

ELIXIR EOSC STRATEGY 2022



ELIXIR EOSC Focus Group

ELIXIR EOSC Strategy 2022

Authors: Jonathan Tedds, Salvador Capella-Gutierrez, Justin Clark-Casey, Frederik Coppens, Gavin Farrell, Celia van Gelder, Björn Grüning, Katharina Heil, Jessica Lindvall, Peter Maccallum, Ludek Matyska, Fotis Psomopoulos, Patrick Ruch, Susanna-Assunta Sansone

on behalf of the ELIXIR EOSC Focus Group.

Introduction	3
ELIXIR, Open Science and the Cloud	3
EOSC for the life sciences	4
ELIXIR's existing engagement with EOSC	5
ELIXIR's strategy for engagement with EOSC	6
Vision	6
Objectives	6
Principles	8
Implementation of the Strategy	9
Resource investment	9
Activities	9
Ownership & Review of this strategy	12
Ownership	12
Review	12
Update	12
Appendix A - EOSC and the ELIXIR Platforms	13
ELIXIR Compute Platform	13
ELIXIR Data Platform	14
ELIXIR Interoperability Platform	15
ELIXIR Tools Platform	16
ELIXIR Training Platform	17
Appendix B - EOSC and the ELIXIR Communities	18
Appendix C - ELIXIR and the EOSC Association Task Forces	18
Advisory Group on the Implementation of EOSC	19
Advisory Group on Metadata and Data Quality	19
Advisory Group on Research Careers and Curricula	19
Advisory Group on Sustaining EOSC	20

Advisory Group on Technical Challenges on EOSC	20
Appendix D - ELIXIR in EOSC Projects	21
EOSC-Pilot	21
EOSC-Enhance	21
EOSC-Life	21
EOSC-Future	21

Introduction

ELIXIR, Open Science and the Cloud

'[ELIXIR](#) unites Europe's leading life science organisations in managing and safeguarding the increasing volume of data being generated by publicly funded research. It coordinates, integrates and sustains bioinformatics resources across its member states and enables users in academia and industry to access services that are vital for their research.'

From its foundation, ELIXIR has been a leader in its vision of a network of Open Data resources and the services required to connect and make use of them. Key ELIXIR Members were among the authors of the original [FAIR principles](#) and the Life Sciences with their large and complex data resources collected through collaborative worldwide activity have long been at the forefront of the technical requirements and solutions for data storage, structuring and reusability.

ELIXIR has set out to provide the necessary infrastructure to support this activity at a European scale, organising across three dimensions - national, technological, and disciplinary.

The national aspect of Life Science data provision is provided by the network of **ELIXIR Nodes**. Each member state provides a Node as a national organisation, which provides a direct connection to funders and end users to respond to local research & training priorities and funding opportunities. In the vision of a connected network of service providers and consumers connected through a Research Cloud the Nodes provide a natural channel to connect European resources and leverage trans-national capabilities for large scale undertakings.

The technological aspects of the infrastructures are provided by the ELIXIR Platforms. ELIXIR technologies are grouped into the **ELIXIR Compute Platform** - the provision of physical computation and storage resources and the technologies which enable them to be connected and used for complex life science problems, **ELIXIR Data Platform** - the key underpinning life science data sources and the technological frameworks which enable them to be curated, accessed and maintained, **ELIXIR Interoperability Platform** - the technologies which allow complex activities to be composed from multiple Data, Compute, and Tools components, **ELIXIR Tools Platform** - the identification, collection, description and delivery of reusable high quality analysis software and workflows and the **ELIXIR Training Platform** - the Europe-wide coordination of bioinformatics training material development, reusability and discoverability. Together the ELIXIR Platforms are designed to provide a connected set of resources required to enable reuse of data, efficient use of resources and reproducibility of research in a trans-national network.

The disciplinary focus of the ELIXIR infrastructure is provided by the **ELIXIR Communities** which bring together researchers around specific experimental data types (for example the *ELIXIR Metabolomics Community*) or cross-platform technology user groups (for example the *ELIXIR Galaxy Community*). The ELIXIR Communities have proven to be a powerful way of engaging and enabling co-design between the Platform technologies and the requirements of the wider user base.

Together, the combination of **ELIXIR Nodes**, **ELIXIR Platforms** and **ELIXIR Communities** make up a federation of national, technological and expert services and people networks which enable

the ELIXIR user community to reuse data and resources at every scale from the individual researcher needs to large scale pan-European projects.

EOSC for the life sciences

The ambition of the [European Open Science Cloud](#) (EOSC) is to develop “Web of FAIR Data and services’ for science in Europe. EOSC will be a multi-disciplinary environment where researchers can publish, find and re-use data, tools and services, enabling them to better conduct their work.’

The [EOSC Association](#) was formed in 2020 and is the legal entity established to govern EOSC and has since grown to over 200 Members and Observers. The Association membership is jointly responsible for delivering the objectives agreed in the Memorandum of Understanding signed by the European Union and EOSC Association to form the official [EOSC Partnership](#). The EOSC ecosystem is being co-created in a series of funded projects and initiatives from Member States and Associated Countries. The EOSC Association plays an important role in helping to coordinate and steer these investments via its Task Forces and other governance structures.

