomics and Health. The initiative works together with the global GA4GH community to improve GA4GH Cloud API standards, develops Proof of Concept implementations to address real-world use cases, promotes standards within and beyond the ELIXIR community and conducts interoperability tests with

other implementers in academia and industry. These efforts are combined in the establishment of the ELIXIR::GA4GH Cloud, a loosely coupled, federated FAIR compute infrastructure for the execution of containerized data analysis workflows in the cloud, based on GA4GH Cloud API standards. After successfully showcasing the execution of a demonstrator workflow making use of all four cloud API standards at the end of 2020, the Compute Platform consolidated the ELIXIR::GA4GH Cloud during 2021. This included co-leading the approval of the Task Execution Service (TES) API v1.0⁴⁸ and by developing and extending services towards a planned limited public release of the environment to end users during 2022.

⁴⁷ <https://github.com/elixir-cloud-aa/elixir-cloud-aa>.

⁴⁸ <https://www.ga4gh.org/news/ga4gh-tes-api-bringing-compatibility-to-task-execution-across-hpc-systems-the-cloud-and-beyond>.



Training Platform

Strategic Objective

1

Ten simple rules for making training materials FAIR



Adapted illustration from Luc Wiegiers and Celia van Gelder (<https://doi.org/10.5281/zenodo.3593257>) published in the article 'Ten simple rules for making training materials FAIR' in PLoS Computational Biology.

Following the strategy outlined in 2020 to address the global pandemic, the Training Platform focussed on delivering online training events throughout 2021. The Platform delivered over 130 events in 2021 through individual Nodes, with most as fully virtual and only a few as blended or face-to-face. TeSS⁴⁹ now lists over 12,277 training events (since 2008) with 1,583 training materials publicly available.

The activities of the Training Platform and training and capacity building Work Package of ELIXIR-CONVERGE

are closely aligned. In 2021, ELIXIR-CONVERGE WP2 delivered 37 courses (involving 1359 participants and 14 Nodes) related to data management and stewardship.

The second round of analysis of training needs and gaps for all ELIXIR Platforms and Communities was completed, enabling a longitudinal assessment of evolving needs. A number of hackathon events were organised in 2021 with the goal of producing new training material on topics such as machine learning, high performance computing, containers

and workflows. Of particular note was the hackathon on developing FAIR training materials organised in the context of the ELIXIR-GOBLET partnership and the ELIXIR Train-the-Trainer project, which featured in the 2021 ELIXIR BioHackathon. Progress was made on Learning Pathways, including a BioHackathon Europe 2021 project supported by CONVERGE.

An important highlight of 2021 was Train-the-Trainer (TtT) activity, receiving particular commendation by the ELIXIR Scientific Advisory Board. Recognising the critical role of

⁴⁹ <https://tess.elixir-europe.org>.

TtT in capacity building, a dedicated ELIXIR TtT Instructor community was established, supporting discussions on improving training material, such as the planned TtT material repository, and scaling up activities. The Training Platform also established a collaboration with Australia BioCommons to replicate TtT courses in Australia and Asia. Finally, the GOBLET-ELIXIR Professional Guide, co-authored by members of the Training Platform, published *Course*

*design: Considerations for trainers – a Professional Guide*⁵⁰.

During 2021, TeSS was established as an ELIXIR Infrastructure Service, complementing existing adoption by two key EOSC infrastructures (EGI Training Catalogue⁵¹ and PANOSC Training Platform⁵²) and the Digital Research Skills Australasia (DReSa)⁵³ portal. To provide support, the TeSS Club⁵⁴ was created enabling the TeSS development team to

prioritise technical developments based on community feedback. Connections were also made from the CONVERGE RDMkit to TeSS.

Beyond the ELIXIR network, the Training Platform is actively involved in the EOSC Advisory Group on Research (careers and curricula) through the involvement of Training Platform members across three Task Forces.

50 Via, Allegra, et al. Course design: Considerations for trainers—a Professional Guide. F1000Research 9.1377 (2020)

51 <https://catalogue.hopto.org>.

52 <https://pan-training.hzdr.de>.

53 <https://dresa.org.au>.

54 <https://elixirtess.github.io/about>

ELIXIR Communities

Bringing together experts from within particular life sciences domains, ELIXIR Communities develop targeted standards, services and training. They capture specific research needs from across ELIXIR Nodes and partner organisations and translate them into formal requirements to drive the portfolio of services in the ELIXIR Platforms. These strong internal ties ensure that those services developed are fit-for-purpose and serve real research community needs.

ELIXIR Communities provide a mechanism to reach out to defined groups of experts including other research infrastructures of the ESFRI roadmap⁵⁵. In turn, ELIXIR provides Communities with a formal, well-defined structure with access to funding opportunities. They can participate in Community-led Implementation Studies and members of Communities often use the networking opportunities to develop proposals for EU funding.

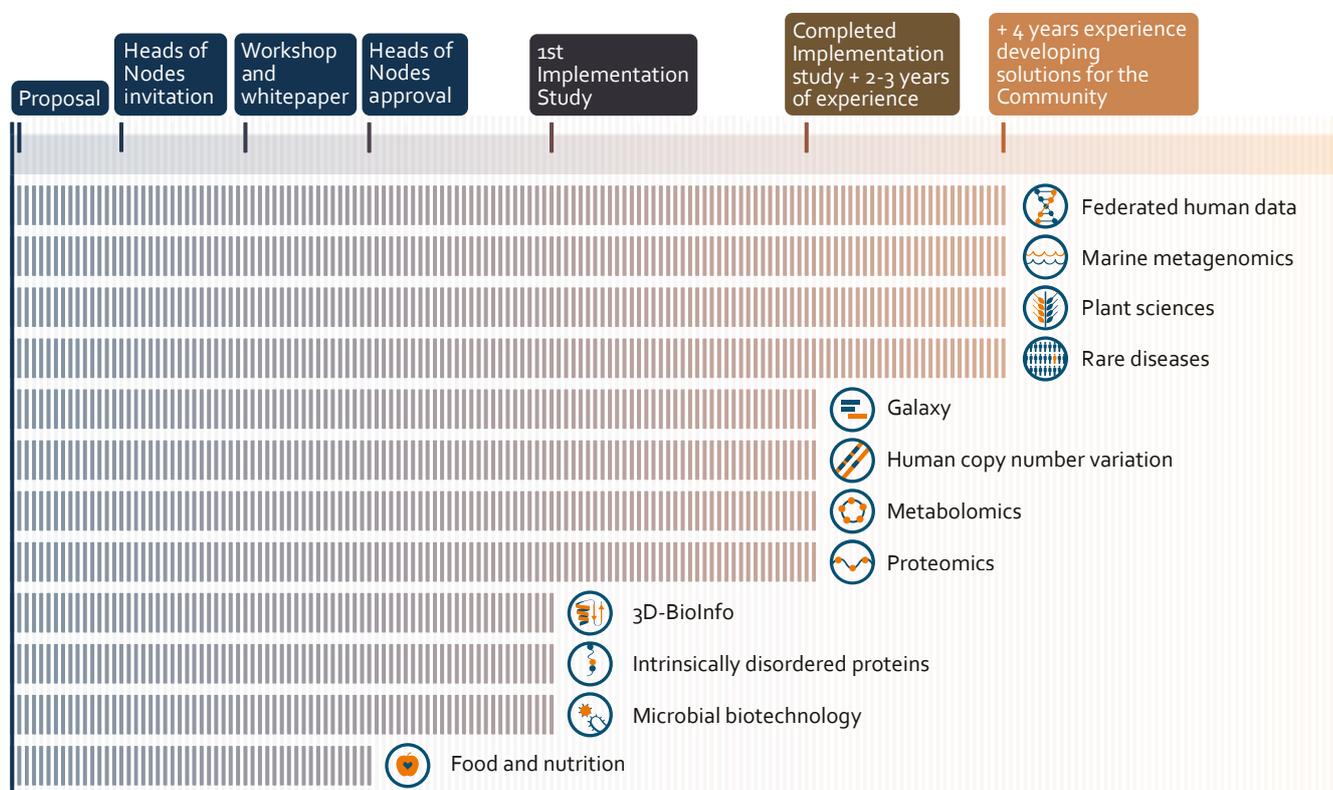
In 2021, there were twelve ELIXIR Communities, with the thirteenth - the Toxicology Community - selected towards the end of the year. There are two emerging Communities - Systems Biology and Single Cell Omics. The selection of new Communities is based on a well-defined process:

1. **Community proposal** - the new Community submits a formal proposal.
2. **Heads of Nodes invitation** - the Heads of Nodes Committee agree to the Community proposal.
3. **Workshop to establish white paper** - the nascent Community develops a white paper to establish the Community's infrastructure needs, describe the roadmap, and make the necessary connections with ELIXIR Platforms.
4. **Community establishment** - the Heads of Nodes Committee agree to the white paper, which has been widely consulted across ELIXIR.

