



Annual Report

2020





Annual Report

2020



Contents

6 Foreword

7 Preface

10

Our EU grants

10 B1MG

11 ELIXIR-CONVERGE

12

Our progress against strategic objectives

14

Our services

14 ELIXIR services

15 ELIXIR
Commissioned
Services

16

Our core structure

17 ELIXIR Nodes

28 ELIXIR Platforms

38 ELIXIR Communities

45

Our flagship events

45 All Hands Meeting – connecting all
ELIXIR members

46 BioHackathon Europe –
accelerating results in
bioinformatics



48

Our collaboration, outreach and industry support efforts

- 48 Strengthening global collaboration
- 49 Investing in people
- 50 Enabling and supporting industry collaboration
- 51 Demonstrating ELIXIR's impact
- 52 Empowering Nodes with ELIXIR F1000Research Gateway
- 53 Sustaining collaboration with virtual events

54

Our people

- 54 ELIXIR Hub staff
- 55 Striving for gender balance in bioinformatics
- 56 Governance

60

Financial data

Foreword



TWO YEARS INTO the current five-year plan and ELIXIR already has several milestones to celebrate. It has gone through a phenomenal journey from the initial six members in 2014 to its 22 members in 2021. The ELIXIR Nodes have become increasingly visible and active within their national life science infrastructure. Although ELIXIR's five-year plan, published in 2018, could not have considered the COVID-19 pandemic, ELIXIR has demonstrated the necessary agility to rise to the challenge and deliver crucial results. Establishing the European COVID-19 Data Platform is the perfect exemplification of this success.

ELIXIR and my team at the Open Science Unit collaborated closely to deliver the European COVID-19 Data Platform. Launched in April 2020 under the umbrella of the European Open Science Cloud (EOSC), it is a joint effort of the European Commission, EMBL-EBI, ELIXIR and other partners, including support from EU Member States. It aims to enable rapid and broad sharing of data and related resources for accelerating the discovery of cures, vaccines and treatments. ELIXIR supports the COVID-19 Data Platform's development through Horizon 2020 projects (e.g. CORBEL, EOSC-Life and ELIXIR-CONVERGE) and the activation of its national Nodes to enable cross-border coordination in sharing relevant data and resources. The Platform benefits from ELIXIR's Galaxy tool for workflow submission.

Now more than ever, the Platform has demonstrated the importance of rapid and open sharing of Findable, Accessible, Interoperable and Reusable (FAIR) data. It fulfils the European Commission's policy for Open Science, mainstreams FAIR data and encourages practices for making science accessible, equitable and reproducible.

Outside the area of infectious diseases, ELIXIR and EOSC have a shared ambition to build a federated ecosystem of data and services that make FAIR and open data sharing a reality – a “commons” for all scientists to access, reuse and build upon. In

EOSC-Life, a joint effort of thirteen biological and medical Research Infrastructures, ELIXIR's coordination role is instrumental in creating an open, collaborative digital space in the life science component of EOSC.

Furthermore, I see the establishment of the ELIXIR Core Data Resources, the ELIXIR Recommended Deposition Databases and the Recommended Interoperability Resources as a key development. They represent a leading European example in coordinating life science resources that other countries around the world can follow. These ELIXIR resources are featured in the Horizon 2020 guidelines for projects under a health emergency, and Horizon Europe guidelines will also continue to recommend them.

Finally, yet importantly, another highlight for 2020 was the launch of the Beyond 1 Million Genomes (B1MG) project, where ELIXIR plays an essential role. It will help realise the ambitious commitment of 24 European countries to make personalised medicine a reality by providing cross-border access to at least one million sequenced genomes by 2022.

I look forward to continuing our collaboration with ELIXIR to reinforce the COVID-19 data Platform and to build the EOSC. I wish ELIXIR the best in achieving many more milestones in the coming years.

Kostas Glinos

Head of Unit for Open Science at
the European Commission

Preface



THE YEAR 2020 HAS BEEN the most challenging year since ELIXIR's formal inception in 2014. The pandemic has affected many people in many different ways, from those suffering personal loss to everyone struggling with the imposed restrictions to curb the spread of the virus and the impact on our working practices. I am immensely proud of the relentless work of the whole of ELIXIR to provide world-class services despite this incredibly challenging operating environment. I look forward to sharing some of these successes with you in this 2020 Annual Report.

As a distributed infrastructure, what makes us strong is our community. We are made up of over 700 scientists and data experts across Europe, and our everyday working practices of face-to-face meetings, workshops and events had to change radically during 2020. Meetings were moved from face-to-face to virtual, and Zoom became the norm. And yet, throughout this, ELIXIR has continued to grow, achieve results and deliver impact.

In many ways, the challenges we faced in 2020 demonstrate the importance of open data. The launch and development of the COVID-19 Data Platform is a case in point. The rapid establishment was possible thanks to EMBL-EBI's leadership, the European Commission's support and the work of many ELIXIR partners. The European Galaxy team, led by our German and Belgian Nodes, has driven a global response, with colleagues in the US, Australia and elsewhere, that quickly established open, reproducible workflows for sequencing and analysis. The strong peer-networks built between ELIXIR Nodes over previous years have provided a firm foundation for rapid collaboration, knowledge sharing and support during the pandemic response.

Although 2020 will be remembered for the year the pandemic first struck, we should not forget the many other achievements that ELIXIR delivered. We launched the EU-funded Beyond One Million Genomes (B1MG) project – a critical component for implementing the Declaration on Access to

a Million Genomes. We launched a major effort where bioinformatics analysis tools are combined into workflows, packaged in 'containers' and distributed between national Nodes in a collaborative computing environment. The work of the pan-European ELIXIR team in the international *Federated Analysis Systems Project* run under the auspices of the Global Alliance for Genomics and Health was uniquely noted for connecting across both institutes and international borders.

ELIXIR has supported research fields across the whole of the life sciences. For example, the ELIXIR Plant Sciences Community has continued to develop the federated data discovery, search and biodiversity retrieval system with Europe's large-scale plant phenotyping centres. In 2020 ELIXIR strengthened links to emerging biodiversity initiatives across Europe with our first EU-funded grant awarded in this area.

As I look to the future, I remain excited about the opportunities ahead. 2021 sees the start of Horizon Europe. In turn, ELIXIR will release its data management toolkit (RDMkit) for researchers – providing easy access to a curated collection of research data management guidelines, tools and services. Together with our network of research data management professionals across our Nodes, we think this will be a key component for the ambition to make all European life science data FAIR.

I end with a heartfelt thanks to all of ELIXIR's partners and collaborators. As you will see from this Annual Report, our achievements in 2020 would not have been possible were it not for the tireless efforts of you all.

Niklas Blomberg
ELIXIR Director
Hinxton,
Cambridge, UK
February 2021

ELIXIR in 2020



DRIVING PROJECTS FORWARD

In 2020, ELIXIR has been involved in eight ongoing European projects:

CORBEL

eTRANSafe

RI Impact Pathways

EOSC-Life

FAIRplus

EU-STANDS4PM

EOSC Enhance

CINECA

And has started coordinating two new projects:

ELIXIR-CONVERGE

€7.8M

to deliver the provisioning of distributed local support for data management across Europe

B1MG

€4.0M

to support implementation of the 1+ Million Genomes Initiative



COVID-19 GLOBAL EFFORT

All ELIXIR Nodes have provided numerous services for studying SARS-CoV-2 and COVID-19 at a national and global level.

12

databases
to access
COVID-19 data

1

unified portal to
access and help store
COVID-19 data

5

resources to make
COVID-19 data findable
and accessible

8

computing resources
gave priority access to
COVID-19 projects

30

COVID-19
peer-reviewed
publications

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.



In 2020, ELIXIR kicked-started two ambitious projects as coordinator:

B1MG

TIMELINE June 2020 – June 2023

BUDGET €4M

WEBSITE b1mg-project.eu

The EU-funded Beyond 1 Million Genomes (B1MG) project primarily supports the goals of the 1+ Million Genomes initiative¹, based on the commitment of 21 EU states, United Kingdom and Norway, to 'access at least 1 million sequenced genomes in the EU by 2022.' B1MG will offer support, coordination and structure to achieve the initiative's objectives.

The project will collaborate with international initiatives and consult with numerous stakeholders to support the creation of a pan-European federated genome-based health data infrastructure. B1MG will generate legal guidance, best practices and recommendations

to create infrastructure. These recommendations will be translated into a B1MG maturity level model that provides concrete guidance on the steps required to implement personalised medicine, a healthcare approach that takes into account a person's genetic make-up, at a local, regional and national scale. Ultimately, B1MG endeavours to facilitate the broad implementation of personalised medicine and its ensuing notable socio-economic benefits.

2020 achievements

- Identified local, national, European and global actors in personalised medicine and organised the first stakeholder meeting.
- Launched the B1MG Stakeholder Portal to obtain broad and efficient engagement of European stakeholders, such as patients, academia, industry or medical authorities.

- Initiated the development of the B1MG maturity level model
- Delivered workshops for two main purposes:
 - To focus on healthcare requirements from three use cases: rare diseases, common complex diseases and cancer
 - To consider the requirements for a synthetic data cohort for testing infrastructure components
- Defined and documented the scope of several topics including, technical infrastructure, industry engagement, and ethical, legal and social implications

¹ <https://digital-strategy.ec.europa.eu/en/policies/1-million-genomes>



ELIXIR-CONVERGE

TIMELINE February 2020 – February 2023

BUDGET €5M, increased to €7.8M following a COVID-19 uplift

WEBSITE elixir-europe.org/converge

The ELIXIR-CONVERGE project builds on the foundations of ELIXIR-EXCELERATE to deliver the provisioning of distributed local support for data management across Europe.

The project will develop national operations to drive good data management, reproducibility and reuse. Over 36 months and with partners from all ELIXIR Nodes, ELIXIR-CONVERGE takes the next step to create a European data federation, where interconnected national operations allow users to extract knowledge from life science's large, diverse and distributed datasets.

2020 achievements

- Started developing the RDMkit, a toolkit for data management support, ahead of its public launch in 2021.
- Brought together ELIXIR's data management experts through a Data Management Network to share knowledge and best practices
- Held outreach events with several countries interested in joining ELIXIR, including a series of webinars with the Romanian research community
- ELIXIR Nodes have developed national data hubs for infectious disease data, led by ELIXIR's establishment of the European COVID-19 Data Portal



COVID-19 in focus

In November 2020, the European Commission amended the ELIXIR-CONVERGE grant with an uplift of €2.8 million for partners to initiate a new set of data-related activities around COVID-19. Of that uplift, €1.5 million was awarded to enhance human and viral data sharing through the ELIXIR Nodes by supporting the development of the COVID-19 Data Platform, with a €1.3 million uplift to support the ELIXIR's Federated EGA activities.

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 into 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. **28**

TRAINING
PLATFORM

P. **36**

INTEROPERABILITY
PLATFORM

P. **30**

TOOLS
PLATFORM

P. **32**

COMPUTE
PLATFORM

P. **34**

2

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

ELIXIR
COMMUNITIES

P. **38**

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.28

GLOBAL BIODATA
COALITION
COLLABORATION

P.48

4

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

DATA
PLATFORM

P.28

INTEROPERABILITY
PLATFORM

P.30

TOOLS
PLATFORM

P.32

COMPUTE
PLATFORM

P.34

5

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

ELIXIR
NODES

P.17

BIOHACKATHON
EUROPE

P.46

INVESTING IN
PEOPLE

P.49

ENABLING AND
SUPPORTING
INDUSTRY
COLLABORATION

P.50

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 2020

ELIXIR AAI

- Over 6200 logins per month
- Enabled over 1100 institutions whose members can use ELIXIR AAI
- Integrated 104 production services in total and an additional 110 in testing
- Had 5374 users

FAIRsharing

- 1740 databases
- 140 policies
- 1518 standards

ELIXIR TeSS

- 1485 training materials
- 22 training workflows
- 79 content providers

BioContainers

- 9.6k tools
- 35.3k versions
- 162.7k containers and packages

The number of services run by ELIXIR Nodes continued to expand in 2020

351

SERVICES IN TOTAL

20

COMPUTE

116

DATA RESOURCES

34

INTEROPERABILITY

195

TOOLS

26

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.

A total of 41 Commissioned Services ran during 2020, engaging activity from across all ELIXIR Nodes. Three of these projects were new and launched in 2020:

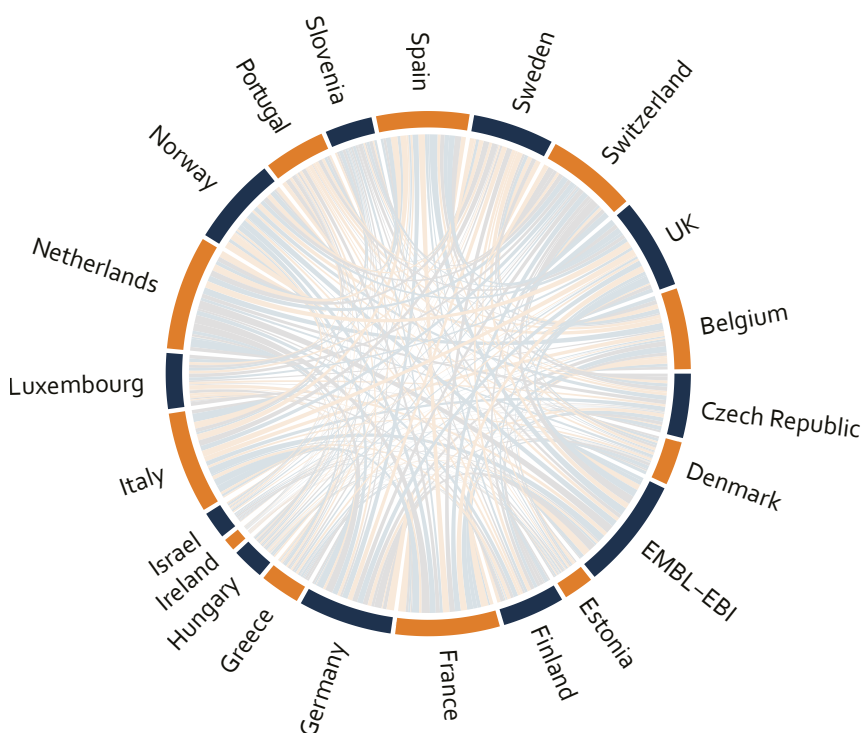
- Exploiting Bioschemas mark up to support ELIXIR Communities (Strategic Implementation Study)
- 3D-Bioinfo Community (Community Implementation Study)
- Microbial Biotechnology (Community Implementation Study)

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.

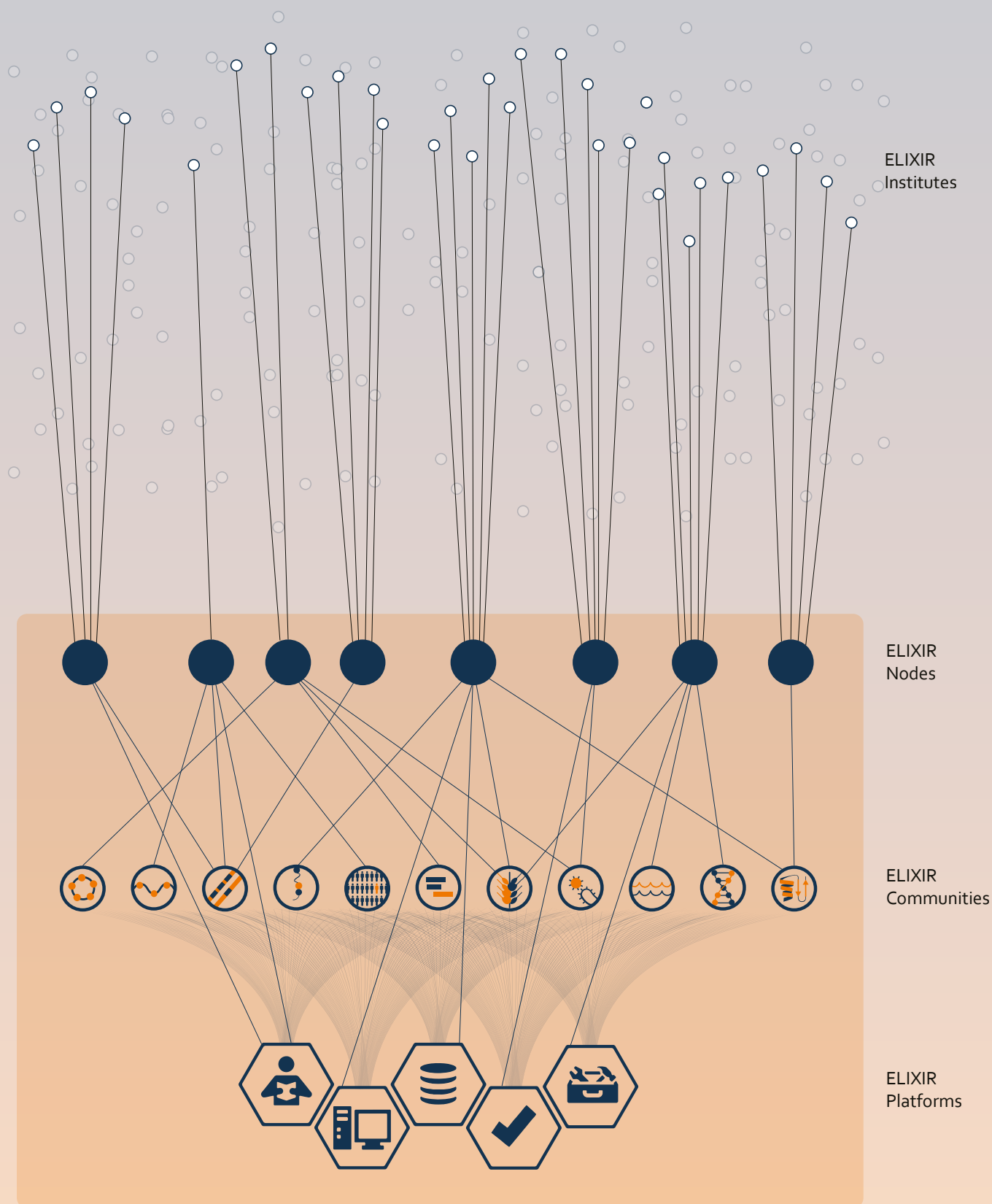
41

COMMISSIONED
SERVICES RAN DURING 2020

Connection between ELIXIR Nodes facilitated by ELIXIR Commissioned Services in 2020



