Organisation: EOSC Life Science Research Node

Project Charter

EOSC Life Science Research Node

1. Project Summary [1]

Four world-leading infrastructures - ELIXIR, EMBL, Euro-BioImaging Eric and Instruct-Eric - are coming together to form an EOSC Life Science Research Node. These infrastructures have actively engaged in the development of the European Open Science Cloud since its inception, including through engagement in EOSC Task Forces and coordination and engagement in numerous EOSC-related EU projects to date.

Each partnering RI has expertise, multiple services and sites/partner institutes across Europe, meaning the collective contribution to open science of partners is significant. The aim is that, by formalising this collaboration, the Node will play an instrumental and significant role in the deployment of the EOSC federation.

Initial use cases will cover the scientific domains of the infrastructures represented, including structural biology, data presentation layers covering a range of biological data-types and the provision of Galaxy analytics workflows to life scientists. The EOSC Life Science Research Node will work closely with other EOSC Nodes of all types, and partners are open to the possibility of onboarding new RIs into the Node and providing advice to other future Nodes.

As a thematic Node comprising 4 RIs, partners are developing their own fit-for-purpose governance and operational structures. These will ensure that the Node functions effectively as a collective yet respects the individual make-up and governance structures of the partners.

2. Value Proposition

- **Main Goal:**

The goal of the EOSC Life Science Research Node is to connect four existing RIs (ELIXIR, EMBL, Euro-BioImaging and Instruct) into a thematic Life Science Research Node that supports the data and software related needs of a large and diverse user community in Europe and beyond.

- **Needs addressed:**

The needs of many different stakeholders will be addressed through an effective Life Science Research Node, including the needs of life scientists, service providers as well as users in other domains. The diverse services provided through the Life Science Research Node will enable life science researchers to access biological data from different modalities and sources, as well as software, tools and other services to fully exploit the data. In addition, solutions developed by the partner RIs, like managing very large databases, specialised formats and standards for image data, working with sensitive data, efficient interface to compute including for AI methods, platform for access and data provenance management, and many others are agnostic of scientific domain and can benefit all researchers across domain.

- **Key Benefits:**

Large volumes of data are being generated through state of the art technologies and facilities provided by life science infrastructures including EMBL, Euro-BioImaging and Instruct. At the same

1

time, many services and resources exist to support the data and software management needs of this community, alongside tools and expertise to support data analysis and visualisation, such as those offered by Instruct, Euro-BioImaging, EMBL and ELIXIR. Integrating these key infrastructures and the services they operate as effectively as possible is a critical first step to ensuring that a major life science component of the EOSC federation develops effectively, for the benefit of users, partners and funders. In addition, the EOSC Node will further increase visibility and uptake of our services across highly diverse user communities in Europe and beyond.

Across the RIs involved in the Life Science Research Node, hundreds of services exist, from foundational services such as Life Science Login and ARIA supporting access and identity management (including for sensitive data) to interoperable data, software and computational services (including those using AI methods) which can support researchers to process, analyse, visualise and annotate data. Ensuring data and digital assets that come from life science research are 'FAIR' and have the maximum opportunities for re-use by other scientists within and outside the life sciences is a primary focus of the EOSC Life Science Research Node. Enabling this interoperability within the context of the European Open Science Cloud through integration with other EOSC Nodes as well as connecting to current and future initiatives that will help to implement other EU Data Spaces is critical. The connection of these RIs via the Life Science Node is an important step to achieve this, given the partners will engage closely with each other, as well as other EOSC Nodes Federation including the EU Node.

- **Who Benefits:**

End Users

- More closely connected and enhanced services offered through the thematic Node will save time and effort, and lead to higher research productivity.
- Further service integration within the life science domain and beyond will enable greater interoperability and facilitate cross-domain collaborations.

The four participating RIs in the EOSC Node

- They benefit because the governance and operational frameworks developed as part of the thematic node will support deeper integration between the expertise and services they provide and increased uptake by new user communities. As a result this will also streamline and improve efficiencies across the RIs.

Funders

- They benefit from efficient implementation of their investments and better connections between the RIs involved, avoiding fragmentation and duplication.