Once a Community is established, it receives funding through an initial Implementation Study to kick-start technical developments and community-building activities. Over time, with annual meetings and collaboration with other external initiatives, ELIXIR Communities grow and mature. Maturity can be described by three levels:

1. Recently selected Communities who are about to or have just begun their first Implementation Study.
2. Communities recently finishing their first Implementation Study that have been implementing the Community roadmap for two-three years.
3. Communities with over four years of work developing standards or technical solutions to meet the Community's needs.

The maturity journey of the ELIXIR Communities



55 <http://roadmap2018.esfri.eu/projects-and-landmarks>



3D-Bioinfo

The ELIXIR 3D-Bioinfo Community aims to integrate protein structure-based data and tools and improve standardisation.

A major achievement in relation to structural/functional annotation was the launch of the 3D-Beacons platform in September. 3D-Beacons is an open collaboration to provide 3D-coordinates and meta-information for experimental and predicted protein models from international resources in a standardised data format. To model the interactome in 3D, the Community established a novel challenging benchmark dataset of physiological and non-physiological homodimers and compiled a list of tools for evaluating protein interfaces. The performance of these tools and of AlphaFold2 (the structure prediction engine by Google DeepMind) in discriminating between the two dimer categories was evaluated.

Work in the area of protein-ligand interactions included the development of a computational pipeline using NextFlow for generating and updating a benchmark dataset to evaluate prediction tools. In the domain of protein-nucleic acid interactions, a panel of 25 experts reviewed the use of nucleic acid valence geometries to avoid conflicting reports and worked with the organisers of an RNA modelling competition, RNAPuzzles, to develop ways of evaluating RNA model quality. Finally, the area of protein engineering continues to make progress capturing experimental data in a standard format in BioStudies at EMBL-EBI.

Over 500 people registered for the AGM in September with all sessions well attended and the event disseminated in a news release⁵⁶.



Food and Nutrition

The Food and Nutrition Community was formally recognised in December 2021, and planning for the first Community Implementation Study, *Food & Nutrition: Microbiome - Diet - Health*⁵⁷ undertaken. During 2021 all Community co-leads were selected and the Community white paper was drafted for publication in F1000.



Galaxy

In 2021, the ELIXIR Galaxy Community continued supporting SARS-CoV-2 data analysis, shifting the focus to surveillance, monitoring and dissemination, for example by running several dedicated webinars and workshops. Capacity building through massive open online courses proved very successful, creating new e-training materials in different formats, targeting different audiences (system administrators, developers, scientists and educators) and reaching out to thousands of attendees worldwide. The Community has continued to run well-attended ELIXIR webinars (10 held in 2021).

Over the past year, more than 2,000 tools have been updated and added to the Galaxy ecosystem, making them available to deployers and users. In particular, several new assembly and genome annotation tools and workflows were developed in collaboration with the ERGA and VGP projects, which have already yielded new high-quality reference genomes. The Galaxy ecosystem is growing, both in the number of users (which almost doubled in 2021) and in the number of new domains, for example pathology, machine learning, biodiversity and climate science. A new Implementation Study started in the second half of 2021 to strengthen the data management capabilities of Galaxy, bringing together partners from 19 institutes and 12 ELIXIR Nodes.

⁵⁶ <https://elixir-europe.org/3DBioInfo-AGM>.

⁵⁷ <https://elixir-europe.org/2022-FoodandNutrition>.



Intrinsicly Disordered Proteins

Intrinsicly Disordered Proteins (IDPs) are characterised by exceptional structural heterogeneity and dynamic behaviour which can be studied by a range of complementary experimental techniques. The ELIXIR IDP Community supports the development of standards, tools and resources to accelerate the identification, analysis and functional characterisation of IDPs. The Community has been active in recruiting new members in this diverse field and started two Implementation Studies in 2021.

The first Implementation Study is *Standardising IDP data*. The main activities are the creation of the IDPcentral hub for data dissemination and the development of formats for standardised data transfer. IDPcentral exploits the Minimum Information About Disordered Experiments standard, which captures experimental details, and Bioschemas, which has been extended to represent positional annotations. The Gene Ontology and the Evidence and Conclusion Ontology have been extended, and the DisProt, MobiDB and PED databases were connected to IDPcentral.

The second Implementation Study is *Improving IDP tool interoperability and integration into ELIXIR*. The CAID benchmarking is being integrated into OpenEBench and the best disorder predictors are being ported into Galaxy. A collaborative project with the 3D-Bioinfo and Proteomics communities is underway to collect tools useful to all Communities and define protein structural states and their ambiguity.



Marine Metagenomics

During 2021, the Marine Metagenomics Community has continued to develop tools, and improve portals and reference databases including the FARification of metagenomic data.

Highlights include contributions to the developed RDMkit tool assembly for the Marine Metagenomics Community and developing a knowledge model for marine metagenomics in Data Stewardship Wizard.

In collaboration with the Biodiversity Focus Group, the Community has co-authored a white paper on Biodiversity, published in the ELIXIR F1000 channel⁵⁸. The Marine Metagenomics Portal has been upgraded to include updates to the MAR databases.



Metabolomics

The Metabolomics Community continued to work with experimental scientists and developers to provide the resources, analysis tools and infrastructure to help metabolite identification. The focus in 2021 was fluxomics.

The 1st ELIXIR Fluxomics Training School ran virtually in October. The course was designed and organised through two hackathons with Implementation Study partners in February and March. A Zenodo space⁵⁹ and YouTube channel⁶⁰ were created to include the programme, and training materials and videos from the lectures and the invited talk. Trainers from 14 Institutions in Europe and the US were invited and MATLAB sponsored the event by providing free access to online MATLAB to all 22 students from 9 European countries and the US. Comments were positive and the Community plans to continue offering the course every two years.

All fluxomics-related bio.tools entries have been included in the fluxomics space recently added to EDAM ontology and bio.tools. A standardised isotopic tracer deposition protocol was developed and initial steps for implementation in the ISA-enabled MetaboLights database have been taken.

⁵⁸ Waterhouse, Robert M., et al. Recommendations for connecting molecular sequence and biodiversity research infrastructures through ELIXIR. F1000Research 10.1238 (2021).

⁵⁹ <https://zenodo.org/communities/fluxomics?page=1&size=20>.

⁶⁰ https://www.youtube.com/channel/UCVsr_7wQL5uk7tiPaDGK2mA.



Microbial Biotechnology

The Community aims to foster a knowledge-based infrastructure for microbial biotechnology. In 2021, it focused on addressing standardisation and interoperability of metabolic models, enzyme representation and semantic ontologies.

The Community played a major role in establishing the RDMkit for microbial biotechnology along the DESIGN-BUILD-TEST-LEARN cycle. RDMkit for microbial biotechnology was the focus of participation in the BioHackathon Europe 2021.

Progress was made on the Community Implementation Study, advancing the specification and implementation of the integration of enzyme-related measurements, provenance and the formal representation of enzymes and metabolic pathways. The model representation of genome-scale metabolic networks were advanced, laying the ground for more accurate semi-automated reconstructions.