Our core structure

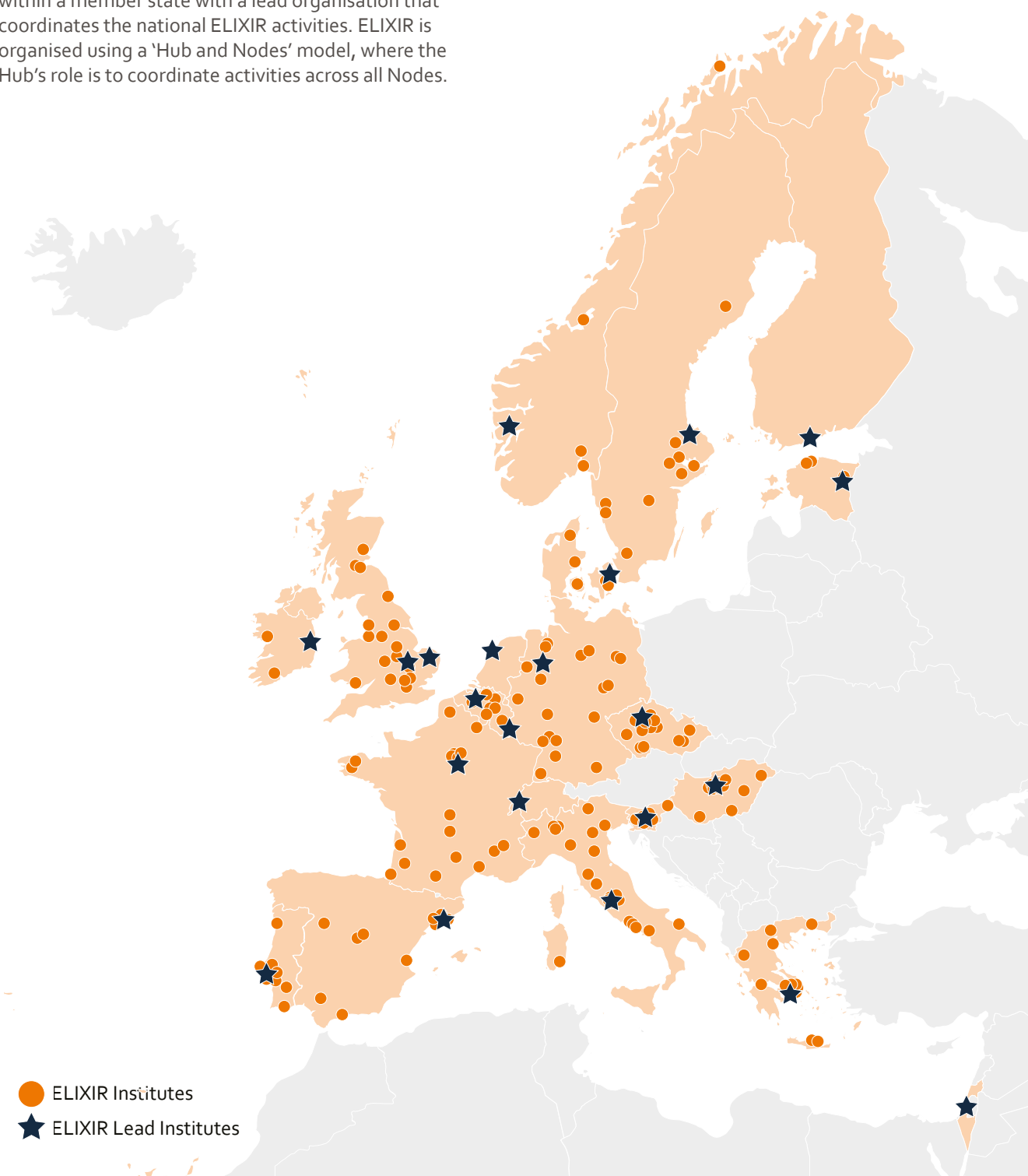


ELIXIR Nodes

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 that coordinates 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





COVID-19 in focus

Developed a tool, its guidance and demo to facilitate the submission of SARS-CoV-2 sequencing data to the European Nucleotide Archive

Helped develop the EOSC-Life's WorkflowHub and integrate workflows to analyse SARS-CoV-2 data and best practices for Galaxy

Supported the establishment of the COVID-19 national coordination team



COVID-19 in focus

Participated in the research team established by IDEA at CERGE-El to model the epidemiological situation in the Czech Republic

Offered computing resources to the testing centre at Palacký University, which hosts EATRIS-CZ, for processing large amounts of test data



ELIXIR Belgium

LEAD INSTITUTE VIB

WEBSITE www.elixir-belgium.org

- Co-developed RO-Crate, a lightweight approach to packaging research data with their metadata
- Co-developed the RDMkit, a one-stop-shop of information, advice and signposting for research data management as part of ELIXIR-CONVERGE
- Assisted the establishment of the Flemish Open Science Board and the Flemish Research Data Network
- Expanded the ELIXIR Belgium Service Delivery Plan to 19 Node Services
- Helped establish the ELIXIR 3D-Bioinfo Community
- Participated in staff exchange events to collaboratively develop the FAIRDOM-SEEK data management platform
- Founded the national community of Research Software Engineers (RSE)
- Reached out to industry and other research institutes with guest lectures and workshops
- Organised 18 training courses and events for a total of 477 participants
- Established national collaborations with LifeWatch, EMBRC and DiSSCo on data management and Galaxy



ELIXIR Czech Republic

LEAD INSTITUTE Institute of Organic Chemistry and Biochemistry of the CAS

WEBSITE www.elixir-czech.cz

- Agreed on the Scientific Strategy for the 2021-2025 ELIXIR CZ Scientific Programme, reviewed by the ELIXIR CZ Scientific Advisory Board and approved by the ELIXIR CZ Board
- Included 14 new services in the Service Delivery Plan
- Delivered 26 training and workshop events on topics such as Galaxy, AI in Life Sciences and In Silico drug design
- Participated in the ELIXIR-CONVERGE Horizon 2020 project, in 12 ELIXIR Implementation Studies and two further ELIXIR Staff Exchanges
- Helped the Institute of Health Information and Statistics of the Czech Republic to create an epidemiological model of the pandemic



ELIXIR Denmark

LEAD INSTITUTE Technical University of Denmark

WEBSITE www.elixir-denmark.org

- Participated in several ELIXIR Implementation Studies and ELIXIR Communities, such as the Proteomics and 3D-Bioinfo Communities
- Expanded their work in the rare diseases area, notably with the efforts in the EJP-RD project
- Implemented and promoted the use of research data management plans in multiple projects, both domestic and international
- Continued working within EU-STANDS4PM and the eTRANSAFE project



ELIXIR Estonia

LEAD INSTITUTE University of Tartu

WEBSITE elixir.ut.ee

- Selected as an Estonian Core Research National Infrastructure
- Created ELIXIR Estonia eQTL Catalogue Browser
- Published a new R package for g:Profiler, an ELIXIR Recommended Interoperability Resource
- Published PAWER, a protein microarray analysis tool
- Joined the management team of the ELIXIR TeSS Training Portal
- Estonia joined the Nordic e-Infrastructure Collaboration (NeIC)



ELIXIR Finland

LEAD INSTITUTE CSC – IT Center for Science

WEBSITE www.elixir-finland.org

- Attracted more than €6.5M of new international funding
- Secured a place on the national research infrastructure roadmap for 2021-2024
- Delivered the moonshot demo at the ELIXIR All Hands Meeting
- Completed the ELIXIR AAI 2019-2020, a 24-month project
- Co-delivered the scoping paper for the technical infrastructure of B1MG project



COVID-19 in focus

Created tracking dashboards to support COVID-19 efforts

Collaborated with national agencies to prepare against infectious diseases

Dedicated parts of the national HPC Infrastructure to COVID-19-related projects

Created a special view of Bio.tools for COVID-19 with an efficient model to populate and update entries



COVID-19 in focus

Developed tools for COVID-19 data analysis and visualisation (koroona.ut.ee)



COVID-19 in focus

Designed the 1+MG federated data infrastructure with support for COVID-19 data

Collaborated on the national COVID-19 host genetic data sharing

COVID-19 in focus

Launched a national COVID-19 task force providing support to life scientists and clinicians, including use of computing facilities, development of databases, tools and web resources, data curation, support in biostatistics, phylogeny inference, and training

Collaborated with national agencies for the surveillance and prevention of infectious diseases



ELIXIR France

LEAD INSTITUTE CNRS

WEBSITE www.france-bioinformatique.fr

- Secured €16.5M of funding for the next five years by the French Ministry of Higher Education, Research and Innovation under the Mutualised Digital Spaces for FAIR Life Sciences (MuDiS₄LS) roadmap
- Launched the IFB/ELIXIR-FR monthly newsletter
- BiImage Informatics Index (Biii) resource became an ELIXIR Recommended Interoperability Resource
- Together with ELIXIR-CH and the ELIXIR Human Copy Number Variation Community, coordinated the publication of the Community roadmap
- Helped create the FAIR-compliant ecosystems of tools and resources
- Extended the ELIXIR France Service Delivery Plan with a joint submission with the Norwegian Node for EDAM
- Co-organised, with ELIXIR-LU, a capacity building workshop on 'ELIXIR Best practices in research data management and stewardship'



ELIXIR Germany

LEAD INSTITUTE Bielefeld University

WEBSITE www.denbi.de/elixir-de

- Signed the ELIXIR Collaboration Agreement
- Extended the ELIXIR Germany Service Delivery Plan to 66 bioinformatics services
- Delivered extensive online training with 40 courses for 1,149 participants
- Participated in fifteen ELIXIR Implementation Studies
- and helped prepare seven proposals for Community-led Implementation Studies
- Contributed to six ELIXIR publications

COVID-19 in focus

Prioritised access to the computing facility de.NBI Cloud for COVID-19 research projects and further support with the European Galaxy server

Established the de.NBI/ELIXIR Germany task force for COVID-19 human omics data management and international dissemination. This task force acts as the national point of contact for omics data for ELIXIR and the COVID-19 Data Portal

Supported the COVID-19 Disease Map consortium through the FAIRDOMHub service



ELIXIR Greece

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

WEBSITE www.elixir-greece.org

- Participated in eight ELIXIR Implementation Studies
- Included a new partner, the National Centre for Scientific Research 'Demokritos'
- Fotis Psomopoulos appointed as a new Executive Committee member for the ELIXIR Training Platform
- Contributed to the moonshot demo at the ELIXIR All Hands Meeting and presented the Compute Infrastructure
- Organised and hosted the ELIXIR Training Platform lesson Hackathon on Machine Learning and Biostatistics
- Hosted an ELIXIR Train-the-Trainer event as a hybrid event
- Delivered training to two other Nodes ('Exploratory Analysis of Biological Data', EMBL/DE and 'Introduction to Machine Learning', ELIXIR-CH)
- Connected a local EGA instance to the testing central EGA server using Node specific credentials, technically ready for submissions



ELIXIR Hungary

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

WEBSITE elixir-hungary.org

- Appointed a new Training Coordinator, Dr Zsuzsanna Dosztanyi, from the Eötvös Loránd University
- Publication of seven new scientific papers with direct ELIXIR Hungary participation having a lead author from one of the ELIXIR research groups
- Established new Galaxy servers at the University of Pécs and at the Semmelweis University
- Started the negotiation process with the EGA to establish a Hungarian EGA server
- Supervised 57 PhD students ELIXIR Hungary research group leaders
- Organised 31 bioinformatic courses and conferences in Pécs, Debrecen, Budapest and online



COVID-19 in focus

Contributed extensively to COVID-19 research, including the setup of a freely-available new Galaxy cluster with COVID-19 best-practice pipelines

Attended the COVID-19 Virtual BioHackathon, leading the project on Machine Learning, and participating in several others



COVID-19 in focus

Developed COVIDoutcome, a COVID-19 platform to estimate disease severity based on SARS-CoV-2 genome mutations



COVID-19 in focus

Science Foundation Ireland's COVID-19 Rapid Response funding included COVID-19 genome sequencing and analysis, a consortium led by Teagasc, and computational design of peptides disrupting coronavirus infection, led by University College Dublin



COVID-19 in focus

Contributed to the national COVID-19 testing activities with a lab data management software (LIMS) platform



COVID-19 in focus

Developed new activities to support COVID-19 research including CorGAT, a novel tool for the functional annotation of SARS-CoV-2 genomes

Started working on the Italian instance of the COVID-19 Data Portal and collaboration with health authorities for national data sharing

Organised a webinar series on 'How Open Science and Data Sharing can help in the fight against COVID-19'



ELIXIR Ireland

LEAD INSTITUTE University College Dublin

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

- Appointed a new Training Coordinator, Dr Simone Coughlan, of the National University of Ireland (NUIG)
- Appointed a new Data Management Coordinator, Dr Graham Hughes, of University College Dublin (UCD)
- Organised the first online PhD student symposium thanks to the excellent collaboration of the participating institutions (NUIG, UCC and UCD) under the umbrella 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

- Co-led the 3D-Bioinfo Community
- Participated in the ELIXIR-carpentries trainers course
- Engaged in multiple ELIXIR events and workshops



ELIXIR Italy

LEAD INSTITUTE CNR Institute of Biomembrane and Bioenergetics

WEBSITE elixir-italy.org

- Contributed to several BioHackathon Europe 2020 projects on different topics such as the ELIXIR service bundle for epidemic response, exploiting bioschemas markup in Community registries, designing a modular learning path or creating a rare disease service bundle for assessing molecular pathogenicity of the genetic variants
- Organised diverse online workshops and conferences on topics including software carpentry, intrinsically disordered proteins and FAIR data
- Launched the Laniakea@ReCaS Galaxy on-demand cloud service
- Established a new Pulsar node in Italy for the Galaxy Community
- Co-delivered the first online 'Train the Trainer' course and delivered the second edition
- Co-developed a simple, affordable process for impact assessment of research activities together with ELIXIR Norway and ELIXIR Portugal



ELIXIR Luxembourg

LEAD INSTITUTE Luxembourg Centre for Systems Biomedicine

WEBSITE elixir-luxembourg.org

- Co-led three projects at the Biohackathon Europe 2020: Galaxy and disease maps; Beacon for clinical and translational data; and Cheminformatics resources and applications
- Co-organised the Translational Medicine 2020 Community of Special Interest Group (COSI) satellite meeting at ISMB 2020
- As part of FAIRplus project, contributed to the FAIR Cookbook and released the new version of IMI-data catalogue, co-developed with ELIXIR UK



ELIXIR Netherlands

LEAD INSTITUTE Dutch Techcentre for Life Sciences

WEBSITE www.dtls.nl/elixir-nl

- FAIR data stewardship continued to be the main focus of ELIXIR-NL driving numerous activities
- Participated in ELIXIR Communities and Focus Groups including Metabolomics, Proteomics, hCNV, Toxicology, Food and Nutrition and Systems Biology.
- As part of the ELIXIR-CONVERGE project, contributed to the Data Management Network, training activities and the RDMkit
- Contributed to the FAIR Cookbook, as part of the FAIRplus project
- Chaired the coordination group of the European 1+ Million Genomes Initiative, co-chaired the working groups, and participated in the associated B1MG project.
- Formulated the Professionalising Data Stewardship Roadmap of the Dutch National Programme on Open Science
- Co-lead the ELIXIR Interoperability and Training Platforms
- Represented the life sciences field in the FAIRsFAIR project, delivering results on FAIR Data Points and recommendations for FAIR practice
- Delivered recommendations for the EOSC Executive Board through participation in the EOSC FAIR and the EOSC Skills and Training working groups



COVID-19 in focus

Led the development and hosting of the COVID-19 Disease Map, a knowledge repository of the molecular mechanisms of COVID-19

Provided data management, analysis and IT infrastructures to support the national research platform for COVID-19 and national cohort studies



COVID-19 in focus

Trained data stewards from COVID-19 related projects in FAIR data stewardship along with health funder ZonMw and Health-RI

Developed disease models and established a national COVID-19 Data Portal

Co-chaired an RDA group to deliver COVID-19 data sharing recommendations



COVID-19 in focus

Promoted open access to COVID-19 data by supporting the Norwegian Institute of Public Health, sharing viral genomes to the European Nucleotide Archive, providing tailored workflows, establishing an open-access SARS-CoV-2 database, establishing and ELIXIR-NO COVID-19 task-force and setting up the Norwegian instance of the COVID-19 Data Portal



COVID-19 in focus

Contributed to the national health data ecosystem to implement tailored electronic case report forms (eCRF) and a dedicated human data hub to facilitate clinical COVID-19 research projects, such as VITACOV



ELIXIR Norway

LEAD INSTITUTE University of Bergen

WEBSITE www.elixir-norway.org

- Kicked off the national infrastructure project, called BioMedData, to promote FAIR management of life science data from generation to deposition
- Appointed a new Training Coordinator, Erik Hjerde, from University of Tromsø
- Successfully migrated five local NeLS-Galaxy instances into one common national UseGalaxy server, with almost 2000 tools and multiple analysis workflows
- Submitted national proposal for continued structural funding (2021-26) for ELIXIR Norway to the Research Council of Norway
- Expanded the ELIXIR Norway Service Delivery Plan, resulting in 14 new national services