The high-level architecture of EOSC consists of EOSC-Core and EOSC-Exchange. **EOSC-Core** is The set of internal services which allows EOSC to operate. This includes a technical platform which facilitates EOSC operations upon which the researcher-facing resources in the EOSC-Exchange can rely and integrate with as appropriate. It also includes non-technical coordination functions which operate and facilitate the technical platform (such as the service management system and the onboarding and the security coordination). **EOSC-Exchange** is the set of federation services and other resources registered in EOSC by research infrastructures, e-infrastructures, and science clusters to serve the needs of research communities and widening to the public and private sector. Generic services and resources which target heterogeneous scientific domains and research communities are identified as horizontal services. Resources which target users from a specific discipline, community, and/or regional domain are identified as thematic and/or regional resources. The capability to compose resources across horizontal and thematic and/or regional resources in compliance with the interoperability framework is defined as the Execution Framework. While it is expected that the majority of horizontal services are provided by the e-infrastructures (e.g. EGI, EUDAT, OpenAIRE, and GÉANT), generic services and resources offered by the science cluster communities will also be offered as a horizontal service. The architectural framework for EOSC-Core and EOSC-Exchange is directed by a set of [EOSC Association Task Forces](#) for different technology focus areas - such as *Technical Interoperability of Data and Services* and *Long-Term Data Preservation*.

The European Research Infrastructures have worked together via five [ESFRI Science Clusters](#) since 2019 and have strong links with research communities and projects, as well as managing significant data volumes and developing innovative data analytics tools, ensuring effective research data exploitation. The five comprise: EOSC Life for the life sciences; ENVRI-FAIR for environmental research; PaNOSC for multidisciplinary scientific analysis in the Photon and Neutron fields; ESCAPE for astronomy and particle physics; and SSHOC for social sciences and humanities.

ELIXIR's existing engagement with EOSC

ELIXIR has long identified EOSC as a significant strategic partner, with a [position paper](#) in 2019 and the 2019-2023 Scientific Programme highlighting the development of the relationship as a strategic objective.

ELIXIR and EOSC have worked together on many aspects of the underpinning technologies and end users application enabled by the shared vision - [EOSC Pilot](#), [EOSC Enhance](#), [EOSC Life](#) and [EOSC Future](#) have each enabled the sharing and co-development of technologies and the demonstration of utility and impact for the life sciences.

ELIXIR's strategy for engagement with EOSC

Vision

'ELIXIR is a partner for EOSC - in implementation and application'

The [ELIXIR 2019-2023 Scientific Programme](#) has as its Strategic Objective 4 'ELIXIR is the recognised and trusted life science foundation of the European Open Science Cloud'.

ELIXIR's organisational and technical design allows for a more ambitious goal, which has already been reached in some areas. ELIXIR has already developed many of the required features of an Open Science Cloud and can provide expertise, technology and organisational models not just in the application domain of EOSC Exchange but in the underlying technologies of the EOSC Core. ELIXIR will bring its expertise, its proven approach to sustainability and federated, interoperability-centred technical development, and its engaged user communities as a resource to help shape the development of EOSC technical, organisational and social structures - while critically reviewing each new capability based on its impact on the delivery of research objectives, commitment to the principles of Open Science and its fit to international best practices. This approach will maximise the alignment of the EOSC and ELIXIR roadmaps - fostering re-use, co-development and delivering benefit to life science researchers and the societies they serve. In a sense, ELIXIR is EOSC.