There was substantial progress on overcoming bottlenecks in metabolic modelling related to namespaces (published in *Metabolites*⁶¹) and gap-filling (manuscript in preparation). The Community opinion paper has been updated for publication in 2022 and a paper on data management strategies for microbial biotechnology and synthetic biology is being prepared.

Links have been made with other Communities (Systems Biology, Metabolomics, Microbiome, Plant Sciences) on the use of genome-scale metabolic models through a joint workshop at the ELIXIR All Hands meeting 2022.



Plant Sciences

The Plant Sciences Community is an interdisciplinary group composed of experimentalists, biologists and computer scientists. Motivated primarily by the desire to enable exchange, integration, retrieval and long-term storage of plant data, the Community has been involved in several activities on tools and standards during 2021.

The FONDUE Implementation Study was completed in November 2021, with training materials developed and widely disseminated⁶². The Implementation Study, *Increasing plant data findability and reuse beyond ELIXIR*, started in 2021 and made progress on defining tools and their data models for three upcoming service bundles.

A project at the BioHackathon 2021, led by the Plant Community, improved the integration between MIAPPE, ISA and Knowledge Graphs with a connection to the Bioschemas Implementation Study. A contribution was made to the BrAPI Hackathon in October for developing the next version. In June, an online datathon for the submission of FAIR data for plant genotyping was organised by the Community. A tools assembly describing plant genomic and genotypic data has been added to the RDMkit.

An opinion paper on VCF specifications and best practices on data management was submitted to F1000Research, and FONDUE is a use case in a publication on BioSamples. The Plant Community engaged with the international DivSeek initiative and discussed possible future collaborations and strategic orientations.



Proteomics

The Community published a white paper on current practices, recommendations and perspectives on the management of sensitive (clinical) human proteomics data. This is a topic of increasing importance in the field and multiple studies have been published in the last two years. The white paper includes several recommendations, including the development of bioinformatics infrastructure with controlled access capabilities for this type of datasets.

The MAGE-TAB-Proteomics data standard defining experimental design and sample metadata in public proteomics experiments was finalised. The data standard was developed using as the basis the MAGE-TAB format used in transcriptomics. Initial tooling was also developed, with PRIDE supporting the submission of the Sample and Data Relationship Format component of MAGE-TAB-Proteomics. All information is available in a Nature Communications publication⁶³.

Several open data analysis pipelines were developed in NextFlow⁶⁴, covering some of the most popular analysis software in proteomics such as MaxQuant, Compomics, OpenMS and the Trans Proteomic Pipeline.

61 Pham, Nhung, et al. Consistency, inconsistency, and ambiguity of metabolite names in biochemical databases used for genome-scale metabolic modelling. *Metabolites* 9.2 (2019).

62 <https://github.com/PBR/elixir-fondue-datathon>.

63 Dai, Chengxin, et al. "A proteomics sample metadata representation for multiomics integration and big data analysis." *Nature communications* 12.1 (2021)

64 <https://github.com/wombat-p>.



Toxicology

In September 2021, Toxicology was formally accepted by the Head of Nodes Committee as an official ELIXIR Community. The Community white paper has been submitted to F1000⁶⁵ and is awaiting peer review. During 2021, the Bioschemas extension for chemical substances (for example, nanomaterials) was approved in schema.org⁶⁶ and the GitHub repository was migrated to the ELIXIR organisation in GitHub⁶⁷.

Human Data Communities

The ELIXIR Human Data Communities (HDCs) include three Communities (Federated Human Data, Rare Diseases and human Copy Number Variation), two Focus Groups (Health Data and Cancer) and activities around the ELIXIR Beacon Commissioned Service.

Cross-cutting themes were discussed in monthly meetings to progress the ambition of operating a sustainable infrastructure for Human Genomics and Translational data in Europe to support life science research and its translation to medicine. To reach this goal, experts in the Communities have been successful in participating in a number of key EU funded programmes. The Beyond 1M Genomes project (B1MG) is one example, where experts from all Communities are supporting the 1+MG initiative by producing a 1+MG Framework and Proof of Concepts in rare diseases and infectious diseases (with links to ELIXIR-CONVERGE). The 1+MG Framework contains guidelines, recommendations and best practices across the domains of ELSI, data standards, quality and federated secure cross-order technical infrastructure. Other cross-cutting themes include discussing the development and implementation of genomic standards via the strategic partnership between ELIXIR and GA4GH.

65 Martens, Marvin, et al. ELIXIR and Toxicology: a community in development. F1000Research 10.1129 (2021).

66 <https://elixir-europe.org/news/towards-better-findability-bioschemas-meets-schemaorg>.

67 <https://github.com/elixir-europe/toxicology-community>



ELIXIR Federated Human Data Community

The main focus of the Community has been continued support of ELIXIR Nodes to make national human genomic data accessible for secondary research. Support has been provided to Nodes at different stages of their national human data implementation strategies.

The establishment of the Federated European Genome-phenome Archive (EGA) has been of continued importance and accelerated with the start of a human data Work Package in ELIXIR-CONVERGE. The contractual service agreement with the five first wave Nodes has been finalised (Finland, Germany, Norway, Spain and Sweden), and the first end-to-end service demonstrators with the most advanced Nodes completed (Norway, Finland and Sweden).

Regular meetings of the EGA Strategic and Operational Committees have been established and the Community has engaged with upcoming Nodes via the Federated Human Data teleconference and Use Case Survey to better understand use cases and individual needs. The first version of the federated EGA maturity model has been released and input is being actively sought to identify areas requiring clarity and further work to ensure upcoming Nodes are provided with a comprehensive roadmap for developing human data sharing capacity, expertise and services.



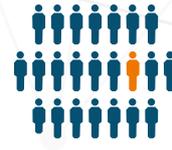
ELIXIR Rare Diseases Community

The Rare Disease Community Infrastructure Implementation Study was finalised, enabling visualisation in the RD-Connect Genome-Phenome Analysis Platform of nearly 10,000 genomic alignments archived at the EGA through GA4GH htsget. This has been key in enabling the H2020 Solve-RD project to expand to include the visualisation of RNA-Seq data.

FAIR metrics for Rare Diseases has been tested, the Proof of Concept launched and progress made to visualise RNA-Seq data processed with the DROP pipeline on transSMART. FAIRification training needs have been surveyed using the ELIXIR Training Platform.

There have been further development of three ELIXIR Service Collections: assessing molecular pathogenicity for rare diseases, systems biology service bundles for rare diseases and the ELIXIR training ground for rare diseases. Close collaboration with the Horizon 2020-funded European Joint Programme on Rare Diseases (EJP RD) has continued, with several ELIXIR resources and tools developed and adapted.

Members of the Community participated in the ELIXIR All Hands meeting, contributing to the Interoperability Platform session on the convergence of FAIRification approaches and RDMToolkit. Contributions have been made to Bioschemas, markup of Orphanet, and preparations for DCAT to Bioschemas mapping. The strategic partnership has continued with the Global Alliance for Genomics and Health, with the EJP RD as a driver project, including Beacon v2, data access models and semantic phenopackets.



ELIXIR human Copy Number Variation

The first Community Implementation Study was completed and two new studies were approved and initiated: *Reference hCNV datasets, use-case workflows and benchmarking*, and *Beacon and beyond - implementation-driven standards and protocols for CNV discovery and data exchange*.

A BioHackathon Europe 2021 project on hCNV detection tools benchmarking and training material was organised and the Community participated in an EOSC Pillar use case definition on Institutional Data repository and Galaxy infrastructure for CNV.

The Community collaborated with ELIXIR Beacon in the Variant Query Scout Team and engaged with other ELIXIR Communities, Platforms and Services, for example, with the Tools Platform through activities in bio.tools, SIS biocontainers, benchmarking, Beacon and Galaxy.

ELIXIR Cancer Data Focus Group

Since being established at the start of 2020, the Cancer Data Focus Group has gathered experts from seventeen different Nodes across ELIXIR and identified ten use cases, along with other challenges in cancer data management in bioinformatics infrastructures.