ELIXIR Portugal

LEAD INSTITUTE Instituto de Engenharia de Sistemas e Computadores

WEBSITE elixir-portugal.org

- Provided crucial expertise for a publication describing MIAPPE v1.1².
- Launched the CorkOakDB portal, an ELIXIR Portugal Node Service for applied studies on *Quercus suber*, with a focus on genetics
- Added four new ELIXIR Portugal Node services: Bioinformatics Docker Images Project (Pegi3S), D-Cellerate, EVOPPI, and 'Ready for BioData Management?'
- Matured the 'Ready for BioData Management?' capacity building programme, with online learning modules hosted by 11 institutions, for 297 participants, from 39 different organisations
- Deployed a prototype of the local EGA.
- Collaborated with The Navigator Company through their first ELIXIR Knowledge Exchange Scheme, marking the first time the MIAPPE standard was adopted by industry
- Developed a simple, affordable process for impact assessment of research activities in the framework of a dedicated ELIXIR Staff Exchange Scheme with other ELIXIR Nodes

² 1. Papoutsoglou, E. A. et al. Enabling reusability of plant phenomic datasets with MIAPPE 1.1. *New Phytol.* 227, (2020). (<https://doi.org/10.1111/nph.16544>)



ELIXIR Slovenia

LEAD INSTITUTE University of Ljubljana

WEBSITE www.elixir-slovenia.org

- Started implementing a €5.3 million national infrastructure project (RI-SI-2) awarded to the Node through European Regional Development Funds
- Launched a new ELIXIR-SI website listing all Node's services and used as a single-entry point for ordering services and booking time slots for all available equipment
- Fully developed an ELIXIR Slovenia e-learning platform with diverse courses and webinars
- Co-lead the ELIXIR Plant Sciences Community, and actively participated in the ELIXIR Compute Platform, the ELIXIR Rare Diseases Community and other human data ELIXIR Communities



ELIXIR Spain

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

WEBSITE inb-elixir.es

- Organised the first virtual edition of the European Conference on Computational Biology (ECCB) with 1,000 participants from more than 50 countries and exemplary participation from ELIXIR Nodes.
- Received a positive evaluation from the ELIXIR Scientific Advisory Board, highlighting the Node's efforts in human data and gender balance initiatives through the Bioinfo4Women programme.
- Contributed to the FAIR Cookbook as part of the FAIRplus project
- Secured national funding within the Spanish IMPaCT (Infrastructure for Personalized Medicine associated with Science and Technology) through the iDATA-MP programme focused on Data Science (2021-2023)
- Co-lead the ELIXIR Tools Platform, and the ELIXIR Communities on Rare Diseases and Federated Human Data, as well as the Cancer Focus Group



COVID-19 in focus

Started developing a COVID-19 registry of patient data to enable online data collection and provide (some) compute resources for national and international research



COVID-19 in focus

Co-lead the newly established work package in ELIXIR CONVERGE on Federated European Genome-phenome Archives for transnational access of COVID-19 host data

Launched the Spanish instance of the COVID-19 Data Portal

Participated in scientific committees to advise the Spanish Government in their COVID-19 response

Developed the CRG COVID-19 Viral Beacon, a tool for finding SARS-CoV-2 variability at genomic, amino acid and motif level



COVID-19 in focus

Set up the Swedish instance of the COVID-19 Data Portal, together with the SciLifeLab Data Centre



ELIXIR Sweden

LEAD INSTITUTE NBIS – National Bioinformatics Infrastructure Sweden

WEBSITE nbis.se

- Continued expansion of the Node to support 267 national principal researchers with bioinformatics support, including drop-in and training sessions that became online in 2020
- NBIS – the SciLifeLab Bioinformatics platform – was favourably evaluated by an international review panel
- Led the formation of the Data Management Network as part of the ELIXIR-CONVERGE project
- Co-led the efforts to create an infrastructure for the European 1+ Million Genomes initiative as part of the B1MG project
- Continued developing the federated EGA, and established the Swedish EGA node at a technical level
- Secured Nordic funding in the NeIC Heilsa Tryggvæddottir (Tryggve3) project (2021-2023)
- Released version 20 of the Human Protein Atlas, an ELIXIR Core Data Resource, to add single cell type specific RNA expression in 53 different cell types
- Worked with partners on a proposal for the EU IMI project, called BigPicture, to create a digital pathology infrastructure, which will start in 2021



ELIXIR Switzerland

LEAD INSTITUTE SIB Swiss Institute of Bioinformatics

WEBSITE www.sib.swiss



COVID-19 in focus

Joined international efforts to accelerate COVID-19 research with numerous dedicated databases and software tools, including a dedicated webinar series on SIB Resources with more than 900 participants

Co-launched an e-workshop for the general public to understand the biology of SARS-CoV-2

- Tailored a secure infrastructure network, BioMedIT, to provide all authorised researchers in Switzerland with easy access to collaborative analysis of confidential data
- Co-launched the first public version of the Swiss Variant Interpretation Platform to offer clinicians a harmonised interpretation of cancer variants
- Released the SPHN Fact-Sheet 2020³ with the Swiss Academy of Medical Sciences
- Joined forces with international, public-private research consortium SOPHIA – Stratification of Obese Phenotypes to Optimize Future Obesity Therapy
- Became a new bioinformatics partner of the international Quality Control for Molecular Diagnostics (QCMD) organisation
- Launched the new version of Expasy.org, the Swiss Bioinformatics Resource Portal, now entirely annotated with the EDAM ontology
- Launched the 11th edition of the Bioinformatics Awards
- Expanded the online bioinformatics training offer, reaching 72 countries
- Created a Diversity Working Group to tackle the topics of equality, diversity and inclusion at community level

3 <https://sphn.ch/2020/11/13/factsheet-2020/>



ELIXIR UK

LEAD INSTITUTE Earlham Institute

WEBSITE elixiruknode.org

- Lead development on RO-Crate, a lightweight approach to packaging research data with their metadata
- Co-developed the RDMkit, a one-stop-shop of information, advice and signposting for research data management as part of ELIXIR-CONVERGE
- Co-developed the FAIR Cookbook as part of the FAIRplus project
- Expanded the ELIXIR UK Service Delivery Plan with six new services: FAIRDOM-SEEK, GWAS Central, HGNC, LIPID Maps, PomBase and PSIPRED, alongside a roadmapped service, KnetMiner.
- Obtained national funding for the BioFAIR feasibility study to gather evidence and make the case for a UKBioFAIR Institute to establish training and a fellowship of stewards on research data management
- Founded and established the WorkflowHub, a registry for workflows. Established Bioschemas Computational
- Workflow community profile for describing workflows for the WorkflowHub, and the Interoperability and Tools Platforms
- Aligned and worked with DISSCo and IBISBA with regards to RO-Crate, ELIXIR Workflows infrastructure and FAIRDOM-SEEK
- Reported on metadata strategy for data cataloguing for inclusion in the EOSC, through EOSC-Enhance
- Leading the FAIR Computational Workflows activities resulting in one publication and four events
- Appointed new Training Coordinators and launched two new national Working Groups on data management and training
- Leading on the co-founded Terms4FAIRskill terminology initiative
- Co-lead the ELIXIR Health Data Focus Group and the ELIXIR Galaxy, Microbial Biotechnology and 3DBioinfo Communities



EMBL-EBI

WEBSITE www.ebi.ac.uk

- Hosted and integrated FAIR data stewardship as part of the EMBL-EBI Competency Hub activities in ELIXIR-CONVERGE
- Took part in activities of several ELIXIR Communities and corresponding Implementation Studies
- In relation to plant sciences, implemented the MIAPPE checklist for validation at submission time and new versions release time
- Achieved widespread use of JSON validator across EBI resources
- Supported submissions of samples' metadata into BioSamples and EV.
- Ran an EOSC-Life remote learning series for supporting research infrastructures to provide remote training
- Contributed to the FAIR Cookbook as part of the FAIRplus project



COVID-19 in focus

Led the development of COVID-19 resources, including the COVID-19 WorkflowHub, a FAIRsharing collection of COVID-19 resources and a KnetMiner for COVID-19

Supported the hosting of the COVID-19 Disease Map through the FAIRDOMHub



COVID-19 in focus

Established the European COVID-19 Data Platform, which facilitates data sharing and analysis in order to accelerate coronavirus research

Initiated the National Coordination meeting series to facilitate the national COVID-19 activities across Europe

Launched COVID-19 Vocabulary (COVoc), an ontology containing terms related to COVID-19 research

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 coordinate the development and operation of ELIXIR services 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

Strategic Objectives



COVID-19 in focus

The launch of the COVID-19 Data Portal is an example of the Data Platform's overarching aim. The Portal is underpinned by many of the Platform's resources, such as PDBe, SWISS-MODEL, Ensembl, Europe PMC, ENA, EGA, UniProt/Swiss-Prot, Cellosaurus and many others.

First and foremost, 2020 saw the ELIXIR Data Platform aim to develop services to cater to the diverse requirements and needs of ELIXIR Communities

The Data Platform finalised round three of the ELIXIR Core Data Resources (CDRs) selection process and successfully included three new ones – Rhea, Reactome and Cellosaurus. With these three new additions, there are now 21 Core Data Resources covering even more data types.

The databases included in the selected list increase their likelihood of achieving long-term support. The ELIXIR Core Data Resource list acts as an indicator, a benchmark in Europe. Its global impact is also worth mentioning; the selection process used in ELIXIR will serve the Global Biodata Coalition as a best-practice model to develop a global list of key life science databases.

Complementing the work in coordinating invaluable data resources, the Data Platform launched a request for proposals on scalable curation⁴. This request is a biocuration exercise to add annotations to otherwise primitive data for advanced data interpretation. The call is expected to support up to eight Implementation Studies for strengthening scalable data resource management, such as semantic annotations of scientific publications, curation process or data integration.

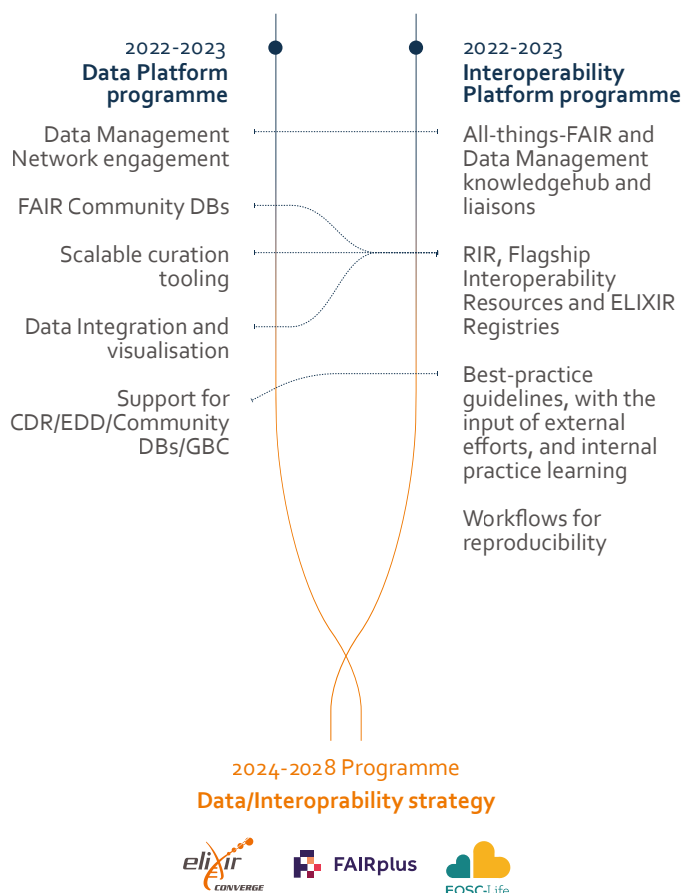
The year 2020 also saw the Data Platform embrace the virtual format for its annual face-to-face meeting to discuss the strategy for the remaining years of the 2019-2023 programme.

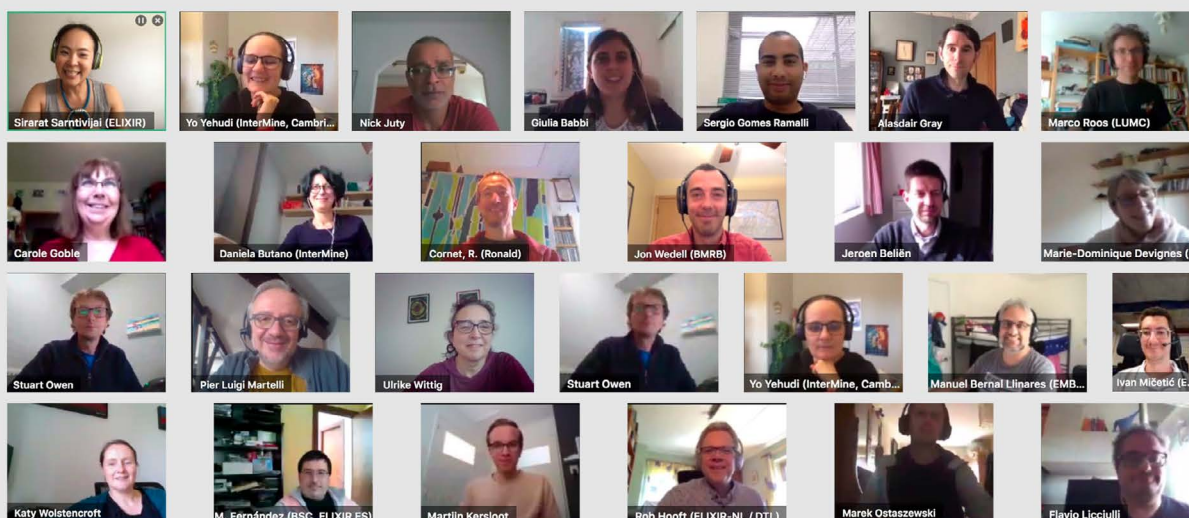
⁴ https://drive.google.com/file/d/19l5-lKvLkiFqGGrl7YBwKGxSHgghv_wO/view

Work was also initiated to bring the Data Platform and Interoperability Platform closer together. This ongoing cooperation included two principal joint efforts in 2020:

1. An alignment between the Data Platform resources and the use cases for the Interoperability Platform Bioschemas Strategic Implementation Studies. Use cases in the area of rare disease, plant informatics, and intrinsically disordered proteins have been selected from the corresponding data resources, Orphadata, CorkOakDB and IDPCentral.
2. A collaboration in various data management activities, such as developing the CONVERGE Data Management Expert Network and RDMkit or delivering the ELIXIR All-hands Data Management Workshop.

Data Interoperability – the complementary view





Participants of the ELIXIR Interoperability Platform Face-to-Face Meeting transformed into a virtual event.



Interoperability Platform

Collaboration drives the ELIXIR Interoperability Platform, reflected in the multitude and diversity of achievements in 2020.

FAIR practice guidance and implementation are critical ELIXIR-wide activities to support sustainable data – a mission many Interoperability Platform members have advocated. The year 2020 has provided several examples of these efforts to support FAIR data:

- Opened a call for new ELIXIR Recommended Interoperability Resources (RIR) and completed the review process. The results of this process were announced in 2021 with four new resources – Plaza, FAIRtracks, OmicsDI and BIII – adding bioimaging,

genomic track data, plant informatics to the already diverse set of data types that the RIRs support. The Platform will continue to review applications to fill the FAIRification gaps identified through the FAIR Service Reference Framework.

- Several standards and funding organisations have recommended using ELIXIR RIRs and flagship resources, including FAIRsharing, BioSchemas and OLS, in their guidance documents. For example, the RDA COVID-19 Recommendations and Guidelines on Data Sharing⁵, the Report on FAIR requirements for persistence and interoperability produced by the Horizon 2020 FAIRsFAIR project⁶, and

the European Commission's Reproducibility of scientific results in the EU⁷ and Open Research Data and Data Management Plans⁸.

- InterMine, another ELIXIR RIR, received funding for a collaboration with PomBase to create a cloud-based FAIR database for *Schizosaccharomyces pombe* by using the InterMine Platform on the EOSC-Life OpenStack.
- Co-led the development of the ELIXIR-CONVERGE RDMkit, scheduled to launch in 2021, to help generate research data management plans with FAIR service resources. It offers use cases and enables capacity building in research data management for Europe.

Strategic Objectives

1

4

⁵ https://www.rd-alliance.org/system/files/RDA%20COVID-19%20Recommendations%20and%20Guidelines%2C%2030%20June%202020_Endorsed-Final_o.pdf

⁶ Riungu-Kalliosaari, L., Hooft, R., Kuijpers, S., Parland-von Essen, J. & Tana, J. D2.4 2nd Report on FAIR requirements for persistence and interoperability. <https://zenodo.org/record/4001631> (2020) doi:10.5281/ZENODO.4001631.1.

⁷ https://cdn2.euraxess.org/sites/default/files/reproducibility_of_scientific_resluts_in_the_eu.pdf

⁸ https://erc.europa.eu/sites/default/files/document/file/ERC_info_document-Open_Research_Data_and_Data_Management_Plans.pdf

- Co-led the development of the FAIR Cookbook and FAIR Capability Maturity Model, as part of the FAIRplus project, to provide a best practices handbook of operations on data FAIRification.
- Tackled the research reproducibility question by connecting Computational workflows by Common Workflow Language (CWL) to the ELIXIR Galaxy Community, WorkflowHub, and RO-Crate.
- FAIRsharing signed an agreement with OpenAIRE, an open and sustainable scholarly communication infrastructure, to share and integrate metadata on databases, standards and data policies. OpenAIRE will allow a broader reach of the manually-curated and community-verified metadata from FAIRsharing.