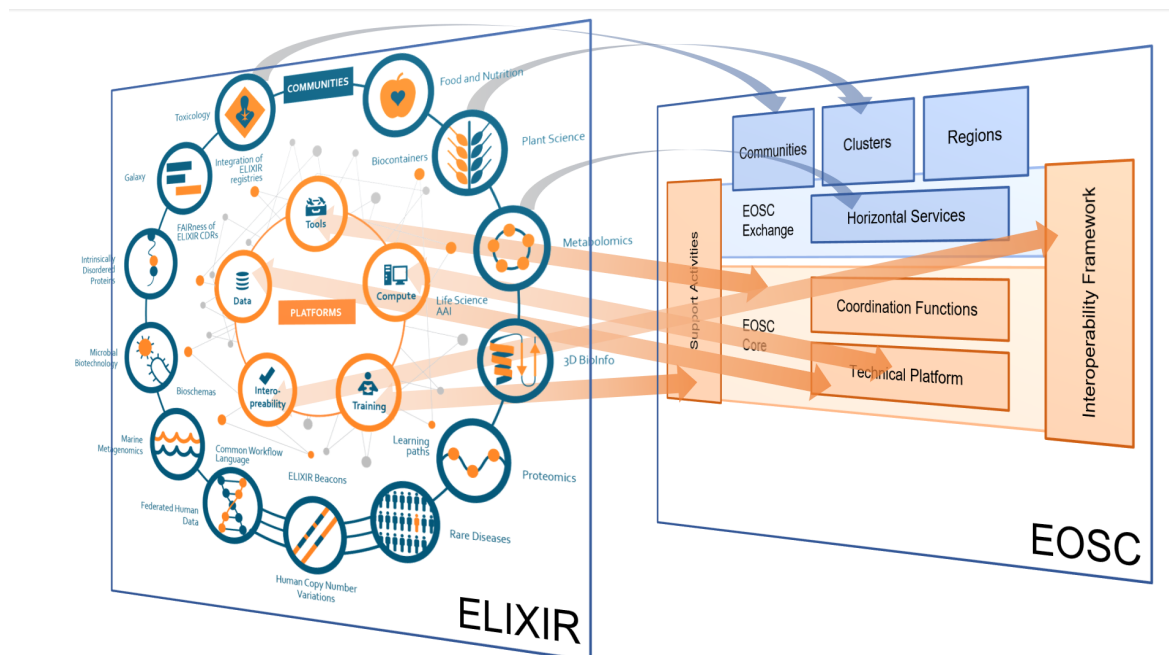


Figure 1 - ELIXIR and EOSC cross-map at all levels - not just in the Science Cluster activities and provision of domain specific data and software resources, but throughout the EOSC Core and Interoperability frameworks. These relationships will be bidirectional, with best of breed solutions being adopted by both sides. Through this process ELIXIR Communities will over time adopt EOSC Services and become driving user communities for both organisations.

Objectives

The [ELIXIR Scientific Programme 2019-2023](#) identified specific objectives for the end of 2023:

- **(O1) Define and develop EOSC for the life sciences**

ELIXIR, in partnership with BMS RIs and other stakeholders, will have defined and developed EOSC for the life sciences, allowing users to access fully operational services via ELIXIR portals and to be supported by ELIXIR Node experts.

- **(O2) Recommended Resources become core parts of the EOSC**

ELIXIR Recommended Resources, such as Core Data Resources, Deposition Databases and Recommended Interoperability Resources, become core parts of the EOSC for life sciences

- **(O3) Provide sustainable, cloud-based workflows**

ELIXIR will provide sustainable, cloud-based workflows that represent the gold standard for workflow-based biological data analysis, demonstrating the central role of ELIXIR in the European life science data landscape.

- **(O4) Build the capabilities of the user communities and infrastructure operators**

ELIXIR will build the capabilities of the user communities and infrastructure operators required to make the best use of the EOSC across all our Nodes.

At the time of writing (2022) these objectives remain valid as goals for the end of the 2019-2023 Programme; the vision and principles will be adopted and are expected to shape the objectives for 2024 and beyond. Many other ELIXIR activities, in particular participation in large scale EU projects, have delivered additional outcomes which will contribute to and influence EOSC.

Principles

In order to consolidate its position as the recognised and trusted life science foundation of the European Open Science Cloud, ELIXIR will adopt these principles:

- **(P1) ELIXIR and EOSC have a common direction of travel.**

ELIXIR will align our own infrastructure with the EOSC architecture, EOSC Core and EOSC Exchange, developing our own technology and standards where required. We will not pause the development of our own current required solutions for future EOSC services.

- **(P2) Bidirectional exchange of technology and services.**

ELIXIR will merge its services with selected mature EOSC services and mature ELIXIR services will become EOSC components. It will not adopt EOSC solutions which are not an improvement on existing practice, or where solutions are already in place in the ELIXIR communities; it will not continue development of ELIXIR services which are sustained and of better quality in EOSC.

- **(P3) Participation at every organisational level.**

ELIXIR will leverage our network-wide membership of EOSC organisational structures to monitor, assess and influence and in some cases lead EOSC strategic and technical direction for the benefit of our research communities.

- **(P4) Participation at every technical level.**

ELIXIR will participate in EOSC as technology developers as well as end users and domain service providers - both EOSC Core *and* EOSC Exchange. We will not invest in demonstration and requirements activities that can not demonstrate impact in the practice of life science research.

Together, these principles are designed to ensure that ELIXIR's underlying vision and purpose are not diluted but rather enhanced by commitment to participation in EOSC. ELIXIR and EOSC share a common end vision of an ecosystems of FAIR Data and accessible, interoperable services and ELIXIR can be a powerful enabling partner for EOSC in the Life Sciences, broker linkages between EOSC and Researchers, and co-develop missing capabilities without losing its technological edge in some disciplines and its focus on its end user communities.

By connecting at every level organisationally and technically, ELIXIR can communicate best practices in both directions and quickly identify shortcomings and mismatches between approaches for efficient resolution.

Implementation of the Strategy

Resource investment

ELIXIR will coordinate the implementation of this strategy from the Hub, with an EOSC Coordinator and dedicated effort from the team who are participating in EU funded EOSC projects. We will communicate the strategy to the ELIXIR Board and Heads of Nodes so they can choose how to apply their own national resources. ELIXIR Platforms and ELIXIR Communities will align their existing funded activities with EOSC in line with the **Vision** and **Principles**. ELIXIR participation in EU-funded EOSC projects will be based on the Principles outlined here with the aim of contributing to the overall strategic Objectives.