In 2021, experts from the Group came together to contribute to two HORIZON-INFRA-2021-EOSC-01 Research and Innovation Action proposals: *A European-wide foundation to accelerate data-driven cancer research* and *Providing cutting edge cancer research services across Europe*.

The Group continues to hold monthly meetings and is working towards a white paper addressing the challenges in cancer data analysis in research and clinical practice.

ELIXIR Health Data Focus Group

Convened in February 2020, the Health Data Focus Group was formed by members across eight ELIXIR Nodes and three external organisations. During 2021, the Group identified common challenges facing the secondary use of health data for research and policy-making purposes.

The Group facilitated a successful knowledge exchange workshop, during which 10 ELIXIR Nodes with established health data initiatives, outlined their successes, challenges and learnings on the re-use of health data.

The Group also participated in the ELIXIR Data and Interoperability Platform meeting and led an informative panel discussion, providing an opportunity for future collaborations across different ELIXIR structures.

ELIXIR Beacon project

The first Beacon Implementation Study was completed and four new studies approved and initiated: a new Beacon Implementation Study, Beacon Infrastructure, Galaxy and hCNV.

Using the Beacon Reference Implementation developed in the context of ELIXIR, Beacons are currently being deployed both in clinical settings (for example, the La Marató project in Catalunya) and by international cohorts (for example, the CINECA project gathering cohorts from Europe, Canada and Africa). The Beacon team provides webinars and training sessions to support external partners in the deployment of Beacon instances. A BioHackathon Europe 2021 project was organised and a Beacon team participated in (1) a prototype implementation of electronic health record data, and (2) the linking of Beacon to a semantic data model for rare diseases. An article describing Beacon has been accepted for publication in *Human Mutation*.

The ELIXIR Beacon team is tightly connected to the Global Alliance for Genomics and Health, and the members actively participated in the elaboration of the Beacon V2 standard (which the ELIXIR Beacon Reference Implementation derives from). The standard has been submitted in November 2021 and is now under review for approval.

Our flagship events

All Hands Meeting – connecting all ELIXIR members

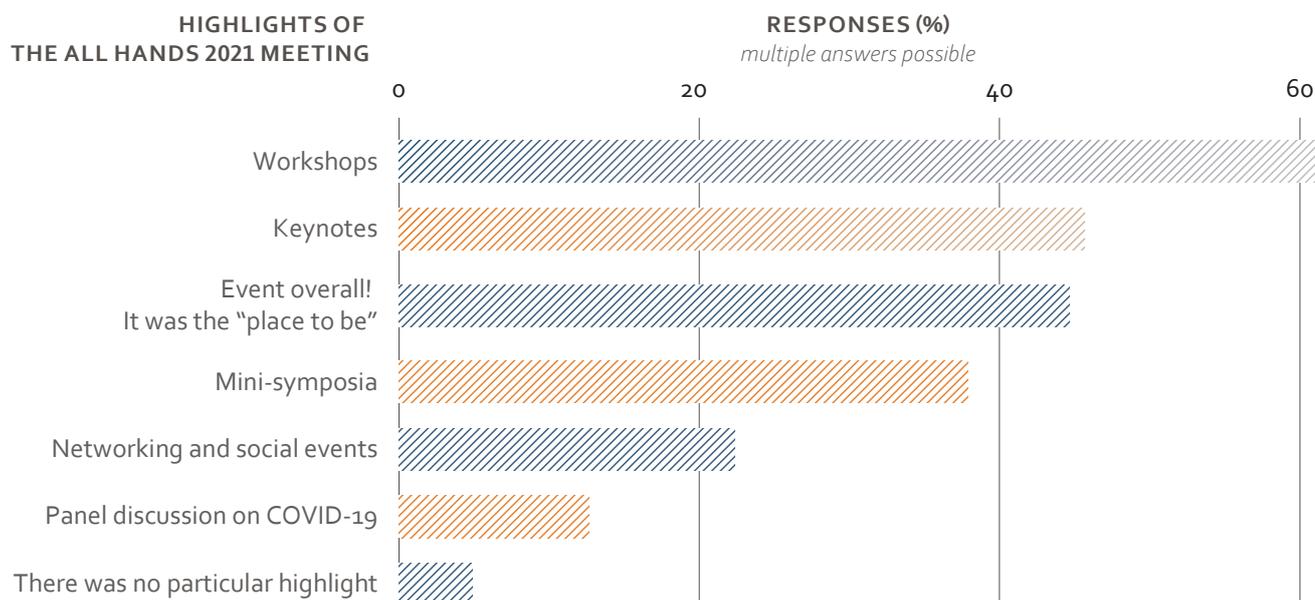
The ELIXIR All Hands meeting is an annual event that brings together the ELIXIR family to share updates, develop new plans and strengthen personal connections.

Due to the uncertainty around the COVID-19 pandemic, ELIXIR All Hands 2021 was planned and run as a fully virtual event. This enabled the event to welcome in excess of 500 participants, including attendees from across the world.

Taking place over five days, the event built community relations through a diverse programme of workshops and mini-symposia focused on emerging and established topics across a range of ELIXIR’s activities. Social events and virtual support enabled a community feel around the event despite the limitations of a virtual environment.

500
PARTICIPANTS

Survey responses from the All Hands Meeting 2021



BioHackathon Europe – ELIXIR’s foray in hybrid hacking

Following the fully virtual and highly successful BioHackathon Europe in 2020, a hybrid event was planned for 2021 to safeguard against the ever-changing circumstances of the pandemic. Held from 8 - 12 November in Barcelona, Spain, ELIXIR welcomed 154 face-to-face and 170 virtual participants to the event. With participants working collaboratively over tables, screens

and time zones on 37 projects, new data, features and friendships were moulded over the course of an intensive week of hacking.

Projects were submitted from teams and individuals from both academic and industry institutes and covered a diverse range of topics from investigating demographics and biases in scientific research to

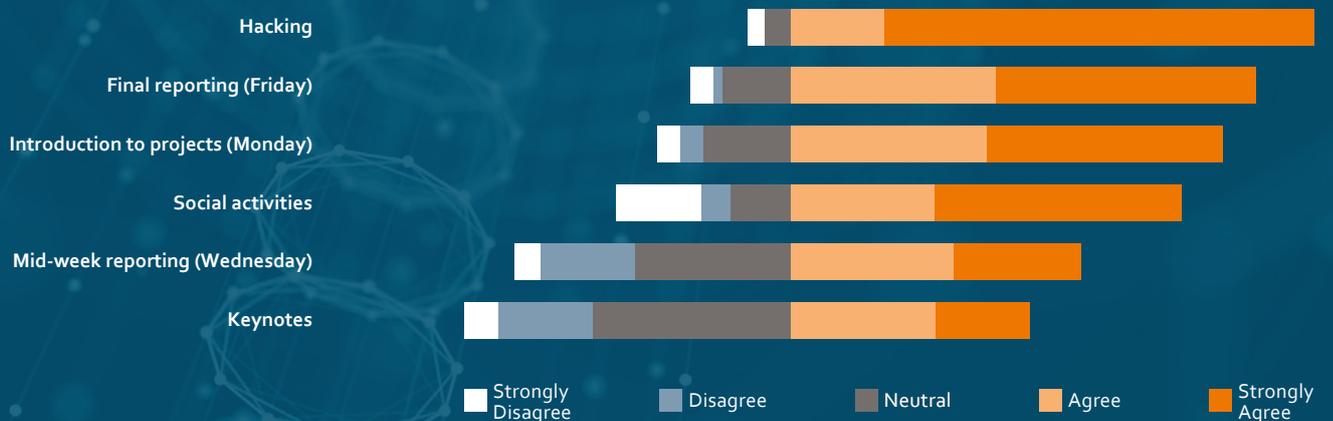
creating a new ontology for wine tasting. Several new features for services and training materials were developed and advanced during the course of the event.

The event was well received, with approximately 97% of surveyed participants stating that they would recommend this event to their colleagues.