Additionally, supporting better indexing, Bioschemas has kept improving data discoverability and interoperability. In 2020, a continued expansion of the number of resources using Bioschemas markup brought the total number to over 110. The new resources using the Bioschemas indexing strategy include IDPcentral, the ELIXIR Training Portal (TeSS), Bio.tools and the European COVID-19 Data Portal.



COVID-19 in focus

Many interoperability resources have been critical for supporting for COVID-19 research and data sharing efforts, with partners often establishing COVID-19 resources or collections in existing repositories. These include the FAIRsharing COVID-19 collection⁹, COVID-19 Disease Map using WikiPathways¹⁰, FAIRDOM Hub¹¹, the COVID-19 Workflows Hub¹², and BridgeDB¹³. Additionally, the COVID-19 Data Portal has adopted BioSchemas in their proof-of-concept to be implemented in their production.

⁹ https://fairsharing.org/collection/RDACovid19WG?q=&selected_facets=registry_exact:Database
¹⁰ covid.wikipathways.org/
¹¹ <https://fairdomhub.org/projects/190>
¹² <https://covid19.workflowhub.eu/>
¹³ Kutmon, M. & Willighagen, E. *BridgeDb: Human and SARS-related corona virus gene/protein mapping database derived from Wikidata*. <https://zenodo.org/record/4297574> (2020) doi:10.5281/ZENODO.4297574.



Tools Platform

The Tools Platform helps scientific communities find, register, and benchmark software tools and improve software sustainability and quality through best practices. These tools help researchers access, analyse and integrate biological data to drive scientific discovery across the life sciences.

In 2020, the Tools Platform focused on improving the BioContainers registry, defining a Software Management Plan, enhancing OpenEBench services, expanding the Bio.Tools registry and implementing the integrated/ interoperable infrastructure of the Tools Platform services, known as the ELIXIR Tools Platform Ecosystem.

Including over 9500 tools, the BioContainers registry has improved security by implementing a vulnerability scan report for all containers. BioContainers has also seen two other major achievements in 2020 with the GA4GH standard Tool Registry Service deployment and

publishing an article in the Journal of Proteome Research¹⁴.

The Software Best Practices Group released the first draft of the Software Management Plan for consultation by the community. They defined software stages and key stakeholders, and discussed and developed its adoption roadmap at the BioHackathon Europe 2020.

Given the difficulties of starting benchmarking, the OpenEBench service has been consolidated and extended to provide basic tool operability tests and alleviate reimplementations of abstractable workflows. The collaboration with WorkflowHub enabled the deposition of benchmarked workflows and new alliances with scientific communities, such as Quest for Orthologs, resulting in a publication in the Nucleic Acids Research Journal¹⁵.

OpenEBench and the Software Best Practices Group have continued

to work together on FAIR for Research Software indicators and the development of the Software Management Plan alongside global communities, such as RDA and ReSA.

Bio.Tools has also been prominent in the Platform's activities in 2020. Community-specific domains were created for 3D-Bioinfo, proteomics and rare diseases, and a special focus was given to tools that support COVID-19 research. The content in the registry has increased to over 10,000 tools, supported by a text-mining tool for publications using Pub2Tools, developed by ELIXIR Estonia. The Tools Platform also released the Bio.Tools schema version 3.3¹⁶.

The ELIXIR Tools Platform Ecosystem aims to create a central repository providing metadata rich, technology-agnostic software containers for its use and deployment. It will facilitate data and metadata integration, homogenisation and curation within GitHub.

Strategic Objectives

1

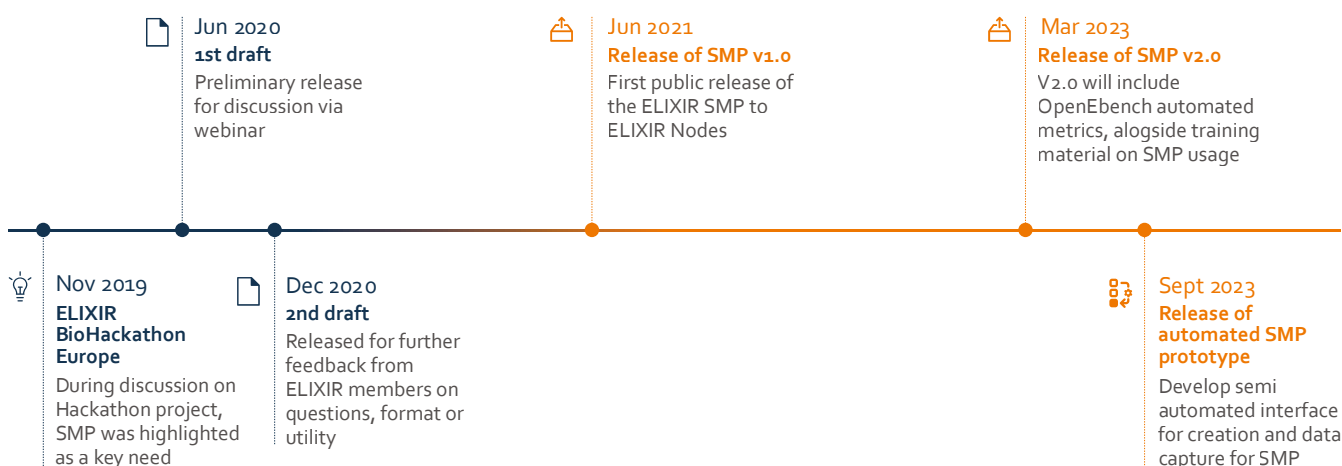
4

¹⁴ Bai, J. et al. BioContainers Registry: Searching Bioinformatics and Proteomics Tools, Packages, and Containers. *J. Proteome Res.* 0, (2021).

¹⁵ Altenhoff, A. M. et al. The Quest for Orthologs benchmark service and consensus calls in 2020. *Nucleic Acids Res.* 48, (2020) (doi.org/10.1093/nar/gkaa308).

¹⁶ Ison, J. et al. biotoolsSchema: a formalized schema for bioinformatics software description. *Gigascience* 10, (2021) (doi.org/10.1093/gigascience/giaa157).

Roadmap of the ELIXIR Software Management Plan (SMP)



The initial efforts will integrate content from Bio.tools, Biocontainers, OpenEBench and Galaxy, and, in time, facilitate the inclusion of new data and metadata producers (e.g. Bioconda, Bioconductor) and consumers (e.g. GA4GH TRS, WorkflowHub). The ambitious work that began in 2020 has already accomplished several milestones:

- Included Bio.tools in GitHub prototyping and offered BioSchema profiles for its entries.
- Automated metadata exchange between OpenEBench, Bio.

Tools, Debian Med and BioImage Informatics Index.

- Agreed on a roadmap for a sustainable Tools Ecosystem within GitHub to start in 2021.
- Established, launched and incorporated WorkflowHub into the Tools ecosystem. Currently holds 82 workflows, 41 of which are classified as Galaxy, and has partnerships with Nextflow (nf-core), Snakemake and CWL.
- Established Workflow-RO-Crate as a metadata and packaging interchange format for workflows in the Tools Ecosystem.



COVID-19 in focus

ELIXIR Nodes involved in the Tools Platform have been active in mobilising support for COVID-19 research. For example, ELIXIR Belgium developed a Galaxy tool¹ for submission of SARS-CoV-2 sequences and associated metadata to public repositories, such as the European Nucleotide Archive.

¹ elixir-europe.org/news/ENA-new-tool-COVID-19-data

The three perspectives of the Tools Platform Ecosystem





Compute Platform

The ELIXIR Compute Platform develops and integrates cloud, compute, storage, access and identity management services for the life-science research community. These allow researchers to easily access, share and analyse data and workflows from different sources using powerful shared IT facilities.

ELIXIR Nodes provide considerable cloud and compute e-infrastructure services for life science research. In 2019, this comprised 80,000 compute cores; 50,000 TB storage; and 3,100 users from the resources available directly within the ELIXIR Nodes or through their local providers. These services are typically accessed by users following national or organisational access policies, through specific international collaboration projects, or through payments.

A major 2020 highlight was the

operation of the common ELIXIR Authentication and Authorisation Infrastructure (AAI) service and its test phase implementation as the basis of the European Life Science Login (AAL) as part of the EOSC-Life project. Many research infrastructures need to manage authentication and access rights to protect confidential resources, such as samples from human patients or information about research projects. This requires providing sufficient information on who the users are (identity proofing and user authentication), whom they are representing (affiliation with a home organisation) and what resources they can access (authorisation). Researchers using ELIXIR AAI and Life Science Login will save time and reduce complexity by using their single home identity to access multiple data, tools and compute resources across life sciences and beyond. By December 2020, the

ELIXIR AAI production service had enabled 5863 users to use 102 production scientific services as a reliable and secure source of identity and entitlement.

By working closely with a range of infrastructure partners through key projects such as EOSC-Life, the ELIXIR Compute Platform is developing sustainable and supported resources for accessing, analysing and transferring data including sensitive human data. Experts from ELIXIR's AAI are leading technical and policy Working Groups to coordinate its testing and implementation with key infrastructures, including INSTRUCT-ERIC, BBMRI-ERIC, GEANT and EGI.

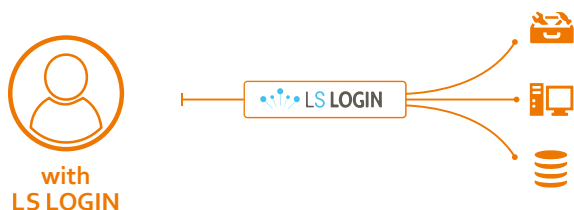
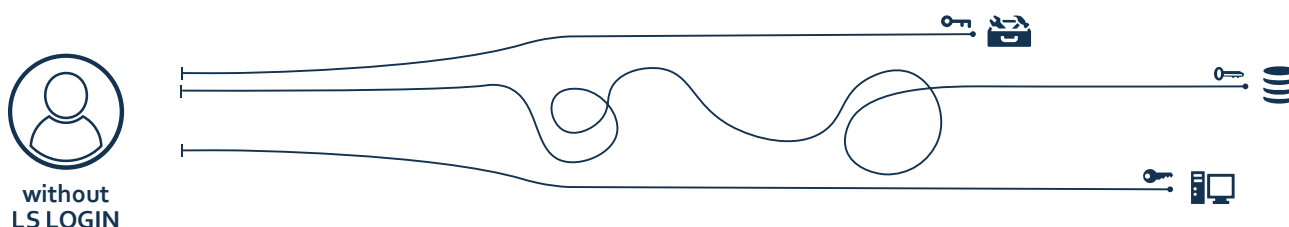
The year 2020 saw ELIXIR co-develop a series of major developments in Global Alliance for Genomics and Health (GA4GH)

Strategic Objectives

1

4

Time saved accessing resources with LS LOGIN



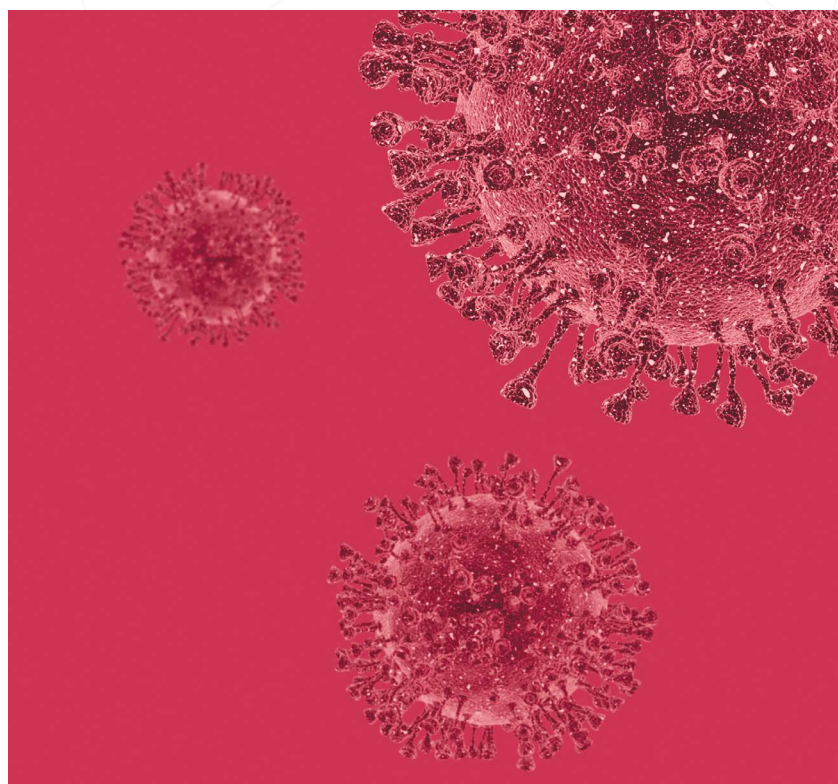
time spent

international standards for cloud computing and researcher identity, including:

- A Moonshot demonstrator of federated cloud compute for genomic data using GA4GH cloud standards across five ELIXIR Nodes showcased these developments.
- The GA4GH Passport v1.0 identity and access standard was implemented by ELIXIR and Google and demonstrated in a GA4GH Webinar series in June 2020.

The Compute Platform investigated mechanisms for data transfer between core biological data resources. These included protocols (e.g. GridFTP, FTS) and services (e.g. Reference Data Set Distribution Service (RDSDS)). New protocols such as FTS3 and htsget, a genome data specific standard, have been integrated into the Compute Platform. The Platform also explored GA4GH Data Repository Service (DRS) as an emerging data access standard with a preliminary implementation in RDSDS.

The Moonshot project to develop harmonised and interoperable API specifications to users was successfully demonstrated at the ELIXIR All Hands 2020. In a subsequent GA4GH Federated Analysis Systems Project (FASP) demonstrator, the ELIXIR Compute Platform was able to execute and manage containerised applications on cloud platforms across four different Nodes (ELIXIR-CZ, EMBL-EBI, ELIXIR-FI and ELIXIR-GR) to run an analysis on a public genomic dataset (1000 Genomes). Development started on a centralised service registry for ELIXIR-based GA4GH (Cloud) services, which is currently being deployed by SIB-Swiss Institute of Bioinformatics (ELIXIR-CH).



COVID-19 in focus

ELIXIR Compute Resources used in support of COVID-19 response

ELIXIR Nodes responded to the pandemic by providing access to compute services for COVID-19 research projects. Several of these offered access to Docker Orchestrators, including Mesos and OpenStack access, Kubernetes/OKD and GPUs where needed. Since the initiation of the COVID-19 Biohackathon, researchers across Europe were not only provided with light touch access to the compute resources but advanced technical support from Node experts in the Compute Platform to implement their analysis workflows and access databases on these resources.

Examples of compute resources made available include:

- de.NBI (ELIXIR Germany) and CSC (ELIXIR Finland) provided COVID-19 projects with priority access to cloud services.
- e-INFRA CZ (ELIXIR Czech Republic) offered relaxed access conditions to supercomputer resources, storage services and distributed compute resources.
- IFB (ELIXIR France) provided a federated set of high performance compute and cloud resources, including national and regional servers.
- The European Galaxy server, an open, web-based platform for data intensive research, gave access to compute and storage resources. It offered more than 2,500 different scientific tools, specific COVID-19 training materials, and workflows to guide users through COVID-19 data analysis.
- EMBL-EBI contributed to EMBASSY Cloud resources as detailed on the European Open Science Cloud (EOSC) Marketplace.
- ELIXIR Italy's on-demand platform, Laniakea, made available a specific Galaxy COVID-19 instance for genomic analysis.
- SIB (ELIXIR Switzerland) provided a ready-to-use slurm workload manager with a scientific software stack via the ExPASy SIB Portal.



Training Platform

2020 brought multiple challenges, yet the ELIXIR Training Platform moved swiftly to identify problems, propose solutions and adapt the training delivery process to a fully-virtual format.

Highlights of the Training Platform in 2020:

- Delivered over 160 courses through the individual ELIXIR Nodes.
- Expanded the ELIXIR Training Portal (TeSS) offer to a total of 11,900 training events since 2008 to researchers, developers and trainers.

- Increased the number of public training materials to 1,480, including 44 e-learning resources with a new quick e-learning tab in TeSS.
- Delivered four Train-the-Trainer (TtT) courses entirely online and extended the pool of ELIXIR TtT Trainers to 12.

In addition to running training activities, the Platform closely monitors the life science community's evolving bioinformatics training needs. Following ELIXIR-wide gap analysis in 2019, the Platform developed new training materials thanks to a biostatistics and machine


learning hackathon. Plans have been made for additional ones for 2021. To further identify the scientific community's training needs, the Platform engaged with ELIXIR Communities, particularly with the Galaxy and Metabolomics Communities.

Two Focus Groups also played a significant role in the Platform's activities. The Certification Focus Group took the first steps towards an ELIXIR Certification of courses and training material. At the same time, the FAIR Training Focus Group published in PLoS Computational Biology the ten

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.





simple rules for making training materials FAIR¹⁷. This work is already making training materials easier to find, reuse and adapt.

2020 also marked the start of the ELIXIR-CONVERGE project, in which twenty-two Nodes are involved in training and capacity building. This work has resulted in the delivery of twenty data management and stewardship courses using all the elements of the Training Platform: the ELIXIR Training Coordinators Group, TeSS, the Quality and Impact strategy and the ELIXIR Trainer database. Additionally, the ELIXIR-CONVERGE project used the 2019 assessment

of training needs on management and operations to develop the Node Coordinator's curriculum.

Beyond the ELIXIR network, the Training Platform maintained tight connections with many relevant initiatives at the European and global level, such as:

- **EOSC Training and Skills Working Group.** Several Platform members were heavily involved in the Working Group, delivering the report on 'Digital skills for FAIR and Open Science' with several links with the Platform, such as TeSS.