Activities

- **(I1) Coordinate information sharing**

ELIXIR will coordinate information sharing through the ELIXIR EOSC Focus Group, open to all ELIXIR Members and key partners with regular meetings where attendees at EOSC meetings, project participants and Task Force members can report significant updates and have the activities and membership recorded centrally for the benefit of all ELIXIR participants.

Key results:

1. Tracking of ELIXIR Membership of EOSC structures
2. Collection of reports and reference material from ELIXIR EOSC participation

- **(I2) Communication plan**

The ELIXIR Focus Group will develop a communication plan for bidirectional communication between ELIXIR and EOSC; including the production of shareable briefings and summary documents for ELIXIR Members, partners and EOSC and the organisation of information sharing events. This will support the promotion of ELIXIR services and expertise into EOSC, and the adoption of EOSC successes as new capabilities for the ELIXIR Platforms & Communities.

Key results:

1. Open material on ELIXIR contributions to EOSC to share with partners
2. Briefings and news for ELIXIR Members
3. Information sharing events

- **(I3) Development of EOSC Appendices for ELIXIR Roadmaps**

ELIXIR will not be 'implementing' EOSC (except through participation in EOSC Projects) and so will not have an EOSC implementation Roadmap; we will rather develop a section in each existing ELIXIR Roadmap - Platform, Communities, and Human Genomics and Translational Data - which maps the planned activities onto the equivalent EOSC structure, where they exist, and detailing the ways in which alignment and convergence will be managed. The set of EOSC appendices, plus the portfolio of ELIXIR EOSC Project involvement, will together represent the full Roadmap for ELIXIR's EOSC implementation and adoption activities.

Key results:

1. An appendix to each ELIXIR Roadmap on EOSC alignment
2. A summary of the key outputs and milestones of funded ELIXIR projects, past and future

- **(I4) Co-development of EOSC-Core and ELIXIR Services**

In the absence of EOSC (and in activities pre-dating EOSC in its current form), ELIXIR would still seek or develop all of the services required to provide a platform for life science data-centric research. In line with principle P1, this will continue. ELIXIR will gather requirements, evaluate technical solutions, and where required propose and implement key services for distributed Open digital research. Where those services have features not present in EOSC services, we will seek to have the additional capabilities introduced into the architecture and implementation of EOSC; where EOSC has capabilities which supersede ELIXIR equivalents, we will work with EOSC to replace the redundant ELIXIR components.

Key results:

1. ELIXIR offers its flagship services to EOSC projects and co-develops converged solutions
2. ELIXIR participates in EOSC projects on successful adoption of mature EOSC products

- **(I5) EOSC integration of existing ELIXIR resources**

ELIXIR already provides a rich ecosystem of data and services for life science researchers. Building on the work of the Science Clusters and EOSC-Life, we will support the registration of existing mature ELIXIR services in the EOSC Marketplace and make sure there is a distinct ELIXIR presence in the EOSC Service Portfolio as it develops.

Key results:

- ELIXIR mature resource collections (CDR, EDD, RIR) are registered in the EOSC Marketplace
- ELIXIR Nodes have shared guidance on how to represent their services
- ELIXIR services and key Service Collections can be found in the EOSC Registries & Marketplace

- **(I6) Engagement with EOSC projects**

Together these Objectives and Principles can be used to triage and select which EU funded EOSC projects we should prioritise.

ELIXIR as an organisation and through its Platforms will seek to join EOSC-Core projects as development partners. The success of LS Login will be the template here, but our services in FAIR data, training and throughout the tools and compute ecosystems will also contribute.

ELIXIR will contribute to the leadership of coordinated domain-specific activities in EOSC-Exchange and the Science Cluster projects. The model of EOSC-Life will be continued, and other Horizontal Services such as Workflow management will be developed and shared.

Nodes and Service providers will be encouraged to join EOSC projects and participate in EOSC activities proposed by Association Task Forces of which we are members. Opportunities to align national programmes and EOSC will be identified and developed.

Key results:

1. ELIXIR contributes key technologies for data-centric research through EOSC Core
2. ELIXIR plays a leading role in future EOSC Exchange/Science Cluster Projects
3. ELIXIR Nodes make optimised contributions to EOSC projects - benefiting life science and maximising impact
4. Relevant national programmes are involved in EOSC activities through ELIXIR Node.

Ownership & Review of this strategy

Ownership

This strategy has been developed through the design of the ELIXIR 2019-2023 Scientific Programme and discussions in meetings of the ELIXIR EOSC Focus Group and the ELIXIR EOSC Focus Group co-chairs.

Review

ELIXIR Focus Groups continue their operation subject to review by the ELIXIR Heads of Nodes and the ELIXIR Director; we anticipate that progress in implementation of this strategy will be reviewed by the Heads of Nodes during 2023 and again after the end of the 2019-2023 Scientific Programme in 2024.