Survey responses from BioHackathon Europe 2021



THE FOLLOWING PARTS OF THE PROGRAMME WERE PARTICULARLY USEFUL/GOOD



Our collaboration, outreach and industry support efforts

Strategic
Objective

3

Strengthening global collaboration

ELIXIR-GA4GH strategic partnership



Global Alliance
for Genomics & Health
Collaborate. Innovate. Accelerate.

ELIXIR continued its mutually beneficial collaboration with the Global Alliance for Genomics and Health (GA4GH) throughout 2021. GA4GH is global in scope but given the challenges in building relationships with countries individually, ELIXIR coordinates the pan-European activities they have in common. In addition, experts from across the ELIXIR Nodes provide essential technical and strategic input into the GA4GH standard development, ensuring that they are suitable for European implementation, including via the use of ELIXIR resources such as Beacons.

In 2021, GA4GH-approved standards were included within workflows established to access data across borders, notably as part of the flagship Beyond One Million Genome project. That same year, a Special Issue on responsible sharing of biomedical data was published in the research journal *Cell Genomics*⁶⁸, in which ELIXIR's contribution was considerable, for example in the articles *GA4GH: International policies and standards for data sharing across genomic research and healthcare*⁶⁹ and *The Data Use Ontology to streamline responsible access to human biomedical datasets*⁷⁰.

Collaboration strategy with the Australian BioCommons



Australian BioCommons

This year has seen early stage outcomes of the implementation of ELIXIR's Collaboration Strategy with the Australian BioCommons, ELIXIR's sibling research infrastructure. The collaboration was formed to enable the coordination of mutually-beneficial activities and identify a framework for long-term cooperation.

Despite the geographic and time-zone challenge and the difficulties of travel, interactions have been productive, friendly and insightful. A jointly convened session was held at the ELIXIR All Hands meeting and there was a strong ELIXIR presence at the Australian BioCommons equivalent. Interactions to date have covered a wide breadth and depth of topics, including tools, training, Galaxy and impact evaluation.

Benefits to ELIXIR from the collaboration are significant, and include knowledge exchange and improving the targeting and reach of resources to a broader range of users.

Global Biodata Coalition collaboration



GLOBAL BIODATA COALITION

ELIXIR has continued to support the development of the Global Biodata Coalition (GBC, globalbiodata.org). The GBC was created to better coordinate and share approaches for efficient management and growth of biodata resources worldwide. It aims to ensure sustainable financial support for the global biodata infrastructure and offers prioritised long-term support to sustain a set of Global Core Biodata Resources (GCBRs)

In 2021, a series of documents was prepared describing the process for selecting the first set of GCBRs⁷¹. Selection of GCBRs is a natural progression of the ELIXIR Core Data Resources, offering research funders and researchers a better understanding of the global infrastructure of biodata resources. Niklas Blomberg will chair the GCBR review committee on behalf of the GBC.

68 <https://www.sciencedirect.com/journal/cell-genomics/vol/1/issue/2>

69 Rehm, Heidi L., et al. GA4GH: International policies and standards for data sharing across genomic research and healthcare. *Cell genomics* 1.2 (2021)

70 Lawson, Jonathan, et al. The Data Use Ontology to streamline responsible access to human biomedical datasets. *Cell genomics* 1.2 (2021)

71 Global Core Biodata Resources: Concept and Selection Process

Enabling and supporting industry collaboration

Impact report



Open access life science resources such as data and software are fundamental to breakthrough discoveries, scientific excellence and entrepreneurial endeavours. In 2021, we published a report, *Open data: A driving force for innovation in the life sciences*, to showcase the benefit of open data to industry and innovation. The report contains testimonials, an extensive survey and an analysis of the hotspots of European innovation.

ELIXIR surveyed over 50 bioinformatics companies to understand how they operate, which public resources they use in their business and how they contribute to the economy. Seventy-six percent of these companies state that their product or service would not exist without data in open repositories. Aside from examining the digital foundations of these companies, the report also looks toward the physical ecosystems that nurture new small and medium enterprises (SMEs). Close proximity to customers, large pools of highly qualified graduates, and existing infrastructures all make for attractive environments for SMEs to set up shop.

76%

STATED THAT WITHOUT DATA SHARED ON OPEN REPOSITORIES, THEY WOULD NOT BE ABLE TO OFFER THEIR PRODUCT OR SERVICES.

89%

STATED THAT A PRODUCT OR SERVICE HAS MORE FEATURES BECAUSE OF ACCESS TO SHARED OR OPEN REPOSITORIES.

63%

STATED THAT WITHOUT ACCESS TO REGISTRIES, ONTOLOGIES, AND DICTIONARIES PUBLISHED ON OPEN REPOSITORIES, THEY WOULD NOT BE ABLE TO OFFER THEIR PRODUCT OR SERVICE.

92%

STATED THAT A PRODUCT OR SERVICE HAS MORE FEATURES BECAUSE OF ACCESS TO REGISTRIES, ONTOLOGIES, AND DICTIONARIES SHARED ON OPEN REPOSITORIES.

Engaging with industry



ELIXIR support members to connect with industry experts through networking events such as the Innovation and SME Forum. Due to the COVID-19 pandemic many planned events were held virtually or as hybrid events. The ELIXIR Innovation and SME Forum was hosted by the Swiss Institute of Bioinformatics (SIB) at their [BC]₂ Basel Computational Biology Conference. The event explored the technological advancements in the field of personalised medicine with a focus on the learning ecosystem of health.

With the sustained travel restrictions and lockdowns of 2021, the event offered hybrid participation, bringing together bioinformaticians from Switzerland and beyond to learn and discuss the ecosystem of health. At the core of the event were short talks by start-ups and SMEs who develop products and services to connect the different components of the learning ecosystem, from treatment to tracking to translation. The companies included: BITAC (Spain), Gene Predictis (Switzerland), HealthECCO (Germany), HeartGenetics (Portugal), IQVIA (UK), Nebion AG (Switzerland) and Novartis (Switzerland). The event brought together more than 230 attendees with 80 participants making use of the virtual option.

Facilitating collaborations

ELIXIR supports collaborative partnerships between industry and ELIXIR members to drive innovation in bioinformatics. The ELIXIR Knowledge Exchange Scheme covers travel and accommodation expenses, as well as some contributions to ELIXIR members (individuals or teams from one institute within ELIXIR) to give them the opportunity to work on a collaborative project with an industry partner. Two projects were submitted to the scheme in 2021, which has been running as a pilot since 2019:

1. Health data anonymisation, synthetic data and pseudonymisation service (ELIXIR Finland). This project aims to validate the use of technologies of VEIL.AI (an SME) as part of the IT Center for Science's sensitive data solutions, and to formalise public infrastructure in partnership with an information and communications technology start-up.
2. Systems biology: bridging the gap to industry needs (ELIXIR Sweden). This project helps meet the needs of industry in the area of genome-scale metabolic models in Metabolic Atlas (a service of ELIXIR Sweden).

Demonstrating ELIXIR's impact

Publications

Using a text-mining approach with search terms relating to funding and ELIXIR-related achievements, 143 ELIXIR-supported publications were identified in 2021 (peer-reviewed articles and preprints). This continues an upward trend (increasing from 136 in 2020). The publications relate to the development and operation of bioinformatics resources from databases to training, highlighting ELIXIR's scientific legacy as a research infrastructure.