Other EOSC Nodes

- Opportunities for closer collaboration and information exchange particularly with other EOSC Thematic Nodes.
- Integrated deployment of diverse Life science services on national infrastructure through EOSC National Nodes.

Second phase EOSC Nodes

- They benefit as they will be able to adopt good practice in forming EOSC Nodes or connect to the Life Science Research Node

3. Use Case(s)

The following use cases are presented as initial examples of potential applications. At this early stage, they should not be considered as firm commitments from either party, as they may evolve or be refined as the Node progresses and further stakeholder engagement occurs. Equally, it is envisaged that additional use cases will be developed and included in the coming period.

Use Case Description 1: Towards accelerated drug discovery and improved human health via *in silico*, *in vitro*, and *in vivo* approaches (Instruct-ERIC, EMBL, ELIXIR, EuroBioImaging)

This use case highlights how the European Open Science Cloud (EOSC) facilitates collaboration across multiple research infrastructures to build on existing findings, such as a novel protein interaction in 'Disease X', with the ultimate aim of supporting therapeutic development.

Beginning at the EOSC EU node, a researcher analyses pre-existing data via the EOSC resource hub, or raw data they collected themselves, (e.g. a proteomic dataset) using Jupyter notebooks and software resources via the EOSC EU node. Once a potential interaction between two proteins is identified, protein structures are retrieved from databases such as the PDB, and computational modelling of the interaction is conducted using EOSC Node resources provided through e.g. the EOSC EU node. This is further validated by using AlphaFold / ELIXIR Galaxy to ensure a robust *in silico* investigation.

To validate the interaction *in vitro*, the researcher uses Instruct-ERIC's ARIA service, where access to appropriate experimental techniques are arranged to confirm the predicted interactions experimentally. Structural data is then processed, refined (PDB-REDO) and deposited into PDB (EMBL) rendering it openly accessible. Data management is handled through federated storage solutions, such as that made available via ICSC, NFDI, or EOSC-PL nodes, with Instruct FandanGO used to index this data and associate it with the original proposal.

Simultaneously, the interaction is confirmed *in vivo* using techniques such as Förster Resonance Energy Transfer (FRET) Microscopy. Image data obtained from these experiments is harmonised and made cloud ready via conversion to the community driven OME-Zarr file format, enabling it to be easily visualised and analysed (EuroBioImaging). This analysis could be performed using existing pretrained models from the BioImage Model Zoo, implemented in ZeroCostDL4Mic or Galaxy.

Finally by utilising in vitro services listed by Instruct, and *in silico* approaches that make use of software available in the EOSC EU node resource hub, the research progresses towards therapeutic intervention development, with the information collected thus far contributing to the identification of small molecules perturbing this interaction.

- **Value proposition:**
    - Benefitting from EOSC
        - Benefit from federating capabilities and interoperability framework to build cross RI and cross node workflows
        - Visibility through EOSC
        - Sharing of good practices with European communities
        - Assistance on the implementation of FAIR
    - Benefit to EOSC
        - ARIA access management platform (Corbel, EOSC-Life), which has already been used outside of the life sciences (CESSDA data access – EOSC Future, non-Structural Biology projects and Research Infrastructures)
        - Provenance tracking of raw and processed data, linked to the original science case
        - Provide Structural Biology & Life Science feedback into Interoperability Framework
        - Structural biology tools and resources to assist structural biology and wider life science research including simulation, visualisation, data processing,

fitting, analysis, annotation, modelling and model refinement, data management and orchestration.
        - Educational resources/training materials
    - Audience
        - Researchers (data users and providers) to use services
        - Infrastructures to handle users and user data in a federated and FAIR way

- **In-scope:**
    - ARIA to provide catalogue for tools and services within the Instruct network
    - Expansion of integration between EOSC-AAI and ARIA-IDSS
    - Where appropriate services will be onboarded to the EOSC AAI via ARIA-IDSS for federation.
- **Out of scope:**
    - Storage is proof of concept at this stage, we will not be offering long term storage for all Life Science users

Use Case Description 2: Linking data-resources to foster therapeutic discovery (Instruct, EMBL, ELIXIR)

Protein Data Bank Europe (PDBe) hosted at EMBL-EBI is part of the wwPDB, which represents the major repository for Instruct-ERIC's Structural biology community and is a key resource of the Elixir network. These data were recently utilised to develop AlphaFold. This AI-based structure prediction system is revolutionising all life science domains, highlighting the importance of open, FAIR repositories for interdisciplinary research, developing AI methods, and reusing research outcomes.