- **GOBLET – Global Organisation for Bioinformatics Learning, Education and Training.** The Platform is also fully aligned to GOBLET's efforts, with active participation in its annual general meeting.
- **The Australian BioCommons Training Group.** The Platform made the first steps towards a strategic collaboration with the Training Group of the Australia BioCommons. TeSS, the FAIR Training Focus Group and the Train-the-Trainer programme were the key elements identified as common ground.

¹⁷ Garcia, L. et al. Ten simple rules for making training materials FAIR. *PLoS Computational Biology* vol. 16 (2020) (doi.org/10.1371/journal.pcbi.1007854)

ELIXIR Communities

Bringing together experts from within particular life sciences domains, ELIXIR Communities develop targeted standards, services and training. They capture specific research domain's 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 communities' needs.

ELIXIR Communities provide us with a mechanism to reach out to defined groups of experts, including other research infrastructures of the ESFRI roadmap¹⁸. 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 will often use the networking opportunities to develop proposals for EU funding.

In 2020, there were eleven ELIXIR Communities, with the twelfth – the Food and Nutrition Community – selected at the end of the year. The selection of new Communities is based on a well-defined process:

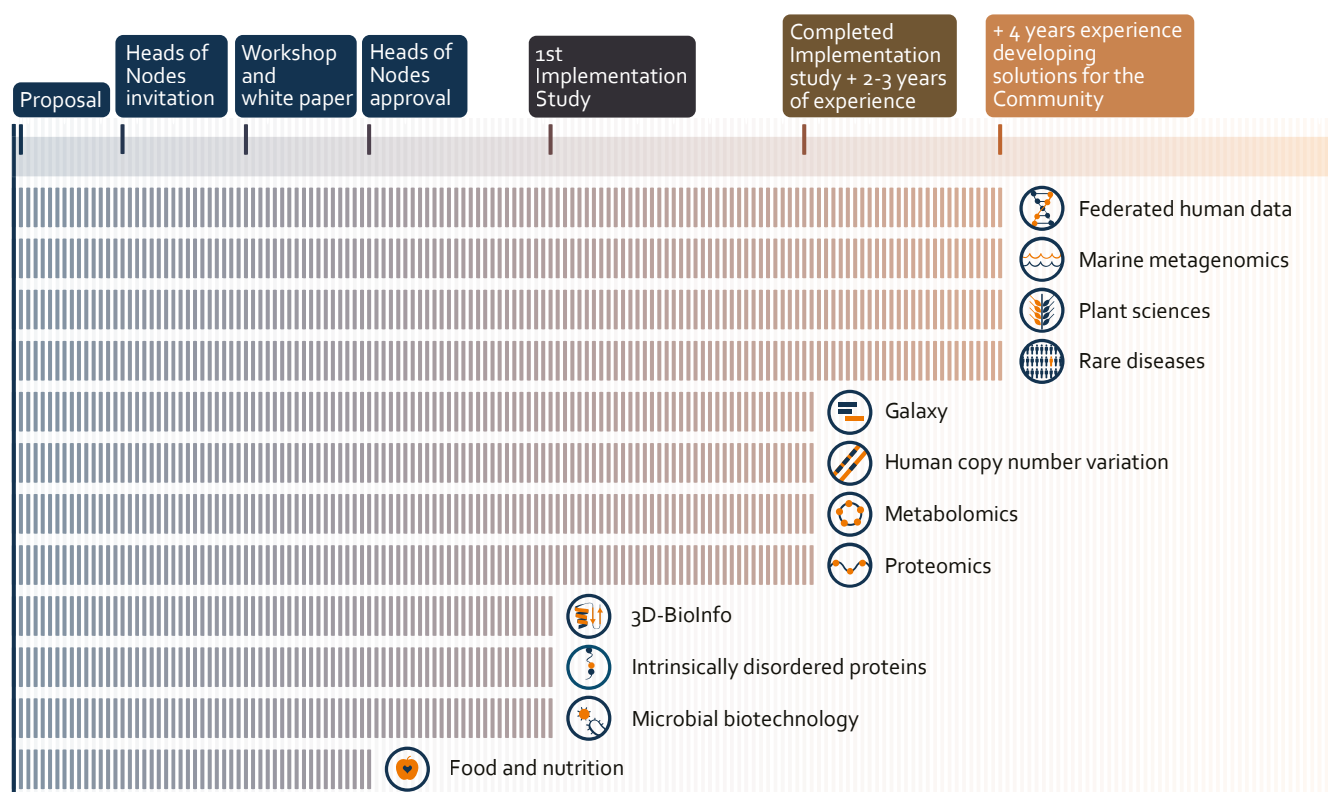
1. **Community proposal.** The new Community submits a formal proposal.
2. **Heads of Nodes invitation.** The Heads of Nodes Committee agree to the Community proposal.
3. **Workshop to establish white paper.** The nascent Community develops a white paper to establish the Community's infrastructure needs, describe the roadmap, and make the necessary connections with ELIXIR Platforms.
4. **Community establishment.** Heads of Nodes Committee agree to the white paper, which has also 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. The maturity levels are described in three groups:

1. Recently selected Communities who are about to or have just begun their first Implementation Study.
2. Communities that have recently finished their first Implementation Study and have been realising 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.

¹⁸ roadmap2018.esfri.eu/projects-and-landmarks/

The maturity journey of the ELIXIR Communities





3D-Bioinfo

ELIXIR 3D-Bioinfo Community endeavours to integrate protein structure-based data and tools and improve standardisation across Europe. In 2020, the Community conducted an extensive survey to identify gaps and areas for future activities. It also launched a community-wide Implementation Study on benchmarking criteria to distinguish physiological macromolecular interfaces from non-physiological ones.

The Community also started building a large-scale benchmark set of ligand-biomacromolecule complexes to improve the current structure-based drug design methods. One hundred parameters and descriptors of proteins, ligands or complexes' binding sites have been identified and stored in a dedicated database.

Experts from major software packages, databases and the crystallographic community have set up a uniform dictionary for nucleic acid valence geometry parameters. The resource will allow the development of tools focused on modeling, refining and validating nucleic acid structures.



Galaxy

The ELIXIR Galaxy Community has focused in 2020 on supporting the needs of other ELIXIR Communities by providing data analysis workflows to bioinformaticians and scientists. It has supported the ELIXIR Human Data Communities, and experts in chemoinformatics and imaging through online meetings and virtual hackathons.

Encouraging open science best practices is also key for the Community. In 2020, its open science dissemination efforts culminated with a webinar series of five sessions to highlight the importance of coordinated open data and infrastructure in tackling global challenges. The series brought international attention for its pandemic preparedness focus.



COVID-19 in focus

The Galaxy Community's quick response to the pandemic deserves a special mention. It has:

- Provided workflows to analyse COVID-19 data.
- Expanded the Pulsar network to enable researchers to perform analyses, such as molecular modelling.
- Made public data available in collaboration with the COVID-19 Data Portal and supported the re-analysis for diverse research purposes such as the CRG Viral Beacon.
- Added new features and tools to cope with the unprecedented amounts of data and support data sharing for researchers.



Intrinsically Disordered Proteins

The Intrinsically Disordered Proteins (IDP) Community supports the development of standards, tools and resources to accelerate the identification, analysis and functional characterisation of IDPs.

In 2020, the Community worked on solutions for the IDP annotation exchange and the implementation of the MIADE standard – the Minimum Information About Disordered Experiments standard.

In collaboration with biocurators and experimentalists, the Community has also improved and standardised the Intrinsically Disordered Proteins Ontology (IDPO). The MIADE and IDPO standards are also among the main objectives of the Proteomics Standards Initiative (HUPO-PSI) and the Intrinsically Disordered Proteins working group of the Human Proteome Organization – an international effort to promote proteomics.

By collaborating with the ELIXIR Interoperability Platform, the IDP Community has also implemented a prototype Bioschema profile for region based annotations. Bioschemas markup is already integrated into MobiBD, DisProt and PED databases, and its adoption by other IDP resources is one of the objectives of the first Implementation Study of the IDP Community, which will start in January 2021.



Marine Metagenomics

Microbiome research is expanding rapidly, providing ever more thorough insights into the 99% of microbes that are yet to be cultured (and sequenced), their organisation into communities and their adaptation to various environments. There is also an increasing interest in how these microbes can be harnessed to rescue a perturbed community or how the proteins they encode can be utilised for medical or biotechnological applications.

The ELIXIR Marine Metagenomics Community recognises that most of the tooling and resources used for the sharing and analysing microbiome data are generalisable, regardless of the biome. Therefore, in 2020, the Community began expanding to encompass a comprehensive range of biomes covering environmental habitats and engineered environments. This broader approach aims to develop a sustainable bioinformatics infrastructure for microbiome resources to deeply understand the function and taxonomy of viruses, eubacteria, archaea and single cellular eukaryotes.

Furthermore, microbiome research is starting to employ a wider range of omics techniques, such as proteomics and metabolomics. This Community will remain focused on nucleotide techniques, and only consider the integration of the outputs from other omics techniques represented by other Communities. This expansion process is expected to be finalised in 2021.



Metabolomics

In 2020, the Metabolomics Community continued working on two Implementation Studies on metabolite identification and fluxomics

The Community has organised two workshops: metabolomics-dedicated workshop at the virtual ELIXIR All Hands and a joint one for the two Implementation Studies in January 2020. At this joint workshop, the Community discussed the outcomes of the fluxomics tools gap analysis and proposed a standard methodology for depositing fluxomics data in the MetaboLights database.

There has been a marked increase in submissions to EMBL-EBI MetaboLights database from the Metabolomics Community. EMBL-EBI outreach activities have greatly increased the availability of FAIR standard data for the benefit of the wider metabolomics community.

Together with ELIXIR Sweden, the Community also helped develop the MetaboIGNITER, a start-to-end workflow that leverages well-established tools to provide an automated, highly customisable metabolomics data pre-processing pipeline.

Interoperability of data and results has also been widely discussed in the Community for the challenges of integrating metabolomics data with other studies or other omics approaches.



Microbial Biotechnology

Microbial biotechnology designs and builds new strains of microbes that can produce chemical building blocks, food supplements, novel medicines and alternatives to fossil fuels.

However, microbial biotechnology has no formal framework to manage and manipulate strains, samples, knowledge, data and metadata. The Design-Build-Test-Learn (DBTL) cycle provides a conceptual framework for developing tailor-made microbes and biological systems.

Aiming to narrow the gap, the ELIXIR Microbial Community has taken the DBTL cycle as a starting point for defining a formal representation of enzymes and metabolic networks. Through an Implementation Study, initiated in June 2020, the Community strives to achieve the objectives through technical definitions and community-oriented workshops. Agreed standard representations will build the foundations of the proof of concepts while ensuring alignment with relevant ELIXIR Communities, such as Metabolomics, Proteomics, 3D-Bioinfo and Marine Metagenomics.



Plant Sciences

In 2020, the ELIXIR Plant Community formalised its 2020-2023 action plan with a recently published roadmap¹⁹, and made noteworthy advances.

On data standardisations, the Community published a paper²⁰ on the MIAPPE standards – the Minimum Information about Plant Phenotyping Experiment data standard – and improved its outreach with broad communications and training sessions. The MIAPPE standards were also used to improve several data management tools, such as FAIRDOM/SEEK, pISA-tree, with MIAPPE compliance for better metadata. The effort aimed to tackle one of the main challenges of the research data lifecycle, the data capture with exhaustive metadata.

Throughout the year, the Community continued a fruitful collaboration with EMPHASIS and CGIAR, on Breeding API and general data sharing activities. They also continued working on improving tool interoperability through genotyping data brokering service bundles and ISA-centred API compatibility.



Proteomics

The Proteomics Community finished in 2020 the Implementation Study on 'Crowd-sourcing the annotation of public proteomics datasets to improve data reusability', led by EMBL-EBI.

In that context, the Community also developed SDRF-Proteomics, a file format for capturing sample and experimental design metadata for proteomics approaches. It is based on the Sample to Data RelationShip (SDRF) format developed by the EMBL-EBI Functional Genomics team for transcriptomics data.

The current specification document is freely available on a GitHub page alongside a list of manually curated datasets from PRIDE (Proteomics IDentification Database). The Community is also developing related tools to support the format. SDRF-Proteomics files can already be submitted and visualised in the PRIDE web interface.

Additionally, the Community has continued their work on the Implementation Study on 'Comparison, benchmarking and dissemination of proteomics data analysis pipelines', led by ELIXIR Denmark. They have also agreed on initial metrics for benchmarking purposes and developed several pipelines implementing popular proteomics data analysis workflows, such as peptide/protein identification or protein quantification.

¹⁹ Pommier C, Gruden K, Junker A et al. ELIXIR Plant sciences 2020-2023 Roadmap [version 1; not peer reviewed]. *F1000Research* 2021, 10(ELIXIR):145 (document) (doi.org/10.7490/f1000research.1118482.1)

²⁰ Papoutsoglou, E. A. et al. Enabling reusability of plant phenomic datasets with MIAPPE 1.1. *New Phytol.* 227, (2020). (doi.org/10.1111/nph.16544)

Human Data Communities

In 2020, the ELIXIR Human Data Communities (HDCs) expanded their portfolio to include two new Focus Groups on cancer and health data and held monthly meetings to discuss cross-cutting themes.

All activities within the Human Data Communities had a shared focus on establishing the coordination of the B1MG project to support the implementation of the EU Declaration for accessing one million human genomes across Europe by 2022. Yet, individually, each activity of the HDCs has achieved several goals and milestones.



ELIXIR Federated Human Data Community

The establishment of the Federation of the European Genome-phenome Archive (FEGA) has been the core focus of this Community in 2020. The priorities have been to:

- Identify the five first wave Nodes – Finland, Germany, Norway, Spain and Sweden;
- Set the technical infrastructure; and
- Plan the contractual service agreements.

The work behind these efforts included prototyping a unified and secure data deposition interface for files and metadata. The interface uses a shared quality control, validation, authentication and authorisation framework to ensure privacy and security.



ELIXIR Rare Diseases Community

In 2020, the Rare Diseases Community delivered an integration across the European Genome-phenome Archive (EGA) and RD-Connect Genome Phenome Analysis Platform (GPAP) for the visualisation of data archived at EGA via a genome browser at Genome Phenome Analysis Platform (GPAP).

The Community also focused on FAIR data by publishing the requirements for FAIR metrics assessment in the rare diseases domain. This was defined in conjunction with representative stakeholders and demonstrated through deployment on representative data sources.

Together with the European Joint Programme for Rare Diseases (EJP-RD), the Community also ran a survey for e-learning requirements to present a targeted training offer in 2021.



ELIXIR human Copy Number Variation

The ELIXIR human Copy Number Variation (hCNV) Community continued to build and maintain a list of services to manage and analyse hCNV data. The TransBioNet project, led by ELIXIR Spain, also used the list to benchmark some of these services in clinical data analyses.

The Biohackathon Europe 2020 became a landmark event for the Community. At the event, their collaboration with the ELIXIR Galaxy Community was key to providing integrated computational workflows for hCNV variant calling and extending the GA4GH Beacon protocol for hCNV data.

ELIXIR Cancer Data Focus Group

Established at the start of 2020, this Focus Group gathered experts from seventeen different ELIXIR Nodes and identified ten use cases for cancer data management in bioinformatics infrastructures. Over the course of the year, the group invited six clinicians to discuss these use cases and other infrastructure challenges seen in real-world settings. This discussion will be the substrate for the group's work in 2021.

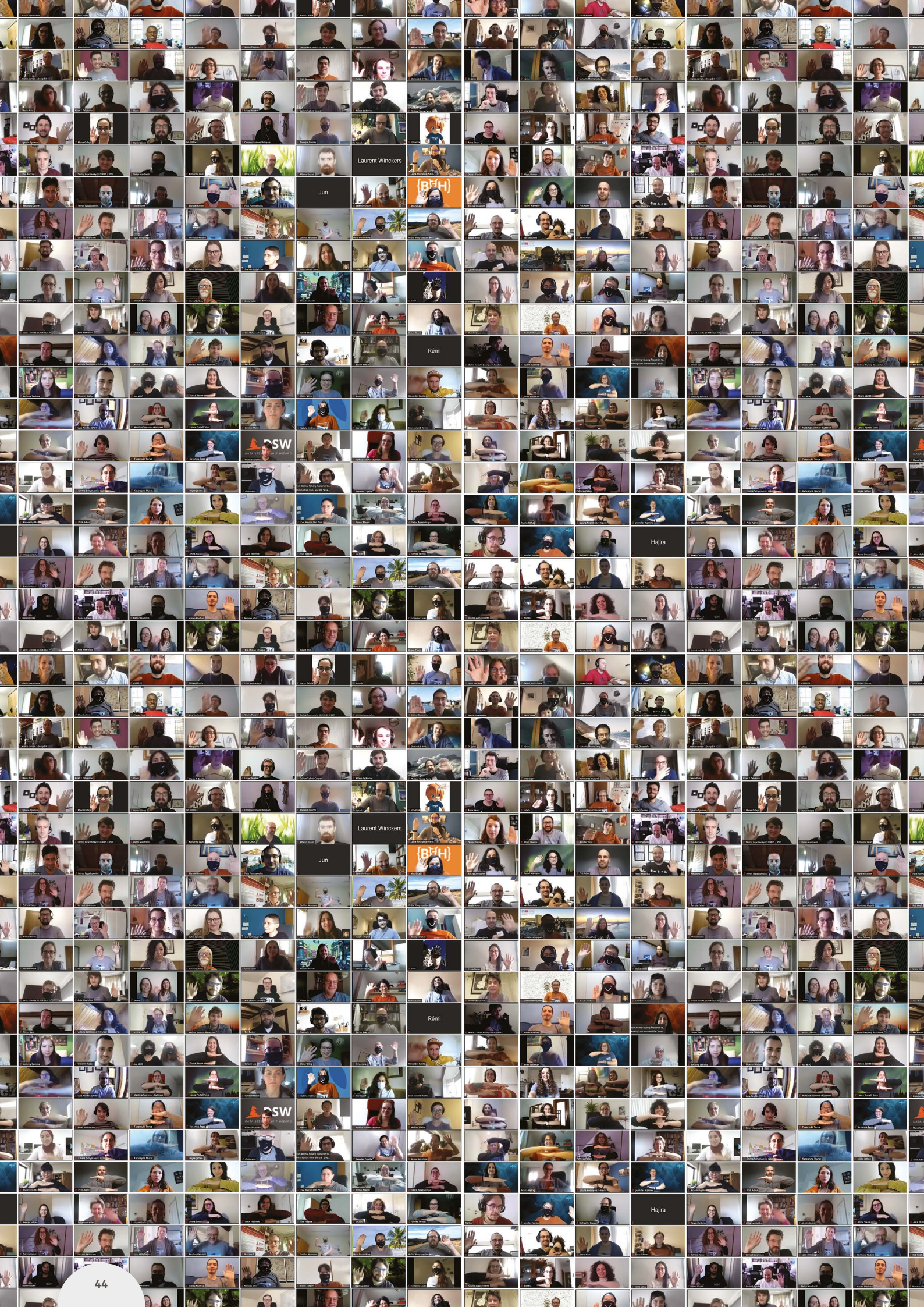
ELIXIR Health Data Focus Group

Convened in February 2020, the Health Data Focus Group was formed by members from eight ELIXIR Nodes and three external organisations. They aim to discuss health data management and their secondary use for research purposes. By identifying common challenges, the Focus Group has built the foundations of their 2021 plans to enhance the use of health data in research across all ELIXIR Nodes. The group has also collaborated with the ELIXIR Galaxy Community and the UK Node of BBMRI to submit a proposal on the use of Synthea for the creation of synthetic data, as part of the EOSC-Life Digital Life Sciences Open Call.