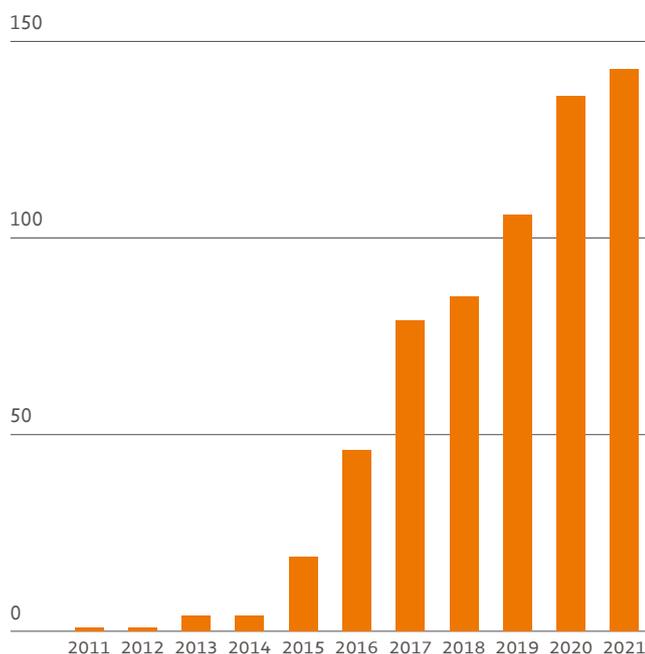
Publication highlights for 2021 include *A proteomics sample metadata representation for multiomics integration and big data analysis*⁷² (in Nature Communications), which acknowledges funding from an ELIXIR Implementation Study, and *Solving patients with rare diseases through programmatic reanalysis of genome-phenome data*⁷³ (in the European Journal of Human Genetics), also funded through ELIXIR as well as the European Union (via the European Joint Programme on Rare Diseases). ELIXIR's Machine

Learning Focus Group provided significant input in the development and publication of *Recommendations for supervised machine learning validation in biology*⁷⁴ (in Nature Methods), an area of growing interest across ELIXIR and beyond.

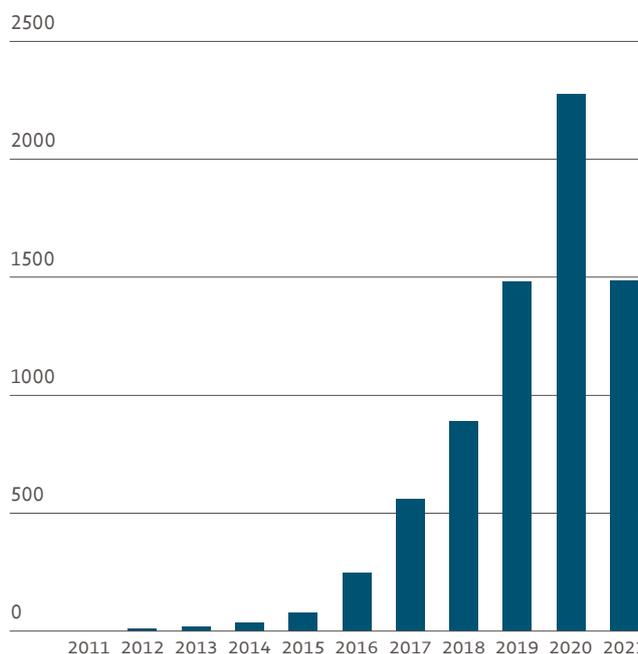
Since ELIXIR was conceived in 2007, ELIXIR-supported publications amount to 624, with 7091 citations (in open access journals), demonstrating the extent of ELIXIR's contribution and appreciation by others.

Impact of ELIXIR related funding (publications and citations)

PUBLICATIONS



CITATIONS



72 Dai, Chengxin, et al. A proteomics sample metadata representation for multiomics integration and big data analysis. Nature communications 12.1 (2021): 1-8

73 Matalonga, Leslie, et al. Solving patients with rare diseases through programmatic reanalysis of genome-phenome data. European Journal of Human Genetics 29.9 (2021)

74 Walsh, Ian, et al. DOME: recommendations for supervised machine learning validation in biology. Nature methods 18.10 (2021)

Stakeholder engagement

ELIXIR remains a highly visible and acknowledged research infrastructure in policy circles. This results from the provision of high quality public-funded bioinformatics services and work programmes and the efforts made across ELIXIR to showcase usefulness and impact beyond research. Key policy stakeholders targeted include a range of EU-related bodies (for example, the European Commission, the European Centre for Disease Control), the European Strategy Forum for Research Infrastructures (ESFRI), the Organisation for Economic Development and Cooperation (OECD), national governments and funding bodies, the G7 and the United Nations (the World Health Organisation and

UNESCO in particular).

Overton, the world's largest searchable index of policy documents (including guidelines, think tank publications and working papers), was used to complement manual tracking of sources mentioning ELIXIR projects, services and achievements in the policy sphere. Forty-five such sources were identified in 2021, continuing an upward trend exemplifying influence in policy and the public value of ELIXIR. Four OECD documents named ELIXIR, for example, *Building and sustaining collaborative platforms in genomics and biobanks for health innovation*⁷⁵ mentioned the distributed infrastructure, the European Genome-phenome Archive (an ELIXIR

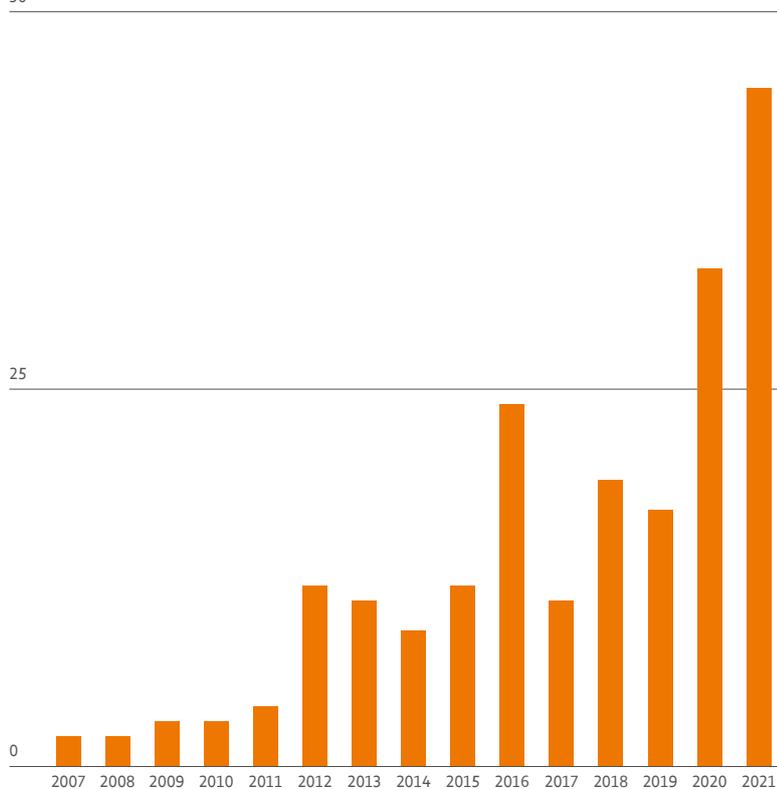
Core Data Resource), and the Strategic Partnership between ELIXIR and the Global Alliance for Genomics and Health.

Two guidance documents^{76,77}, for EU-funding grantees (Horizon Europe, European Research Council), recommended a range of ELIXIR resources for data and software management and interoperability in support of open science. Finally, the Swedish Research Council reconfirmed their positive assessment of ELIXIR membership in their report *National benefits from Swedish membership of international research infrastructures*⁷⁸.

Number of sources (e.g. policy documents, reports, guides for grantees, funding calls) and distribution of categories these sources, mentioning ELIXIR, its projects, resources and key achievements, as a reflection of ELIXIR's visibility and influence in the policy sphere.

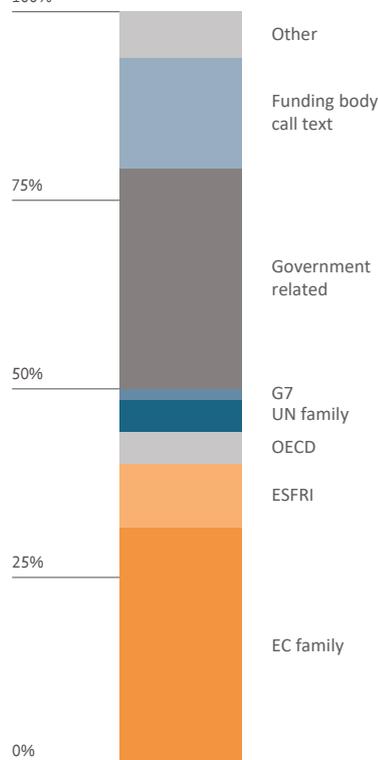
SOURCES MENTIONING ELIXIR

50



CATEGORIES OF SOURCES

100%



75 https://www.oecd-ilibrary.org/science-and-technology/building-and-sustaining-collaborative-platforms-in-genomics-and-biobanks-for-health-innovation_11d960b7-en.