Update

Our expectation is that it will remain relevant in its current form until the start of the 2024-2028 Programme, at which time the outcome of the review and design of the new programme will indicate an update is required. In particular the 2023 Objectives will be replaced with fresh 2028 Objectives and a revision will reflect those.

Appendix A - EOSC and the ELIXIR Platforms

ELIXIR Compute Platform

EOSC-Life is developing the [Life Science AAI](#), an evolution of the ELIXIR AAI that is widely deployed and used in the life sciences for identity and access management. We want the **EOSC AAI** to seamlessly integrate with the Life Science AAI to better enable cross-domain science and data access. **EOSC interoperability and operational requirements** on ELIXIR community **service providers** need to be lightweight and easily compatible with existing infrastructure and practices. These include aspects such as **AAI, monitoring, metrics, usage accounting, certification, helpdesk and security**.

In the life sciences we increasingly have situations where data cannot be moved and so needs to be processed in place. For example, sensitive personal data may be forbidden from leaving its host institution or jurisdiction, or the volume of data may be too large to move. We want to see EOSC integrate with the distributed computation APIs, workflow orchestrators and hybrid cloud infrastructure being used and developed in the life sciences, such as the cloud APIs of the [Global Alliance for Genomics and Health \(GA4GH\)](#) and workflow systems such as [Nextflow](#), [Snakemake](#) and [Galaxy](#).

We are in a world where access to scientific computing resources needs to be negotiated, allocated and funded project by project and resource by resource - a process that can take months. Sometimes these administrative hurdles stop smaller projects from proceeding at all. We need an EOSC payment mechanism that can allow for more permissionless scientific innovation. This could put in place the necessary persistent legal and contractual plumbing to send money to any EOSC service provider in exchange for services rendered, without each project negotiating their own usage agreement.

ELIXIR seeks to make **platform services sustainable**. Funding ongoing operational requirements such as security fixes, computation and storage infrastructure and third party service dependencies such as container image libraries is a continued challenge. We want to see EOSC continue to work on and promote sustainability solutions.

ELIXIR Data Platform

The goal of the ELIXIR Data Platform is to drive the use, re-use and value of life science data. The ELIXIR [Core Data Resources](#) (CDR) are the data sets of fundamental importance to the life science community; delivering these databases requires secure, scalable technologies, standards for data interoperability, and -as the linkage to human clinical data grows- an increasing requirement for federation and demonstrably secure, legal, and ethical access for research purposes. The ELIXIR Data Platform (EDP) is therefore actively re-evaluating the selection criteria and process, and monitoring the performances of these CDR to reflect the dynamic landscape of data usage requirements. Thus, developing scalable data curation frameworks to support the community of both data resource operators and users. The frameworks include biocuration-support services, as well as efficient access to the literature, including pre-prints and supplementary data (e.g. EuropePMC, SIBiLS). ELIXIR data providers represent some of the most significant international centres of expertise in delivery of datasets at scale; EOSC developments which support the management and operational delivery of these databases and cross-disciplinary approaches to data interoperability, search, security, audit and governance could impact on all of them.

The mature ELIXIR evaluation and recommendation processes, specifying the [ELIXIR Deposition Databases](#) (EDD), and the **ELIXIR Core Data Resources** select the resources we recommend to producers of experimental data from major biological data types to specific niche data, encouraging data to be open from early in its lifecycle. The life science community has invested heavily in this area and would seek to align but does not envision this function being replaced by more general data recommendation frameworks. We depend on tools to allow community users to become data providers, and so tools - including linking with the CDR - for the generation of high quality FAIR data, secure access environments for users to manage their own datasets, and dataset exploration systems are an area of active co-development with EOSC and other EU programmes.

The Data Platform also acts as a community bringing together experts - from the creation of ELIXIR-wide expert groups to specialist networks such as the **ELIXIR Data Management Expert Group**, and the [ELIXIR Biocuration Focus Group](#) - and with the ELIXIR Training Platform connects trainers for the data resources with their users. Tools which support the community curation of data are a focus within ELIXIR, and could be complemented by EOSC capabilities; the networks of people and skills could be connected to similar networks within EOSC.

ELIXIR Interoperability Platform

The [ELIXIR Interoperability Platform \(EIP\)](#) has been established to deal with the challenge of delivering FAIR data, to work with FAIR data, and to enable its actual reuse. Specifically, the EIP:

- delivers a framework of **registries**, the **services** and the **standards** that equip the users with the fit-for-purpose FAIR resources; and
- identifies and promotes interoperability **best practices** for data providers and data integrators.

To achieve this, the EIP works in partnerships with a number of stakeholders, within and outside the ELIXIR Nodes and Communities, ranging from individual data providers and interoperability service developers to international standardisation initiatives and communities. The EIP also adopts, adapts and drives emerging practices and technologies, avoiding ad-hoc and proprietary implementations, and aligns with relevant and related key global efforts. EIP activities broadly align with the [EOSC Interoperability Framework](#).