ELIXIR Beacon project

Throughout the year, the ELIXIR Beacon collected the requirements from diverse stakeholders (ELIXIR Nodes, GA4GH Driver Projects, key Horizon 2020-funded projects) to build the Beacon API specification v.2.0. The development of the specification has positively progressed and has gathered interest from additional potential stakeholders such as the Health-RI initiative in the Netherlands. The Global Alliance for Genomics and Health has kept a close eye on the developments, and plans to approve the specification in 2021.

Additionally, ELIXIR identified ELIXIR Beacon and the Beacon Network as a potential ELIXIR Infrastructure Service.



Our flagship events

All Hands Meeting – connecting all ELIXIR members

The ELIXIR All Hands meeting is an annual event that brings together all members of ELIXIR to discuss the achievements and future plans of the research infrastructure.

In June 2020, all plans had to be quickly adapted to an unexpected remote setup. ELIXIR partners responded with enthusiasm and took advantage of this opportunity, allowing more individuals to join. With almost no limitations in the number of attendees, the All Hands hosted almost 500 participants. Our first virtual All Hands, and first big online conference, was a success.

Across three different days, the sixth ELIXIR All Hands meeting hosted a series of plenary sessions, workshops and mini-symposia, where participants reviewed ELIXIR's achievements, activities and the plans for the future.

The yearly gathering also hosted two open keynote sessions. With Dimitris Koureas, from the Naturalis Biodiversity Center in the Netherlands, ELIXIR members explored museum specimens with a keynote under the title of *Electric Butterflies, from field to bytes and back*.

With Janet Kelso, from the Max Planck Institute for Evolutionary Anthropology in Germany, the audience ventured into the field of archaic genomics²². Her talk on the history of modern and archaic human groups has generated a great impact on the ELIXIR Youtube Channel with more than 20,000 views. It has significantly boosted ELIXIR's visibility since our videos usually average a hundred views. This number is significantly higher compared to other well-established life science channels – more than ten times higher.

23

ELIXIR NODES

501

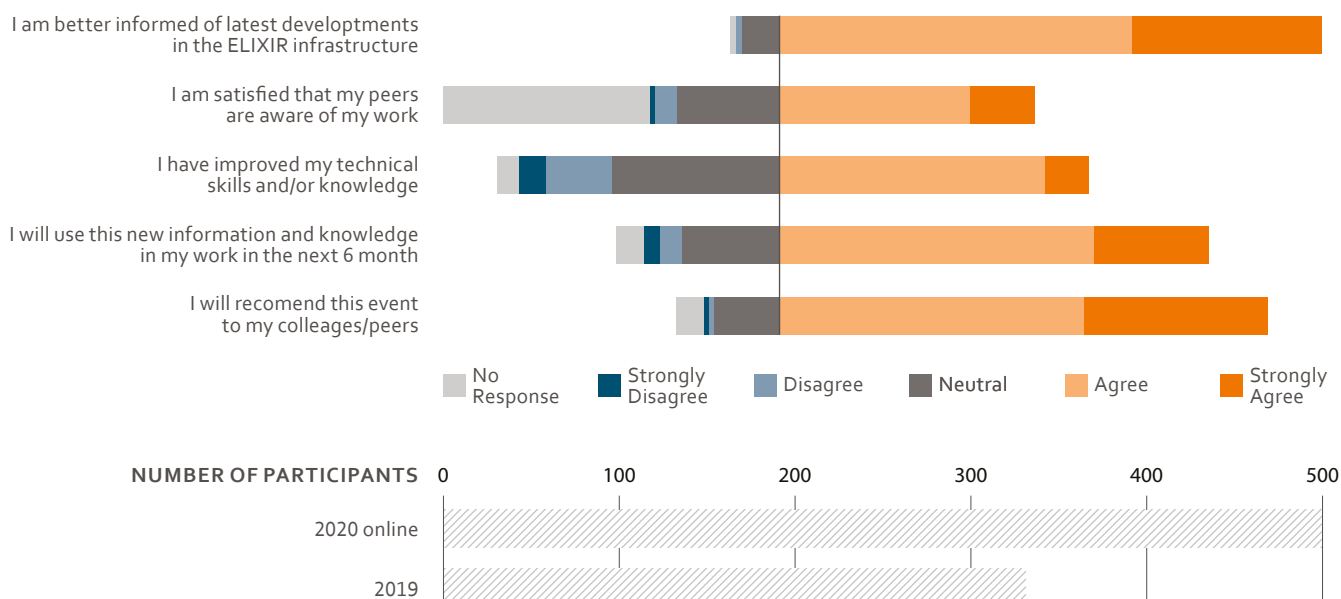
PARTICIPANTS

20K

ON YOUTUBE FOR
JANET KELSO'S
KEYNOTE

²² <https://elixir-europe.org/events/all-hands-keynote-i>

Survey responses from the All Hands 2020



BioHackathon Europe – accelerating results in bioinformatics

Strategic Objective

5

2020 was an unprecedented and yet successful year for BioHackathon Europe. Over 300 people attended the first virtual BioHackathon Europe and the third of its kind. The intensive week of hacking brought together experts from all backgrounds – bioinformaticians, software developers and trainers – to improve bioinformatics tools and resources.

Despite the unique challenges, the collaborative nature of BioHackathon Europe was successfully replicated as a virtual event. Participants hailed from across all time zones, across all our Nodes, Platforms and Communities, from Europe to Japan to the US and Australia, leveraging online tools to work on 40 bioinformatics projects.

The variety of this year's projects was remarkable – from infectious diseases to biodiversity – aligning with ELIXIR's core interests around the life sciences.

Industry also played an important role in 2020, being the first year the industry actively led projects in the BioHackathon. Atos-Bull made their BullSequana Edge capabilities available for the duration of the BioHackathon.



COVID-19 in focus

As one might expect, COVID-19 and pandemic preparedness was the focus of many projects. One such project aimed to produce a service bundle for epidemic response. The team was able to include 143 tools, tailor several concept maps to structure analysis and build a website²¹.

“

It was an amazing experience, very interactive and felt very close to a real life event. Two teams embraced our challenge to demonstrate the validity of edge computing. We reached our goal of connecting with the bioinformatics community and creating awareness for a new technology. No doubt we would participate in the future.

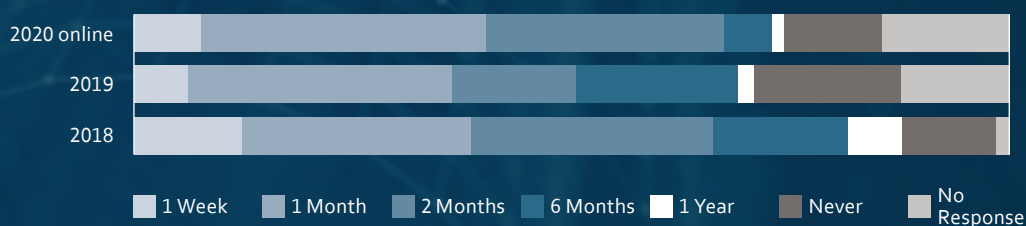
Natalia Jiménez Lozano

Director
Atos HPC

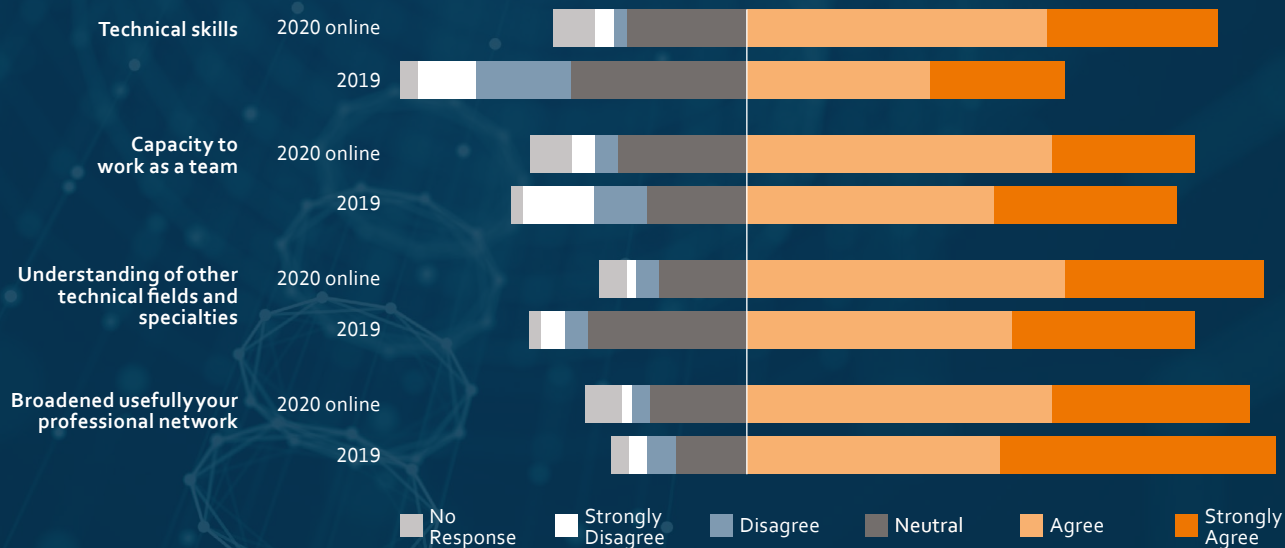
²¹ SB4ER – sb4er.github.io/projects.html

Survey responses

Without the Biohackathon how long would it have taken a single person to reach the same outcome? (%)



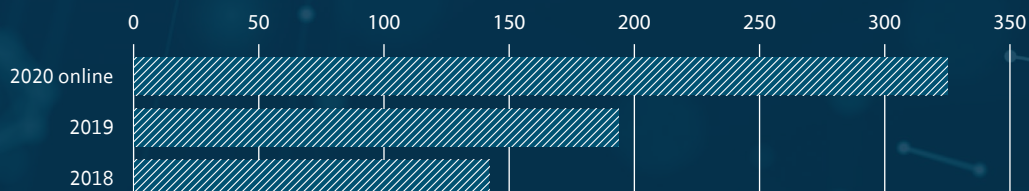
YOU HAVE IMPROVED YOUR (%)



WOULD YOU RECOMMEND THIS EVENT TO YOUR COLLEAGUE? (%)



NUMBER OF PARTICIPANTS



SURVEY RESPONSES (%)



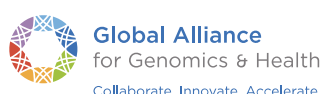
Our collaboration, outreach and industry support efforts

Strategic Objective

3

Strengthening global collaboration

ELIXIR-GA4GH strategic partnership



Across 2020, ELIXIR continued its collaboration with the Global Alliance for Genomics and Health (GA4GH). The partnership achieved several objectives:

- Expanded the partnership with several new projects, such as EUCANCan, European Joint Programme on Rare Disease, Genomics England, and Swiss Personalized Health Network.
- Kick-started working on the Capability Maturity Model for GA4GH products to describe and promote its adoption to ELIXIR members.

Collaboration strategy with the Australian BioCommons



Gaining maximum visibility across ELIXIR, the Australian BioCommons Collaboration Strategy was launched at the ELIXIR All Hands meeting in June 2020. The ambitious collaboration aims to demonstrate mutually-beneficial activities and identify arrangements for long-term cooperation.

Geographical barriers and travel restrictions did not stop this collaboration. Multiple interactions have taken place throughout 2020, some building on existing activities such as Galaxy, where Australian BioCommons engaged in a Community-led ELIXIR Implementation Study.

Elsewhere in ELIXIR, the Australian BioCommons co-leads a task in the FAIR Training ELIXIR Focus Group and is currently working alongside the Training Platform to roll-out a Train-the-Trainer programme in Australia.

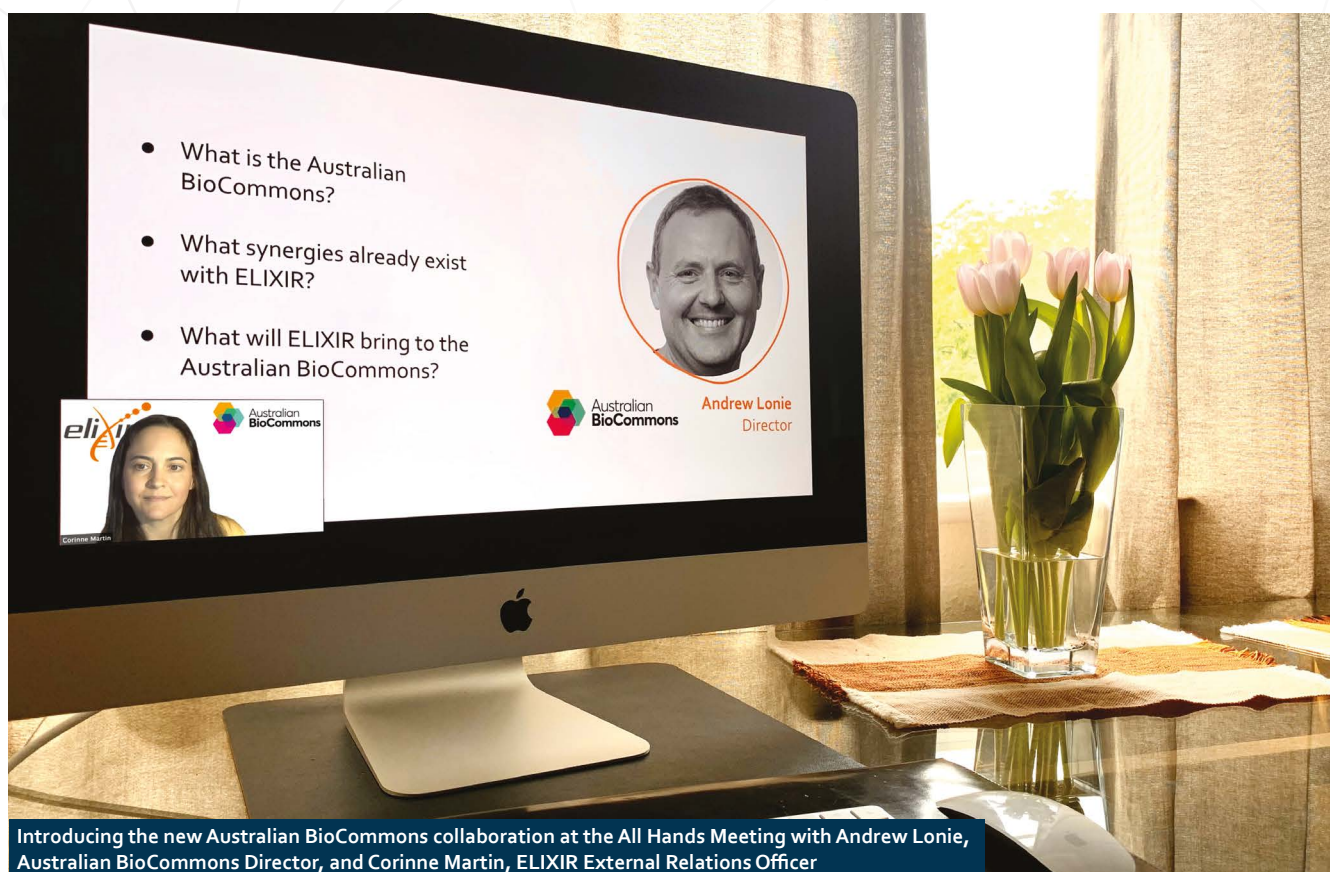
In 2021, the collaboration will continue to expand, for instance, under the Tools Platform and in the area of impact evaluation with the Australian BioCommons as an active participant of a dedicated ELIXIR Strategic Implementation Study.