However, the representation of protein complexes, key for structure-based drug design, is comparatively lower in public repositories. On the other hand, new technology development in fragment screening has enabled application of high throughput Structure-based and fragment-based drug design (SBDD, FBDD) approaches for publicly funded projects. These projects enable exploration of new proteins for developing novel therapeutic and biotechnology solutions, and can generate the amount of drug-like target interaction data to enrich structural databases and enable future development and validation of AI-driven SBDD. Here we propose to combine/integrate Elixir, Instruct, and EMBL data storage and analysis resources to develop a pilot cloud-based service for data deposition and analysis in the context of FBDD and SBDD. This project will link to resources developed through the Instruct-ERIC Use case. The access pipeline will be managed through the Instruct ARIA platform, with data produced during experimentation indexed via the use of the Instruct FandanGO tool and linked to the research record.

Data generated from FBDD and SBDD at the EMBL facilities, along with cloud-based data analysis, facilitates rapid identification of potential fragment hits. Automated tools for large-scale metadata harvesting within the CRIMS software will facilitate deposition of SBDD data into the PDBe repositories following standardised dictionaries enabling data reuse and discoverability by integrating additional biological data from the PDBe Knowledge Base and ChEMBL. There is also potential for sequence variability data available through ELIXIR to be integrated as well.

- **Value proposition:**
    - Proof of concept for automated data collection and metadata harvesting,
    - Potential to translate to other techniques relevant to other nodes.
    - Increase in resources for AI driven SBDD.
- **In-scope:**
    - Pilot connects limited resources to support a number of proof-of-concept projects.

- **Out of scope:**
    - The fragment screening pipeline represented in this use case should be considered a proof of concept, and will not be ready for widespread use within this project.

Use Case Description 3: Harmonising image data in EOSC (Euro-BioImaging ERIC, EMBL, ELIXIR)

The proposed imaging use case will demonstrate the power of open data and FAIR practices by enabling reuse of existing image data. It will be driven by discovering, handling and analysing bioimage data from diverse sources in the EOSC based on high quality standardisation and curation processes. While the use case itself is proposed for bioimage data, the universality of image data itself means the solutions can support other domains, like environmental sciences, astronomy, etc. The use case is also strengthened by access to potential cross node services like those provided by BBMRI Node and NFDI Node.

The use case will follow the steps below:

Access image data: Openly available image data will be accessed through dedicated image data repositories, like the BioImage Archive, IDR and EMPIAR. Here we will explore the potential for extending to additional imaging modalities like digital pathology imaging.

Harmonise image data formats: The image data will be harmonised and converted to community driven OME-Zarr file format, making the data cloud ready and easy to access and analyse.

Curate image data: Curation process will be employed to ensure sufficient and quality metadata for selected datasets. Existing community standards (following REMBI recommendations and ontologies like FBbi, EDAM BioImaging etc.) for image data will be employed in the process.

Analyse image data: The harmonised image data will be analysed and quantified to derive novel information. Existing pretrained models from the BioImage Model Zoo (bioimage.io) will be applied to the data. The analysis pipelines will be implemented in ZeroCostDL4Mic or Galaxy platform and made available in the WorkFlow Hub.

Data and knowledge visualisation: Pipelines for in-browser access and visualisation of image data and analysis results will be developed. BAND virtual desktop will be used for interactive image data analysis.

The use case is based on work being carried out at Euro-BioImaging ERIC, EMBL and ELIXIR, and is supported by current EU projects including AI4Life, OSCARS (Open Call), foundingGIDE, EVOLVE and RI-SCALE as well as building on valuable work and tools from concluded EU projects EOSC-Life and EOSC Future.