The EIP has delivered a sustainable portfolio of [FAIR Recommended Interoperability Resources \(RIRs\)](#), a collection of foundational components recognised and highly recommended for fit-for-purpose FAIRification tasks and actions. Selected via a formal process, by a panel of external reviewers, many RIRs are already embedded within ELIXIR Node Service Delivery Plans and international data management practices, and recommended by the Horizon Europe and the [IMI/IHI Data Management Guidelines](#).

Among the resources in the EIP FAIR Framework those below are already part of the EOSC ecosystem, adopted by EOSC reports, and embedded in EOSC projects;

- [FAIRsharing](#) (cross-disciplinary): registry on (meta)data standards (technical and contextual), data resources and policy.
- [RO-Crate](#) (cross-disciplinary): a web-native FAIR Digital Object implementation; packaging Research Objects for exchanging, archiving, publishing and citing workflows and data.
- [RDMkit](#) (life sciences): an online resource with guidelines and best practices about research data management tasks.
- [FAIR Cookbook](#) (life sciences): an online resource with hand-on recipes that help users to make and keep data FAIR.
- [Bioschemas](#) (life sciences): using schema.org to enhance the findability on the Web of life sciences resources such as datasets, software, and training materials.

Current EIP ongoing activities particularly relevant to EOSC include:

- Creation of a FAIR Services Architecture Framework and the definition of **interoperability stories** demonstrating impact, benefits and actions; and
- Delivering the **Knowledge Hub** we will use the RDMkit and the FAIR Cookbook as components of the knowledge dissemination platform.
- Developing resources for **Research Data Management**, tailored by the Interoperability stories, which use the complementary FAIR Cookbook and RDMkit.
- Enhancing **Discoverability of the resources**, using Bioschemas and schema.org, also via the [collaboration between OpenAIRE and FAIRsharing](#).

ELIXIR Tools Platform

The Tools Platform helps communities find, register and benchmark software tools in the Life Sciences domain. As a reference place for assisting researchers to access, analyze and integrate biological data, the ELIXIR Tools platform can benefit from further interactions with EOSC.

These tools help researchers access, analyse and integrate biological data, and so drive scientific discovery across the life sciences. We maintain information standards for these tools, and produce, adopt and promote best practices for their development including developing software management plans for life sciences. We also:

- connect tools and data services
- link tools to training materials
- make it easy to download, deploy (e.g. using container technologies), and benchmark tools for performance
- put tools in the context of the common workflows in bioinformatics, including environments such as [Galaxy](#).

We are integrating and cross-linking the core products of the Platform ([bio.tools](#), [BioContainers](#), [WorkflowHub](#) and [OpenEBench](#)) to other ELIXIR resources, such as the training resource [TeSS](#).

The Tools Platform has initiated the development of the “Tools Platform Ecosystem” which implements many of the software management and discovery mechanisms which will be required by EOSC. This metadata exchange platform will coordinate the different registries and services maintained by the ELIXIR Tools Platform, using standards such as [EDAM](#), [Bioschemas](#), and [biotoolsSchema](#). Building on content aggregated and curated over years, it will open up the content and make it accessible beyond API calls. The Tools Ecosystem infrastructure aims to serve infrastructures and open communities beyond ELIXIR. This will be achieved by 1) a royalty free tools ecosystem and 2) by being inclusive to external resources, according to the governance and guidelines defined within the ecosystem. All the collected metadata can be freely exchanged between all resources, yielding over time in rich metadata annotation in downstream and upstream resources.

From the EOSC perspective, the ELIXIR tools platform can contribute to exposing high-quality metadata in software/workflow registries. There are synergies between the platform and the EOSC software task force especially associated with the topics software sustainability, software management plans and software development best practices associated with life sciences.

- Facilitate the cross-walk across the EOSC marketplace and the different registries and repositories available at the platform, e.g. [bio.tools](#), **WorkflowHub**, **Biocontainers** and **OpenEBench**.
- Contribute towards the sustainability of research software in Life Sciences by co-producing best practices for software development. This effort could greatly benefit from the ELIXIR tools platform efforts in this matter.
- Advance in the development of software management plan efforts as a complement to existing efforts around data management plans.
- Connection with **EOSC-Life** tools collaboratory

ELIXIR Training Platform

The **ELIXIR Training Platform** was established to develop a training **community** that spans all ELIXIR member states. It aims to strengthen national training programmes, grow bioinformatics **training capacity** and competence across Europe, and **empower researchers** to use ELIXIR's services and tools. The Platform:

- establishes and implements **best practices** in bioinformatics training through a **training Toolkit**, a collection of generic and open resources, with the aim to be used beyond the Platform, across the ELIXIR Communities and extended to other ecosystems.
- supports training providers across Europe in **creating** and **delivering training** for developers, researchers and trainers (Train-the-Trainer), aligning to the FAIR principles as applied for training.
- builds a **sustainable training infrastructure**, with the ELIXIR Training portal [TeSS](#) as the flagship service. It offers an environment for trainers towards sharing materials and event information, as well as a gateway for trainees to find relevant training events and resources, and to perform specific, guided analysis tasks via customised training workflows.