Global Biodata Coalition collaboration



ELIXIR has continued to support the establishment of the Global BioData Coalition (GBC). This new international effort enables collaboration between international funders to support global biodata resources. The GBC focusses on selecting a set of global core data resources for the life sciences.

Through technical, scientific and outreach activities, ELIXIR has supported the ambitions of the Coalition's Board of national and charitable funders. In the future, ELIXIR is expected to make a significant contribution, both strategic and procedural, to identify Global Core Data Resources, building on the foundations of the ELIXIR Core Data Resources.



Strategic Objective

5

Investing in people

Many ELIXIR Nodes' staff cover several roles and need proficiency in multiple areas for their day-to-day tasks. Staff frequently manage communications, industry engagement and impact assessment tasks, in addition to their research and Node coordination work. Many staff in ELIXIR Nodes have also transitioned from their research roles into these new, unexplored roles for them.

To support the diverse needs of Node staff, ELIXIR has established Focus Groups that meet regularly. These exchange best practices, experiences or knowledge through small meetings, workshops and other internal communications channels. In 2020, ELIXIR strengthened the Innovation and Industry and the Impact Focus Groups and re-established a Communications one. All three have identified knowledge gaps that will set

the foundations for capacity building activities in the coming years.

As part of the ELIXIR-CONVERGE project, ELIXIR has also offered a Workshop Series on Impact, Industry and Communications that started in December 2020 with a session on stakeholder mapping and landscape analysis. The series will continue to deliver training in 2021, focusing on impact assessment, outreach to funders and stakeholder communications.

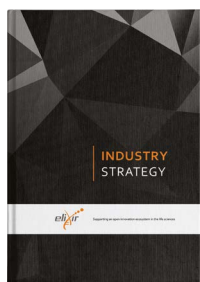
Investing in people's skills lies at the core of ELIXIR. In 2020, ELIXIR continued to support and endorse the Executive Masters Programme for the Management of Research Infrastructures (EMMRI), offered by the University of Milano-Bicocca. ELIXIR covers the course fees and other related expenses to allow

Hub and Node staff to improve their knowledge in research infrastructure management.

The unexpected travel restrictions and remote working in 2020 did impact the outcomes of the EMMRI, yet some ELIXIR Nodes successfully obtained the EMMRI Certificate of Excellence in Research Infrastructure Leadership. This Certificate is offered to Nodes rather than individuals to provide the right combination of competencies for different managers to hand-pick those modules most relevant to their work. Both Ana Portugal Melo (ELIXIR Portugal) and Christine Stansberg (ELIXIR Norway) presented their field projects with the support from an ELIXIR Staff Exchange Project, in which ELIXIR Italy (Francesca de Leo) also participated.

Enabling and supporting industry collaboration

Paving the way for an open innovation ecosystem with a revamped Industry Strategy



2020 saw the publication of an updated ELIXIR Industry Strategy. The revamped strategy focuses on creating an open innovation ecosystem driven

by collaboration between industry and ELIXIR Nodes. As a long-term objective of Strategy, ELIXIR will aim to increase industry usage and integration of the public life science resources provided by ELIXIR Nodes, leading to new services, products and economic growth.

Facilitating collaborations

Collaboration is supported through initiatives such as the ELIXIR Knowledge Exchange Scheme, which offers financial support to ELIXIR Nodes working on collaborative projects with industry partners. In 2020, two pilot projects were planned to go ahead but travel restrictions imposed by the COVID-19 pandemic reshaped the scope of the projects. Despite the limitations, ELIXIR Knowledge Exchange Scheme provided us with some outcomes.

ELIXIR Luxembourg, in collaboration with a UK based company, Petagene, held a virtual workshop on best practices for storing genomic data. Additionally, ELIXIR Portugal progressed in their project that will see the ELIXIR plant data infrastructure brought to the Portuguese pulp and paper industry.

Demonstrating impact

Following up from the report on 'open data resources as a business model for SMEs', ELIXIR has engaged in a research collaboration with an economist. Throughout 2020, work was done to explore the innovation ecosystem of open life science data in greater depth. The initial results of the study were presented at the European Open Science Forum in a virtual session in the Science to Business Programme, visited by 195 attendees. The final study report will be published in 2021.

Engaging with the private sector



ELIXIR supports its members in connecting to industry experts through networking events, such as the Innovation and SME Forums. Due to the COVID-19 pandemic many planned face-to-face events had to be postponed until 2021 or held virtually.

The ELIXIR Bioinformatics Industry Forum (EBIF) is a great example of adaptation to a virtual format. In 2020, with sustained travel restrictions and lockdowns, the EBIF 2020 moved virtual, bringing together bioinformaticians from numerous countries to learn and discuss the latest research advances and visionary ideas.

The event was divided into a series of five different sessions to cover diverse themes, bringing together 480 attendees from around the globe. Inspired by the voices of multiple experts in the field of bioinformatics, all virtual sessions tackled the latest challenges in the wider bioinformatics community. Academia and industry experts collaborated in different sessions to explore hot topics such as artificial intelligence, machine learning or quantum computing applications. The interactive nature of the events allowed participants to feedback directly on questions that will help develop ELIXIR's service portfolio in the future.

“

The first online industry forum by ELIXIR provided a fresh new approach to connect bioinformatics expertise from industry, academia and nonprofit organisations at a time when COVID-19 challenges our normal social exchange. From community building, reproducibility and biomedical applications in machine learning to open-source software initiatives and quantum computing, a wide and diverse range of bioinformatics ventures bloom thanks to ELIXIR.

Sergio Martinez Cuesta
Senior Bioinformatics Data Scientist,
AstraZeneca

Demonstrating ELIXIR's impact

Since 2011, ELIXIR has supported the publication of more than four hundred peer-reviewed articles and preprints. These relate to the development and operation of bioinformatics resources from databases to training – highlighting ELIXIR's scientific legacy. Carefully monitoring ELIXIR-related publications and continuing to encourage them is an indispensable exercise to understand our impact.

Text-mining ELIXIR's impact

In 2020, ELIXIR put in place a system to identify all its publications and help understand one of the aspects of ELIXIR's impact. This monthly-updated system uses text-mining to search for terms relating to funding streams and other achievements linked to ELIXIR. The search is carried out in EuropePMC – one of ELIXIR's Core Data Resources – and is carefully curated to remove false positives.

The resulting visualisations show that, year by year, ELIXIR publications increase in number and citations.

One of the well cited publications is DisGeNET, a comprehensive platform integrating information on human disease-associated genes and variants²³. It acknowledges funding from ELIXIR-EXCELERATE, the project that was instrumental in building the ELIXIR of today. Another well-cited publication related to the PRIDE proteomics database²⁴, which acknowledges three ELIXIR Implementation Studies.

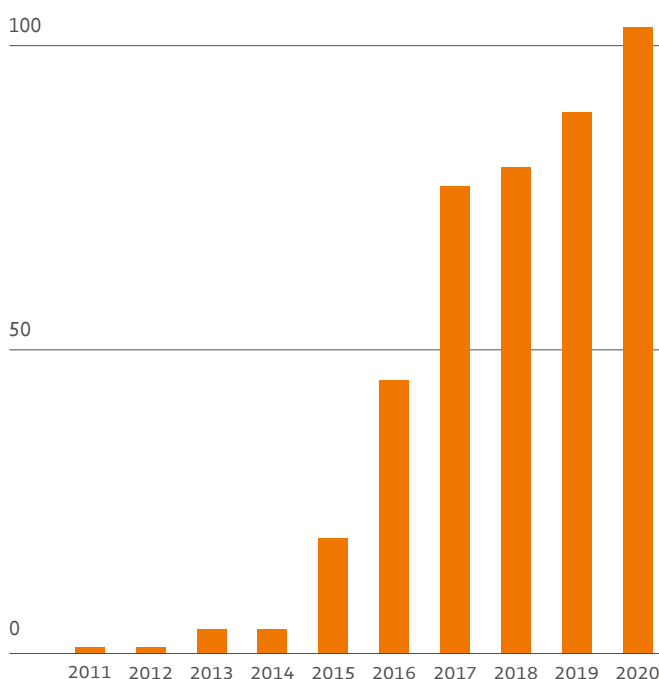
In 2021, ELIXIR will continue to use a range of approaches to assess scientific and socio-economic impact. ELIXIR will endeavour to enable policy- and decision-makers to access and visualise key indicators and other related evidence.

23 1. Piñero, J. et al. The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res.* 48, (2020) (<https://doi.org/10.1093/nar/gkw943>)

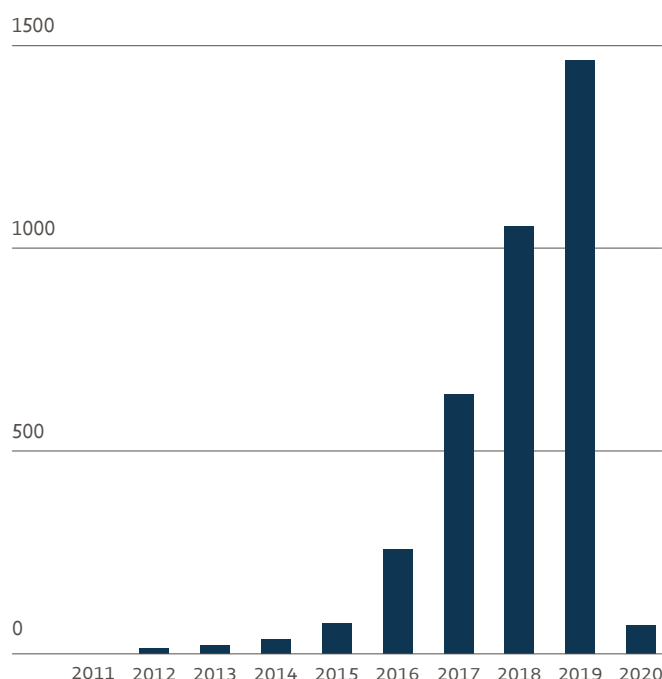
24 Perez-Riverol, Y. et al. The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.* 47, (2019) (<https://doi.org/10.1093/nar/gky1106>)

Scientific legacy of ELIXIR-related funding

PUBLICATIONS



CITATIONS



Empowering Nodes with ELIXIR F1000Research Gateway

Created in 2015, ELIXIR's Gateway in F1000Research provides a portal for publishing general or peer-reviewed ELIXIR-related work. From research articles to papers, roadmaps and slides, it is a 'rite of passage' for newly established ELIXIR Communities to publish a white paper to achieve their official status.

As a service to all ELIXIR partners, ELIXIR Hub budget covers the article processing fee to publish peer-reviewed articles in the Gateway. In 2021, the ELIXIR Hub will continue to offer this support and encourage partners to keep using the service.

Reviewing publication metrics

In 2020, some new publications gained significant visibility, such as a software article on g-profiler²⁵, an ELIXIR Recommended Interoperability Resource. This ELIXIR Estonia article has obtained more than 1,000 views. An ELIXIR Germany (de.NBI) article also ranked at the top of the most viewed papers. This opinion article²⁶ on bioinformatics training also achieved more than 1,000 views.

Alongside this highly-visited articles, other publications also deserve special mention, such as the 3D-Bioinfo Community proposal to integrate structural bioinformatics activities in ELIXIR²⁷, an article on the ELIXIR Implementation Study on human genomic variations²⁸ and the ELIXIR hCNV Community white paper²⁹.

Five years on, the Gateway contains:

193

POSTERS

53

SLIDE DECKS

20

DOCUMENTS

36

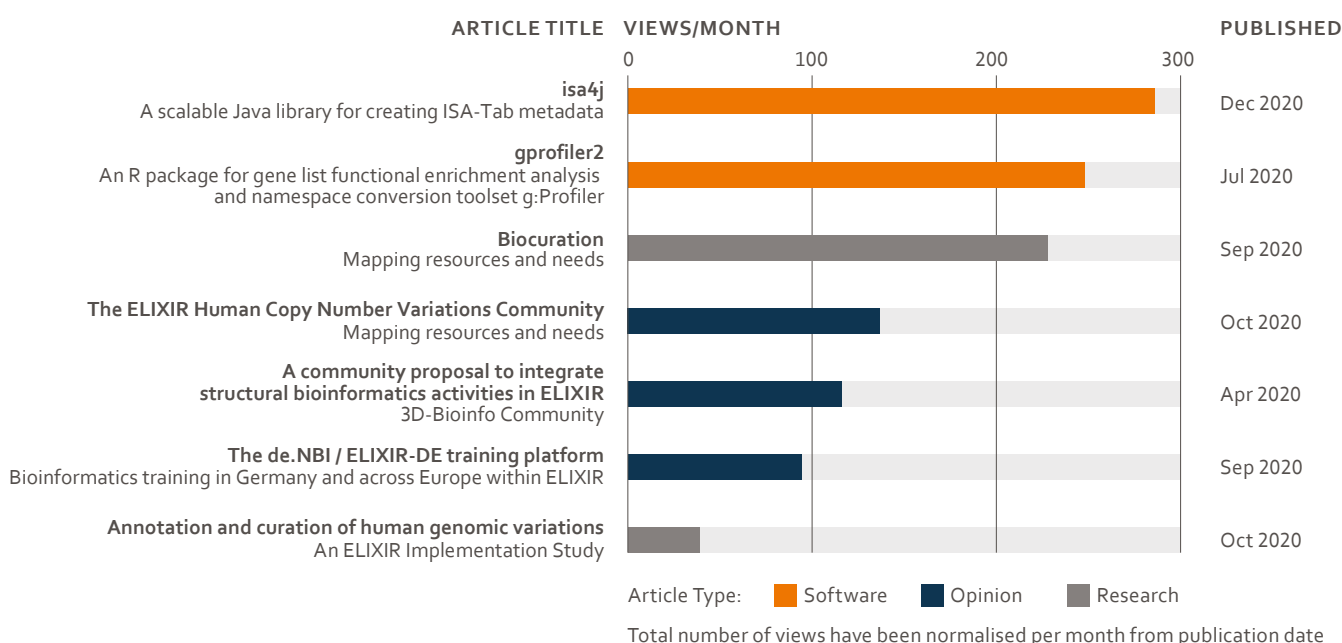
PEER-REVIEWED ARTICLES

360+

CITATIONS COLLECTIVELY GATHERED

- 25 Kolberg L, Raudvere U, Kuzmin I et al. gprofiler2 – an R package for gene list functional enrichment analysis and namespace conversion toolset g:Profiler [version 2; peer review: 2 approved]. *F1000Research* ELIXIR (2020) (<https://doi.org/10.12688/f1000research.24956.2>)
- 26 Wibberg D, Batut B, Belmann P et al. The de.NBI / ELIXIR-DE training platform – Bioinformatics training in Germany and across Europe within ELIXIR [version 2; peer review: 2 approved]. *F1000Research* ELIXIR (2020) (<https://doi.org/10.12688/f1000research.20244.2>)
- 27 Orengo C, Velankar S, Wodak S et al. A community proposal to integrate structural bioinformatics activities in ELIXIR (3D-Bioinfo Community) [version 1; peer review: 1 approved, 3 approved with reservations]. *F1000Research* ELIXIR (2020) (<https://doi.org/10.12688/f1000research.20559.1>)
- 28 David A, Barbié V, Attimonelli M et al. Annotation and curation of human genomic variations: an ELIXIR Implementation Study [version 1; peer review: 1 approved with reservations]. *F1000Research* ELIXIR (2020) (<https://doi.org/10.12688/f1000research.24427.1>)
- 29 Salgado D, Armean IM, Baudis M et al. The ELIXIR Human Copy Number Variations Community: building bioinformatics infrastructure for research [version 1; peer review: 1 approved]. *F1000Research* ELIXIR (2020) (<https://doi.org/10.12688/f1000research.24887.1>)

A visualisation of the articles published in the ELIXIR F1000R Gateway in 2020.





COVID-19 in focus

Unveiling ELIXIR COVID-19 efforts

Across 2020, all ELIXIR Nodes have worked relentlessly to support COVID-19 research with tools, compute, standards, platforms, collections or workflows. These efforts have been reflected in the number of ELIXIR publications in 2020, offering an impactful image of the active role of all ELIXIR partners in enabling scientific research.

30

COVID-19 PUBLICATIONS FROM ELIXIR NODES

ELIXIR also given visibility to these ELIXIR resources to reach the wide scientific community working on COVID-19 research. Our COVID-19 resources page on ELIXIR's website was one of our main efforts in making ELIXIR Node resources visible. The dedicated page was featured as a highlight receiving the highest number of visits on the ELIXIR website – nearly 20,000 views.

Sustaining collaboration with virtual events

In 2020, the ELIXIR Hub was involved in the organisation of 69 events to cater to the needs of externally-funded projects, ELIXIR Platforms, Communities and Focus Groups. The events ranged from governance-related events to industry-related ones, and received almost 2,500 registrations across all event types, peaking at more than 500 for the annual ELIXIR All Hands Meeting and 350 for the ELIXIR BioHackathon Europe.

It is undeniable that the move to an all-virtual format brought significant technical and cultural challenges. Yet, many have reported that it has made attendance to specific meetings easier for those who might not have been able to travel in normal (pre-pandemic) times.