76 https://erc.europa.eu/sites/default/files/document/file/ERC_info_document-Open_Research_Data_and_Data_Management_Plans.pdf.

77 https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/programme-guide_horizon_en.pdf.

78 <https://www.vr.se/analys/rapporter/vara-rapporter/2021-09-22-national-benefits-from-swedish-membership-in-international-research-infrastructures-2016-2019.html>.

Node capacity building

In 2021, ELIXIR stepped up its activities around impact evaluation through a Strategic Implementation Study involving 16 Nodes and the Australian BioCommons. Building on a previously-funded ELIXIR Staff Exchange project and the EU-funded *Research Infrastructures Impact Pathways* project, this 24 month-long study supports the long-term sustainability of ELIXIR Nodes. The study blueprint, *Demonstrating public value to funders and other stakeholders - the journey of ELIXIR⁷⁹* was published in 2021 (in *Annals of Public and Cooperative Economics*).

The study was complemented in 2021 by a series of four virtual and interactive workshops organised under the aegis of the EU-funded ELIXIR-CONVERGE project, in which more than 100 ELIXIR participants learned about *Outreach to funders and policy-makers, Communications, Impact evaluation* and *Organising impactful hybrid conferences*. Many participants attended several events and post-event surveys showed a high proportion of attendees went on to use the new skills in their work.

“The case studies and experiences of the different speakers were very insightful”.

“It worked well to have a practical exercise ... when you are forced to think and write down ideas then many ideas and questions arise”.

“A very good balance between theoretical learnings and practical sessions”.

“A friendly and open conversation where you could ask questions at any time and get an answer straight away”.

Quotes from participants in the ELIXIR-CONVERGE interactive workshops

⁷⁹ Martin, Corinne S., et al. Demonstrating public value to funders and other stakeholders—the journey of ELIXIR, a virtual and distributed research infrastructure for life science data. *Annals of Public and Cooperative Economics* 92.3 (2021)

Our people

ELIXIR Hub staff

The ELIXIR Hub provides, coordinates, and supports all ELIXIR Nodes and members with six different teams:

- Administration and Operations
- External Relations
- Human Genomics and Translational Data
- Project Management
- Technical
- Legal

Together with the ELIXIR Director, the Heads of these six teams form the management team of the ELIXIR Hub, alongside the Head of Programme and Strategy.

The staff in the Hub represent a diverse mix of nationalities and cultures, in turn reflecting the diversity of our Nodes.

Changes and additions

The ELIXIR Hub has significantly evolved and expanded to meet the needs of our partners. In January,

Melissa Konopko joined the Human Genomics and Translational Data team, supporting the coordination efforts in human data. In the spring, Katharina Heil joined the Technical team to coordinate the ELIXIR Communities and Peter Maccallum as the ELIXIR Chief Technical Officer. During the summer, Gary Saunders embarked on a new journey and left the Hub.

In the autumn, ELIXIR's capacity to manage the larger EC funding portfolio was strengthened by Andrea Troncoso who joined the Project Management Office as the BY-COVID project manager. The Legal Services and Technical teams were also bolstered, with Martina Caloi joining as a Legal Officer and Gavin Farrell as Technical Officer. Raj Mitra joined the External Relations team as web developer and at the end of the year, Anamika Chatterjee and Giselle Kerry joined the Human Genomics and Translational Data team.

Flags of all nationalities:



ELIXIR Hub structure



Governance

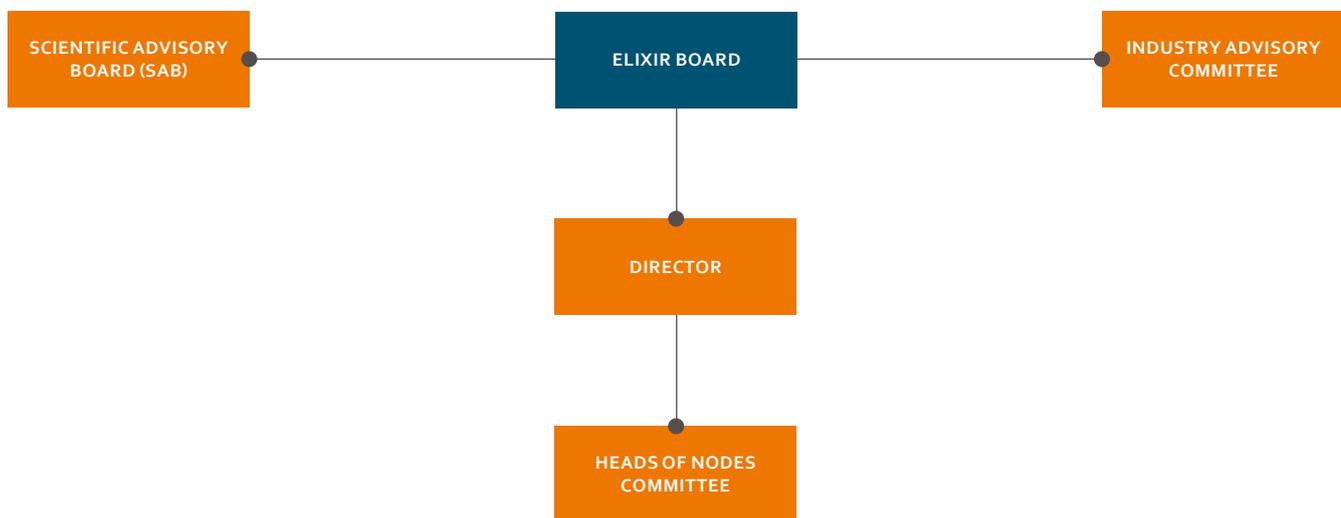
The highest decision-making body in ELIXIR is the ELIXIR Board, composed of representatives of ELIXIR members. The ELIXIR Scientific Advisory Board (SAB) advises the Board on ELIXIR scientific strategy and reviews the applications from ELIXIR Nodes. The SAB is an independent body, made up of leading experts from around the world. The committee also includes two independent ethics advisors to advise on ethical, legal and social

issues related to ELIXIR. The members are appointed by the ELIXIR Board. The Heads of Nodes Committee has a major role in developing and agreeing the ELIXIR scientific and technical strategy. The committee is composed of scientific representatives of each of the ELIXIR Nodes.

The Head of Node is appointed by each Node according to national processes. The Industry Advisory

Committee (IAC) consists of experts from industry users, including SMEs, suppliers and publishers who provide high-level strategic advice and input from industry stakeholders. Members of the IAC are appointed by the ELIXIR Board. The ELIXIR Director leads the ELIXIR Hub and is responsible to the ELIXIR Board for implementing the ELIXIR Scientific Programme. ELIXIR Director chairs the Heads of Nodes Committee.