Across all activities, the Training Platform has implemented a sustainable strategy for Quality and Impact for training and encourages its adoption outside of the Training Platform. Ultimately the Platform supports members in growing their expertise and capacities, therefore empowering the individual Life Science researchers.

In this context, the ELIXIR Training Platform is setting out the route to become **an EOSC Training Provider for the life science domain**, be bringing **our expertise related to training, skills, competences** as well as the flagship **training portal TeSS**, already adopted by other major infrastructures such as training portal for the photon & neutron community ([PANOSC](#)), [Training Catalogue of EGI](#) and the Digital Research Skills Australasia ([DReSA](#)). In EOSC-Future context **TeSS** will be one of the 4 first selected pilots for the EOSC Training Catalogue.

Additional examples of our ongoing activities and outcomes are:

- Participation in [EOSC Working Groups skills & training](#) ([slides](#))
- The **ELIXIR/NPOS work on data stewardship competencies** has been mentioned explicitly as a use case in the Skills chapter of the [SRIA](#) (page 128)
- ELIXIR Training is contributing as TF lead and task lead in the **3 Skills & Training related Task Forces** of the Advisory Group "[Research careers and curricula](#)" of the EOSC Association
- ELIXIR Training is a member of the coordinating team and a use-case for the [terms4fairskills](#) work on a [controlled vocabulary for FAIR skills](#). This work is also linked to CONVERGE WP2.
- ELIXIR Training will **lead the Training Work Package in EOSC4Cancer**, which will be using RDMkit as well as TeSS
- ELIXIR Training is actively participating in the global [Community of Practice of training coordinators](#), which is the premier place where all EOSC Training related initiatives meet and work together outside the scope of their own projects. In the CoP we are putting ELIXIR on the map/on their radar. One of the ongoing activities is on **making training materials FAIR**, which links with the **ELIXIR FAIR Training Focus Group**.
- ELIXIR is active within the [RITrainPlus](#) project, showing a close link to EOSC. The existing link and close collaboration in planning new training opportunities will help to ensure close alignment between EOSC, the ELIXIR TrP and RITrainPlus, also allowing the integration of ELIXIR within the RITrainPlus course.

Appendix B - EOSC and the ELIXIR Communities

ELIXIR Communities bring together experts across Europe to develop and test/explore standards, services and training within specific life science domains, formed around domain experts in ELIXIR nodes. They i) Capture user needs into formal requirements, ii) Drive service developments & uptake (standards, tools, workflows, amongst others in ELIXIR Platforms), iii) support collaborative standards development and maintain these, iv) foster collaboration between ELIXIR Nodes via transnational expert networks, , v) introduce developments from Communities to ELIXIR platforms.

Within the EOSC context, ELIXIR Communities and their members can be advocates for EOSC and specifically ELIXIR's EOSC strategy on local/national level. Communication is bi-directional, allowing ELIXIR to capture EOSC related information from various life science areas and gather or distribute it. Communities can become EOSC use-cases, demonstrators, and apply to development funding opportunities.

EOSC's core activities range around the development of technologies, but they rely on the involvement of end users for focus and purpose. ELIXIR and specifically Community members are in an ideal position to take up such a role, not only testing and feeding back information, but also proactively shaping the wider EOSC (as well as local/national) strategy. Communities and their members enable alignment and collaboration with other ESFRIs and large scale initiatives.

The ELIXIR Communities currently active are:

- [3D-BioInfo](#) - Understanding the 3D structure of macromolecules like proteins and DNA.
- [Food and Nutrition](#) - Understanding the effect of food choices on human health.
- [Galaxy](#) - A Galaxy community in Europe, together with Galaxy resources and training.
- [Intrinsically Disordered Proteins](#) - Develops standards, tools and resources to help identify and characterise IDPs.
- [Marine Metagenomics](#) - A sustainable metagenomics infrastructure to nurture research and innovation in the marine domain.
- [Metabolomics](#) - Resources, analysis and infrastructure to help metabolite identification.
- [Microbial Biotechnology](#) - Helps the development of tailor-made microbes and biological systems.
- [Plant Sciences](#) - Develops an infrastructure to facilitate genotype-phenotype analysis for crop and tree species.
- [Proteomics](#) - Develops and maintains sustainable proteomics tools and data resources.
- [Toxicology](#) - Standards, tools and resources to aid toxicology research projects.
- [Federated Human Data](#) - Develops long-term strategies for managing and accessing sensitive human data.
- [Human Copy Number Variation](#) - Aims to make it easier to detect, annotate and interpret human Copy Number Variations (hCNVs).
- [Rare Diseases](#) - Supports the development of new therapies for rare diseases.

With three new ELIXIR Communities at different stages in the process of foundation:

- **Single Cell Omics, Systems Biology, Biodiversity**

Appendix C - ELIXIR and the EOSC Association Task Forces

The EOSC Association has five Advisory Groups, and these have commissioned Task Forces for the realisation of EOSC, which will liaise with the Board of the EOSC Association and shape strategy. In line with Principles 3 & 4, ELIXIR Members are present throughout these Task Forces to facilitate our bidirectional communication and influence.