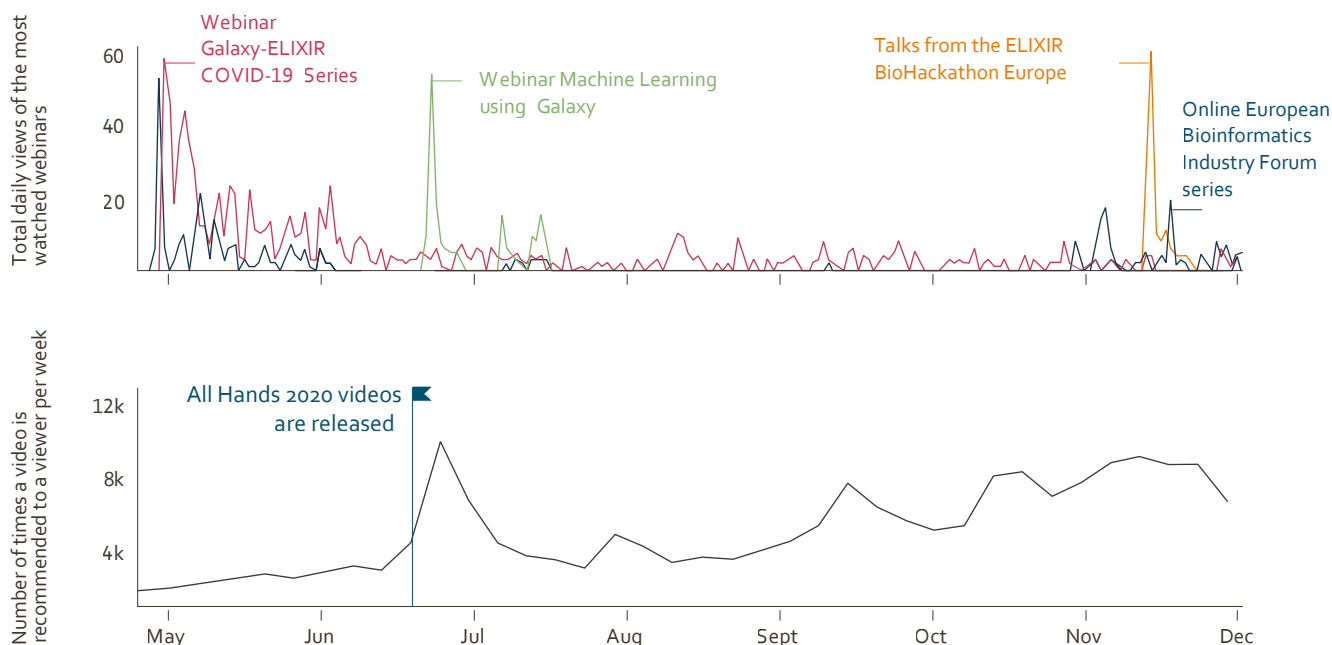
In addition to regular events and meetings, ELIXIR ran 27 different webinars, where COVID-19 became one of the main topics. Already in April 2020, the Galaxy Community swiftly responded to our audience's needs with a comprehensive webinar series on FAIR and open data to tackle COVID-19. The practical series called

the attention of more than a hundred participants per webinar. Its success resulted in the proposal of two further Galaxy COVID-19 webinar series for more advanced users in 2021.

While COVID-19 was the imperative in many conversations, other research areas were never neglected. ELIXIR took advantage of the unexpected remote situation to empower its partners and the global research community. In these short yet intense online sessions, experts talked about machine learning, FAIR training materials, cloud computing or systems biology modelling amongst many others.

2020 has taught us the power of these short webinars – many joined us from countries such as South Africa, Kenya, Saudi Arabia, India, Canada and Australia. Although 2021 might bring normality back to the workplace, ELIXIR aims to continue supporting the portfolio of webinars to keep sharing knowledge and connecting experts across the globe.

The most viewed webinars on the ELIXIR YouTube Channel in 2020 and the growth of the channel.



Our people

ELIXIR Hub staff

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

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

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

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,

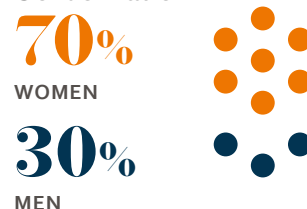
Arshiya Merchant joined the Human Genomics and Translational Data team to support the coordination efforts in human data. At the same time, Xènia Pérez Sitjà joined the External Relations team as Communications Officer as a maternity cover. Her contract was extended due to the increased outreach needs of our EU-funded projects. For the same reason, Ellie Taverner was added to the project management team. Additionally, Agnieszka Egan started a temporary secondment from EMBL-EBI as Operations Project Manager.

Over the course of the year, Jerry Lanfear's role evolved from CTO to Head of Programme and Strategy. The technical team also witnessed other significant changes with John Hancock's retirement as Community Coordinator and Rachel Drysdale leaving her position as Data Platform Coordinator. The recruitment process has started and will be completed in 2021.

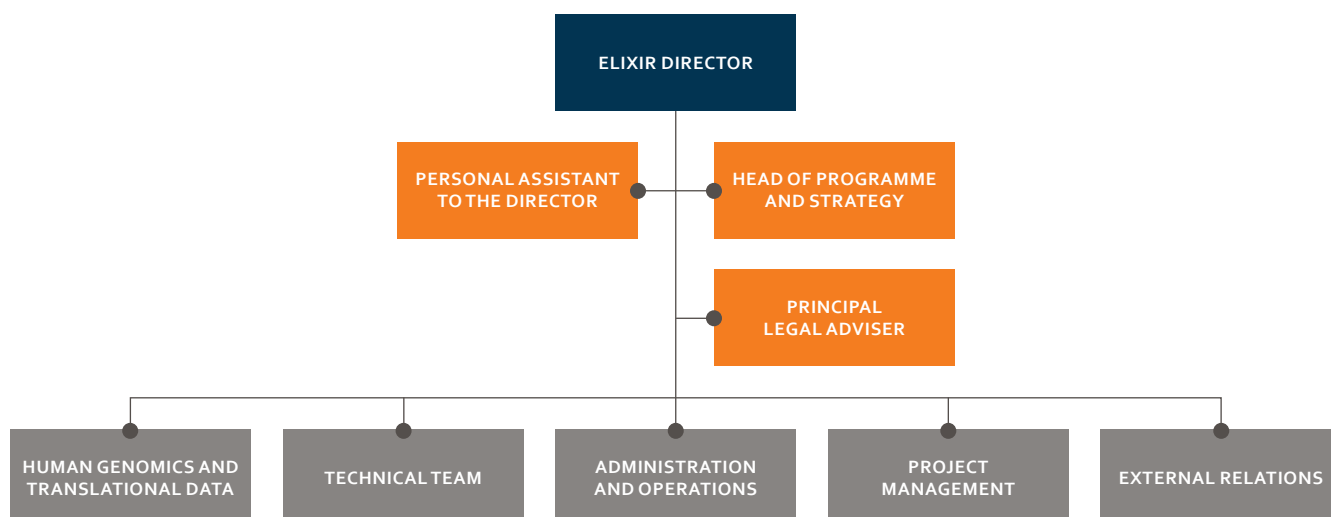
Nationalities



Gender ratio



ELIXIR Hub management structure





Striving for gender balance in bioinformatics

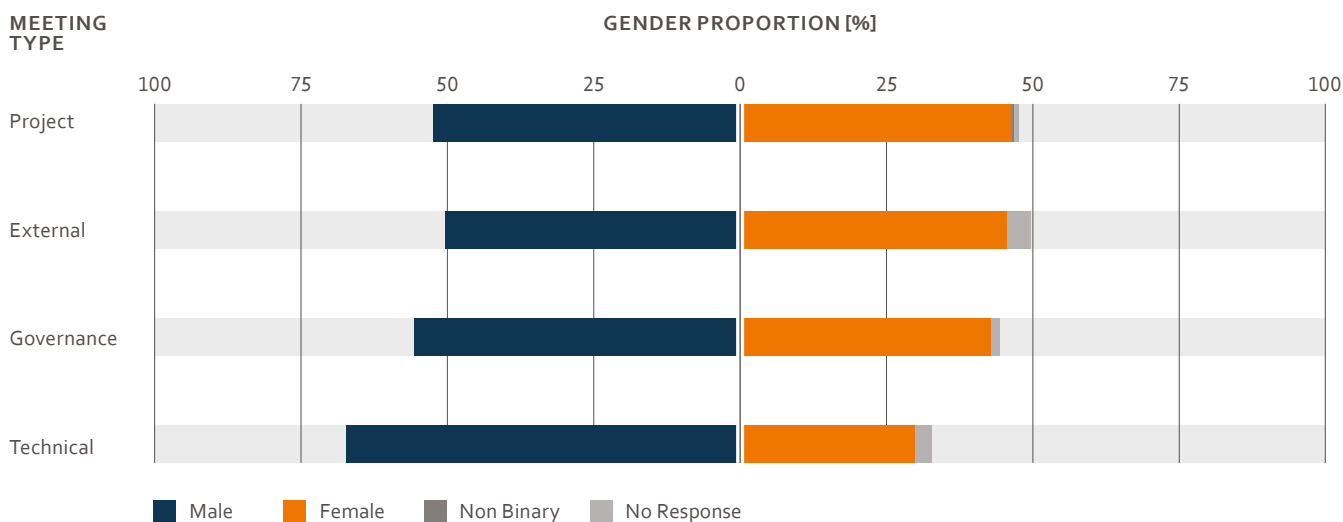
In 2020, ELIXIR started collecting more broadly gender information as part of the registration process in all events. The efforts aimed to support the work on equality, diversity and inclusion, following requirements of important funders, such as the European Commission and its gender equality strategy.

These gender data have presented a helpful snapshot of our infrastructure's profile to help us provide opportunities to under-represented groups during our events – and beyond.

Across all types of events in 2020, 59.7% of participants identified themselves as male, 37.5% as female, 0.3% as non-

binary, and 2.5% preferred 'not to say'. Based on 30 events, technical meetings (Platforms, Communities, All Hands, BioHackathon) had on average lower proportions of females (30%), whilst externally-funded project meetings, governance and externally-facing meetings had relatively higher female representation (43-44%).

Gender proportion at ELIXIR meetings across 2020



Governance

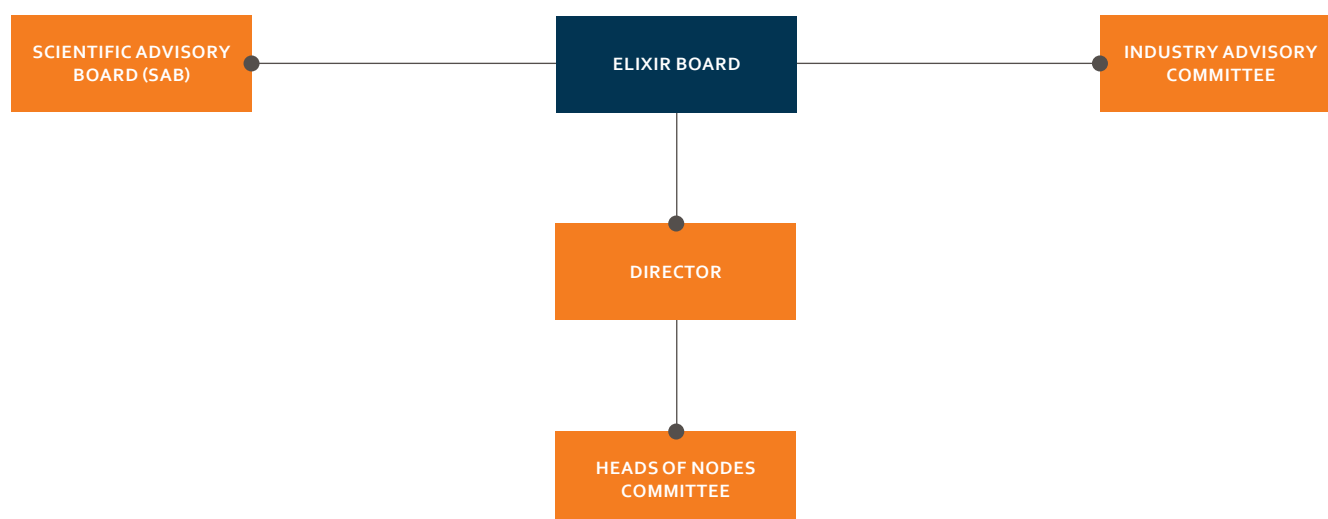
The highest decision-making body in ELIXIR is the ELIXIR Board, composed of representatives of ELIXIR members. The ELIXIR Scientific Advisory Board (SAB) advises the Board on ELIXIR scientific strategy and reviews the applications from ELIXIR Nodes. The SAB is an independent body, made up of leading 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 ELIXIR Scientific Programme. ELIXIR Director chairs the Heads of Nodes Committee.

ELIXIR governance hierarchy



ELIXIR Board members

MEMBER	ADMINISTRATIVE DELEGATE	SCIENTIFIC DELEGATE
Belgium	Laurence Lenoir (stepped down November 2020)	
	Michele Oleo	
	Didier Flagothier	
Czech Republic	Jan Burianek	Jaroslav Koča
Denmark	Line Bekker Poulsen	Anders Krogh
EMBL	Plamena Markova (replaced Silke Schumacher in August 2020)	Alvis Brazma
		Edith Heard
Estonia	Toivo Räm	Lili Milani
	Priit Tamm	
Finland	Riina Vuorento	Per Öster
	Sirpa Nuotio	
France	Eric Guittet	
Germany	Annette Kremser (replaced Johannes Mohr in March 2020)	Alexander Goesmann
		Rolf Backofen
Greece	Maria Gkizeli	Babis Savakis (replaced Artemis Hatzigeorgiou in September 2020)
		Christos Ouzonis
Hungary	Gábor Tóth	László Patthy
Ireland	Garry Purcell	Maria Nash
Israel	Barak Gatenyo (replaced Ilana Lowi in March 2020)	Iris Eisenberg
Italy	Grazia Pavoncello (replaced Salvatore La Rosa in January 2020)	Rita Casadio
Luxembourg	Jean-Claude Milmeister	Regina Becker
Netherlands	Ana de Castro (replaced Bea Pauw in February 2020)	Ruben Kok
Norway		Rein Aasland
		Stig Omholt
Portugal	Marta Abrantes (appointed November 2020)	Isabel Rocha
	Tiago Saborida	
Slovenia	Albin Kralj	Damjana Rozman
Spain	Ignacio Baanante	Ferran Sanz
	Cristina Bauluz	
Sweden	Mikael Borg (replaced Maria Nilsson in September 2020)	Björn Andersson
Switzerland	Doris Wohlfender-Bühler	Christian von Mering
UK	Mark Palmer	Christopher Rawlings
	Amanda Collis	

ELIXIR Heads of Nodes Committee

MEMBER	HEAD OF NODE	DEPUTY HEAD OF NODE
Belgium	Ferderik Coppens	Kim De Ruycck
Czech Republic	Jiří Vondrášek	Ludek Matyska
Denmark	Søren Brunak	
EMBL-EBI	Rolf Apweiler and Ewan Birney	Johanna McEntyre
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 (replaced Babis Savakis in September 2020)	Christoforos Nikolaou (replaced Martin Reczko in September 2020)
Hungary	Balázs Györfy	
Ireland	Denis Shields	Colm Ryan
Israel	Michal Linial	Dan Ben-Avraham
Italy	Graziano Pesole	Silvio Tosatto
Luxembourg	Reinhard Schneider	Wei Gu
Netherlands	Jaap Heringa	Morris Swertz
Norway	Inge Jonassen	Finn Drablos
Portugal	Mário Silva	Ana Portugal Melo
Slovenia	Brane Leskošek	
Spain	Alfonso Valencia	Salvador Capella-Gutierrez
Sweden	Bengt Persson	
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

Pascal Borry

University of Leuven, Belgium
(stepped down in March 2020)

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

		31/12/2020	2020	31/12/2019
		ACTUAL	BUDGET	ACTUAL
		€000	€000	€000
INCOME				
ELIXIR Member state contributions				
Ordinary contributions	(a)	7,233	7,200	7,089
Foreign exchange (loss)/gain on sterling contributions	(b)	(36)	-	(68)
Grant income	(c)	672	1,200	627
Miscellaneous income		100	-	-
Net Income		7,969	8,400	7,648
EXPENDITURE				
Technological Activities				
Salaries		597	600	571
Running costs		129	400	354
Commissioned services		1,402	8,400	3,014
Total expenditure Technological Activities		2,128	9,400	3,939
Directorate and Administrative expenditure				
Salaries		1,178	900	905
Running costs		318	600	322
Total expenditure Directorate and Administration		1,496	1,500	1,227
Support and Admin Infrastructure costs		898	900	674
Grant expenditure incurred		650	900	1,076
Total expenditure		5,172	12,700	6,916
SURPLUS/(DEFICIT)	(d)	2,797	(4,300)	732

(a) ELIXIR Member state contributions

	31/12/2020	31/12/2019
	€000	€000
Belgium	199	195
Cyprus	3	-
Czech Republic	70	69
Denmark	136	134
Estonia	9	9
Finland	98	96
France	1,104	1,082
Germany	1,529	1,499
Greece	89	87
Hungary	47	46
Ireland	81	80
Israel	126	124
Italy	816	800
Luxemburg	15	15
Netherlands	345	338
Norway	199	195
Portugal	84	82
Slovenia	17	17
Spain	536	526
Sweden	228	223
Switzerland	301	295
United Kingdom	1,201	1,177
Total	7,233	7,089

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

The ELIXIR Board approved that, from January 2016, the UK will pay its member state contributions in Sterling (ELIXIR/2015/28). The difference between the value of these contributions valued in Euros at the date of payment and the date of the approval of the 2020 budget was a loss of €36k (2019: loss of €68k).

(c) Grant income

	2020	2019
	€000	€000
Grant funding awarded	7,812	6,144
Grant income earned in the current year	672	627
Grant expenditure incurred in the current year	(650)	(1,076)
Unutilised grant income	2,752	1,734

(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 at 31 December 2020

	Prepaid contributions
	€000
Germany	1,560
Norway	203
Total	1,763



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 adaptation and advancement of ELIXIR regardless of the unprecedented changes during 2020.

Hinxton, UK, May 2021

Published under the CC BY 4.0 licence.



Data for life

ELIXIR HUB

✉ info@elixir-europe.org

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