ELIXIR governance hierarchy



ELIXIR Board members

Chair: Ferran Sanz (Spain)

Vice-Chair: Isabel Rocha (Portugal)

Vice-Chair: Christopher Rawlings (UK)

MEMBER	ADMINISTRATIVE DELEGATE	SCIENTIFIC DELEGATE
Belgium	Virginie Storms	
	Michele Oleo	
	Didier Flagothier	
Czech Republic	Jan Buriánek	Luděk Matyska (replaced Jaroslav Koča in October 2021)
Denmark	Line Bekker Poulsen	Anders Krogh
EMBL	Plamena Markova-Anderson	Alvis Brazma
		Edith Heard
Estonia	Toivo Rääm	Lili Milani
	Priit Tamm	
Finland	Riina Vuorento	Per Öster
	Sirpa Nuotio	
France	Eric Guittet	Hugues Roest Crollius (replaced Frederic Boccard in January 2021)
Germany	Annette Kremser	Alexander Goesmann
		Rolf Backofen
Greece	Argyro Karachaliou (replaced Maria Gkizeli in March 2021)	Babis Savakis
		Christos Ouzonis
Hungary	Klára Horváth (replaced Gábor Tóth in March 2021)	Zsuzsanna Dosztányi (replaced László Patthy in April 2021)
Ireland	Noelle Waldron (replaced Garry Purcell in September 2021)	Maria Nash
Israel	Barak Gatnyo	Iris Eisenberg
Italy	Grazia Pavoncello	Rita Casadio
Luxembourg	Romain Martin (replaced Lynn Wenandy in March 2021)	Regina Becker
	Bruno Rodrigues (replaced Jean-Claude Milmeister in March 2021)	
Netherlands	Ana de Castro	Ruben Kok
Norway		Rein Aasland
		Stig Omholt
Portugal	Andreia Feijão (replaced Marta Abrantes in November 2021)	Isabel Rocha
	Tiago Saborida	
Slovenia	Albin Kralj	Damjana Rozman
Spain	Ignacio Baanante	Ferran Sanz
	Cristina Bauluz	
Sweden	Mikael Borg	Björn Andersson
Switzerland	Doris Wohlfender-Bühler	Christian von Mering
UK	Mark Palmer	Christopher Rawlings
	Amanda Collis	

ELIXIR Heads of Nodes Committee

Chair: Niklas Blomberg (ELIXIR Director)

MEMBER	HEAD OF NODE	DEPUTY HEAD OF NODE
Belgium	Ferderik Coppens	Kim De Ruyck
Czech Republic	Jiří Vondrášek	Karel Berka (replaced Ludek Matyska in August 2021)
Denmark	Søren Brunak	
EMBL-EBI	Johanna McEntyre (replaced Rolf Apweiler in September 2021) and Ewan Birney	
Estonia	Jaak Vilo	Hedi Peterson
Finland	Tommi Nyrönen	Ilkka Lappalainen
France	Jacques van Helden and Claudine Médigue	Anne-Françoise Adam-Blondon
Germany	Andreas Tauch	Alfred Pühler
Greece	Martin Reczko	Christoforos Nikolaou
Hungary	Balázs Györfy	
Ireland	Denis Shields	Colm Ryan
Israel	Dan Ben-Avraham (replaced Michal Linial in May 2021)	
Italy	Graziano Pesole	Silvio Tosatto
Luxembourg	Reinhard Schneider	Wei Gu
Netherlands	Jaap Heringa	Morris Swertz
Norway	Inge Jonassen	Nils Peder Willassen (replaced Finn Drablos in April 2021)
Portugal	Mário Silva	Ana Portugal Melo
Slovenia	Brane Leskošek	
Spain	Alfonso Valencia	Salvador Capella-Gutierrez
Sweden	Bengt Persson	Jessica Lindvall
Switzerland	Ron Appel and Christine Durinx	
UK	Carole Goble and Neil Hall	
Cyprus	George Spyrou	Vasilis Promponas

ELIXIR Scientific Advisory Board members

Chair:

Francis Ouellette

Origin Bioinformatics, Canada

Vice-Chair:

Janet Kelso

Max Planck Institute for Evolutionary
Anthropology, Germany

Philip Bourne

University of Virginia, USA

Ana Sofia Carvalho

Catholic University of Portugal, Portugal
(appointed November 2020)

Jennifer Gardy

Bill & Melinda Gates Foundation, USA
(appointed November 2020)

Robert Gentleman

Harvard Medical School, USA

Melissa Haendel

Oregon Health and Science University, USA

Larry Hunter

University of Colorado, USA

Elina Ikonen

University of Helsinki, Finland

Nicola Mulder

UCT Computational Biology Group (NBN), South Africa

Susan Wallace

University of Leicester, UK

Doreen Ware

USDA ARS, Cold Spring Harbor Laboratory, USA

ELIXIR Industry Advisory Committee members

Chair:

Abel Ureta-Vidal

CMS Ventures, UK

Vice-Chair:

Natalia Jiménez Lozano

Atos, UK

Ian Barrett

AstraZeneca, UK

Thomas Exner

Edelweiss Connect GmbH, Switzerland

Andreas Kremer

ITTM, Luxembourg

Klaus Maisinger

Illumina, UK

Filip Pattyn

Ontoforce, Belgium

Jörg Peplies

Ribocon GmbH, Germany

Elizabeth Reynolds

General Bioinformatics, UK

María Rodríguez Martínez

IBM, Switzerland

Philippe Sanseau

GlaxoSmithKline, UK

Catherine Sirven

Bayer, France

Financial Data

		31/12/2021	2021	31/12/2020
		ACTUAL	BUDGET	ACTUAL
		€000	€000	€000
INCOME				
ELIXIR Member state contributions				
Ordinary contributions	(a)	7,378	7,400	7,233
Foreign exchange (loss)/gain on sterling contributions	(b)	(101)	-	(36)
Grant income	(c)	1,130	1,300	672
Miscellaneous income		(80)	-	100
Net Income		8,327	8,700	7,969
EXPENDITURE				
Technological Activities				
Salaries		706	800	597
Running costs		196	500	129
Commissioned services		2,894	5,100	1,402
Total expenditure Technological Activities		3,796	6,400	2,128
Directorate and Administrative expenditure				
Salaries		1,160	900	1,178
Running costs		297	700	318
Total expenditure Directorate and Administration		1,457	1,600	1,496
Support and Admin Infrastructure costs		883	700	898
Grant expenditure incurred		1,068	1,300	650
Total expenditure		7,204	10,000	5,172
SURPLUS/(DEFICIT)	(d)	1,123	(1,300)	2,797

(a) ELIXIR Member state contributions

	31/12/2021	31/12/2020
	€000	€000
Belgium	203	199
Cyprus	3	3
Czech Republic	71	70
Denmark	139	136
Estonia	10	9
Finland	100	98
France	1,126	1,104
Germany	1,560	1,529
Greece	90	89
Hungary	47	46
Ireland	83	81
Israel	129	126
Italy	833	816
Luxemburg	15	15
Netherlands	352	345
Norway	203	199
Portugal	85	84
Slovenia	18	18
Spain	547	536
Sweden	232	228
Switzerland	307	301
United Kingdom	1,225	1,201
Total	7,378	7,233

(b) Foreign exchange (loss)/gain on sterling contributions

The UK pay its Member State contributions in Sterling (ELIXIR/2015/28) as a hedging mechanism for currency movements, considering that most of the ELIXIR staff expenditures are in Sterling. The nominal loss arises from the difference between the value of these contributions valued in Euros at the date of payment and the date of the approval of the 2021 ELIXIR Budget due to fluctuations of the value of Sterling against the Euro in this period.

(c) Grant income

	2020	2019
	€000	€000
Grant funding awarded	8,135	7,812
Grant income earned in the current year	1,130	672
Grant expenditure incurred in the current year	(1,068)	(650)
Unutilised grant income	2,007	2,752

(d) Surplus/(Deficit)

This surplus is included in the EMBL general reserve, but has been ring-fenced for the use by ELIXIR.

(e) The following countries have amounts prepaid/owing at 31 December 2021

	2021 contribution owing
	€000
Denmark	(139)
Greece	(90)
Total	(229)

	Contribution prepaid
	€000
Germany	1,591
Total	1,591



Credits and Acknowledgments

This report was produced on the direction of the ELIXIR Board by the ELIXIR External Relations team at the ELIXIR Hub.

With a special thanks to all who contributed to the development of ELIXIR in 2021, notably the ELIXIR Heads of Nodes, Platform and Community Leads, Training and Technical Coordinators, Hub staff and members of the numerous working groups throughout ELIXIR.

Hinxton, UK, May 2022

Published under the CC BY 4.0 licence.

A complex network diagram with numerous nodes of varying sizes and colors (shades of blue and grey) connected by thin white lines. A large, thick white arc curves across the right side of the image. The text "Data for life" is written in white, following the curve of the arc.

Data for life

ELIXIR HUB

✉ info@elixir-europe.org

📍 South Building
Wellcome Genome Campus
Hinxton, CB10 1SD, UK