- **Value proposition:**
    - Access to and processing of heterogeneous image data from different sources
    - Further development of data formats, standards and ontologies that users can adopt
    - Embedding of tools and services for image data analysis and visualisation in the EOSC
    - Solutions for large data handling in the EOSC

- **In-scope:**
    - Workflow for pulling large image data from source
    - Curation and conversion of image data to OME-Zarr

    - Image analysis workflows openly available to users

- **Out of scope:**
    - Excludes access controlled data (i.e. sensitive data)

Use Case Description 4: Galaxy multi domain graphical workflows platform used by tens of thousands of researchers (ELIXIR)

ELIXIR offers approximately 500 national Node provided services through their ELIXIR Service Delivery Plans which can be found here: https://elixir-europe.org/services. The exact requirements for EOSC Federation Node service provisioning will impact ELIXIRs choice of service provisioning at an EOSC Node level. A strong candidate as a set of key mature services and associated cross domain use cases in the Life Science Research Node is provided by the ELIXIR Galaxy Community.

Galaxy is an open-source platform enabling researchers to analyse and share scientific data through interoperable APIs and various user-friendly web interfaces. Launched in 2005, the Galaxy project has evolved into a robust framework utilised across diverse research fields, including *omics, biodiversity, machine learning, cheminformatics, NLP, material science, and climate research. Key features of the Galaxy platform emphasise transparency, reproducibility, and reusability. Its multi-user environment supports the sharing of tools, workflows, notebooks, visualisations, and data, facilitating the reproduction of results for verification and enabling future studies to build on previous work.

Galaxy captures all provenance information of a dataset, such as tool versions, parameters, and execution environments, which can be reused or exported using standards like BCO or RO-Crate, as developed in the EOSC Life Project, to public archives. This high level of provenance tracking also ensures traceability and reduces the environmental impact of data analysis. Centralised user support enhances efficiency, allowing simultaneous resolution of tool failures and promoting the use of more efficient or accurate tools.

The INFRA-EOSC project EuroScienceGateway has been implementing a shared vision for an open, collaborative digital space for European scientists by advancing six national Galaxy servers and the European Galaxy server, currently serving 100,000 users. To accommodate future growth, the service has developed technologies enabling scientists to integrate their own compute or storage infrastructure with their preferred Galaxy instance, known as Bring Your Own Compute (BYOC) and Bring Your Own Storage (BYOS).

In the EOSC United project, commencing in 2025, Galaxy will demonstrate the capability to utilise the storage and compute resources of the EOSC EU Node, including connected services such as BlueCloud from the marine domain, in order to meet real user needs, serving the long tail of science with the Galaxy framework. This will include deploying a Pulsar endpoint into the EOSC EU Node.

We plan to submit entire scientific workflows, e.g. the ones stored in the EOSC-Life co-developed WorkflowHub, to the EOSC-EU node Pulsar endpoint. Every Galaxy server connected to this Pulsar endpoint could also provision Notebooks (e.g. Jupyter Notebooks, RStudio instances) using the

EOSC-EU resources. We will also create customised and user-centric training material using the GTN infrastructure

The compute resources made available via Pulsar and the storage resources of the EOSC-EU Node will then serve hundreds of thousands of researchers and enable use cases including the Biodiversity Genomics Europe (BGE)[2].

Out of scope: Processing of sensitive data until appropriate legal and technical frameworks to support this are in place.

Use case description 5: FAIR Image Analysis across Sciences (Euro-BioImaging, ELIXIR, EMBL, Instruct, other EOSC Nodes)

The fragmentation of quantitative image data analysis across disciplines leads to inefficiencies, limiting the reuse of methodologies and computational tools. This use case addresses these challenges by establishing cross-disciplinary FAIR image analysis workflows and integrating best practices from bioimaging, environmental sciences, and astrophysics.

By leveraging existing EOSC services, such as the European Galaxy server and WorkflowHub, we enable researchers to develop, share and reuse interoperable workflows across different EOSC nodes. For example, a workflow optimised for processing large-scale astrophysical data on one node using a Galaxy server can be adapted for bioimaging or environmental science applications on another. Achieving this cross-disciplinary, multi-node collaboration requires robust data interoperability and using ontologies to annotate tools and workflows accurately.