Advisory Group on the Implementation of EOSC

Task Force on PID Policy and Implementation will recommend standards and approaches which could align with technologies in use by data resources; ELIXIR's Data and Interoperability Platforms have significant expertise in this area.

Task Force on Researcher Engagement and Adoption will benefit from association with the ELIXIR Communities and should be a focus for our engagement with them.

Task Force on Rules of Participation Compliance Monitoring will recommend rules for joining EOSC activities and technical infrastructures, and is of particular interest for ELIXIR Service providers and our own compliance frameworks - CDR, EDD, RIR and the Maturity Models around the Federated EGA.

Advisory Group on Metadata and Data Quality

Task Force on FAIR Metrics and Data Quality will recommend classifications which could complement the approaches taken in the ELIXIR CDR and EDD adoption processes. The metrics and quality criteria have the potential to create technical or administrative barriers to data providers joining EOSC projects and catalogues, and ELIXIR can provide a useful benchmark for the costs and benefits of these recommendations.

Task Force on Semantic Interoperability will look at the technical framework for interoperability in which ELIXIR already plays a leading role, through all five Platforms but centred on the Interoperability Services.

Advisory Group on Research Careers and Curricula

Task Force on Data Stewardship Curricula and Career Paths will be relevant for the FAIR skills tracks of the Training Platform's activities and also for the development of the Data Platform's community of experts and users through the Biocuration Focus Group and the CONVERGE Data Expert Network.

Task Force on Research Careers, Recognition, and Credit could provide useful resources for the training and long term career recognition of ELIXIR Members and the bioinformatics community more generally (for example through data citation and work such as APICURON on data curator recognition).

Task Force on Upskilling Countries to Engage in EOSC will be relevant for the local development of ELIXIR capabilities in each Node and overlaps with ELIXIR's overall goal to connect European life science researchers for the benefit of all participating countries.

Advisory Group on Sustaining EOSC

Task Force on Defining Funding Models for EOSC should provide recommendations relevant to the ELIXIR Service providers with a focus on the sustainability of key services and high value data resources.

Task Force on Long-Term Data Preservation will provide useful insights into the long term sustainability of resources beyond the lifetimes of individual projects, with both technical and organisational solutions to the ultra-long retention of life science data.

Advisory Group on Technical Challenges on EOSC

Task Force on AAI Architecture should provide recommended services to allow seamless access by authorised users in complex multi-data-source project environments. Each EOSC technical component or interchange specification has the potential to provide a reusable part of an ELIXIR service or a barrier to entry; each user-focused component has the potential to streamline and improve the user's experience or to add an unwelcome layer of complexity - each needs to be carefully evaluated by data service providers and research community end users.

Task Force on Infrastructure for Quality Research Software will align with existing ELIXIR work through projects and the Platforms on best practices, open source principles, and the development of a software management plan to complement the now accepted concept of a data management plan.

Task Force on Technical Interoperability of Data and Services will provide a component level description of the parts of EOSC which should include profiles for data-centric interfaces and interchange mechanisms, encompassing or complementing the ELIXIR CDRS, EDDs and RIRs as well as the Compute and Tools Platform services.

Appendix D - ELIXIR in EOSC Projects

[EOSC-Pilot](#)

(2017-2019) The **EOSCpilot** project supported the first phase in the development of the European Open Science Cloud (EOSC). The project improved the ability to reuse data resources and provided an important step towards building a dependable open-data research environment where data from publicly funded research is always open and there are clear incentives and rewards for the sharing of data and resources.

[EOSC-Enhance](#)

(2019-2021) **EOSC-Enhance** aimed to improve and enhance the discoverability of the European Open Science Cloud (EOSC) services and resources. It developed and augmented the EOSC catalogue, and enriched the EOSC data offering by connecting the thematic clusters and clouds then in development.

[EOSC-Life](#)

(2019-2023) The **EOSC-Life** project aims to create an open collaborative digital space for life science in the European Open Science Cloud (EOSC). The 13 research infrastructures involved in the project will publish data from facilities and data resources in the EOSC. They will link these resources to open and reusable Tools and Workflows accessible to users via Europe's national and international life-science clouds. EOSC-Life also aims to connect users across Europe to a single login authentication system and develop the policies needed to preserve and deepen the trust given by research participants and patients volunteering their data and samples.

[EOSC-Future](#)

(2021-2023) **EOSC-Future** will make incremental improvements to its EOSC platform during 3 phases:

1. Researchers can access a rich set of resources, including some regional and thematic ones. They can also use a rich and diverse set of services, such as requesting resources through INFRAEOSC-07 projects.
2. Researchers can request access to centrally funded resources from multiple locations as the integration functions are improved, thus enabling sharing between portals.
3. Researchers can access a full spectrum of the available resources, from several funding streams and commercial sources. They can search data, recompose data into new datasets and discover new resources through smart, personalised recommendations. Users have access to services that support the full lifecycle of data processing, storage, and analysis including a growing number of end-to-end workflows.