The EOSC's infrastructure, including communities from Science Clusters (LS RI, ENVRI, and ESCAPE) and research infrastructures like Euro-BioImaging and ELIXIR, provides a strong foundation for integrating AI methods, high-performance computing, and workflow management systems. Embedding FAIR principles at every stage ensures machine-actionability, transparency, and long-term usability of image analysis workflows. It also creates potential for extending the interoperable workflows to domains like structural biology within the EOSC LSR Node, as well as cross EOSC Node interaction with for example, the DataTerra, BlueCloud Nodes on a thematic level, and NFDI and EOSC PL Node on an infrastructure level.

To maximise impact and usability, we will develop comprehensive documentation and training materials for these workflows. Tutorials will be made available through the Galaxy Training Network (GTN), ensuring accessibility and reusability by researchers across disciplines. These tutorials will guide users in adopting and adapting workflows for their specific scientific needs.

This use case demonstrates how the EOSC Federation facilitates interdisciplinary collaboration, enhances computational scalability, and accelerates scientific discovery. By embedding these workflows within EOSC, we establish a sustainable framework for FAIR-driven image analysis, maximising impact across scientific domains.

- Value proposition
    - Enables cross-disciplinary reuse of image analysis workflows by integrating FAIR principles and EOSC services, creating a space for collaboration and reducing redundancy.

[2] https://biodiversitygenomics.eu/

- In-scope
    - Development and adaptation of FAIR workflows (proof of concept)
    - Use of ontologies for tool and workflow annotation
    - Interoperability across EOSC nodes
    - Creation of FAIR training materials and tutorials in the GTN
- Out of scope
    - Maintenance of EOSC services

Use case description 6: Harmonising biodiversity observations (EMBL, ELIXIR, Euro-BioImaging ERIC)

Omics-based observations have transformed how biodiversity is monitored across different environments, with major hubs such the Global Biodiversity Information Facility now taking species observations based on DNA sequences and/or images. However, such approaches present two challenges to research infrastructures: firstly the utility of multi-omics datasets relies entirely on interoperability of omics layers, i.e. on formalised data linking. Secondly, omics-derived data typically lead to computationally expensive analyses, and so rely on the availability of high performance computing or cloud infrastructures. To address this, common workflows have been produced by ELIXIR core data resources (e.g. MGnify[3]), communities (Galaxy[4] and Microbiome[5]) and more broadly. The Biodiversity Portal housed at EMBL-EBI connects multiple ELIXIR Core Data Resources, and is currently expanding to connect genomes to observations. Other omics data types, such as imaging, can be used to link phenotypes and genotypes with sequence data and allow modelling of ecological dynamics and adaptations to environmental conditions (including climate change).

This use case will demonstrate how EOSC federation can be used to execute different Workflows to meet the growing demands of data-driven biodiversity observations. We will develop and showcase data and metadata validation, reshaping, and packaging tooling to support federation between EOSC Nodes and other hubs like GBIF, supporting interoperability through RO-Crate and Darwin Core Archive standards. These developments will enable data presentation across a wide range of Nodes and resources. Current estimates indicate that there are millions of potential DNA datasets that could be systematically processed to produce species observations, demonstrating sharing of the compute burden, and removing redundancy that currently exists in the life-cycle. This use case builds upon a proof of concept established at an ELIXIR hackathon[6].

- Value proposition
    - Demonstration of how federated analysis can be performed in the EOSC cloud.
    - Potential to connect to NFDI (specifically NFDI4Microbiota, de.NBI) and BlueCloud
- In-scope
    - Validation of pipeline outputs via Research Objects (proof of concept)
    - Use workflows (Nextflow and Galaxy) for sequence annotation
    - Multiomics data integration
    - Interoperability across EOSC nodes
    - Presentation of data products suitable for BlueCloud Node
- Out of scope
    - Long-term compute resources

[3] https://workflowhub.eu/projects/9/workflows
[4] https://galaxyproject.org/news/2024-11-15-mgnify-v5/
[5] https://doi.org/10.1093/gigascience/giad078
[6] http://osf.io/preprints/biohackrxiv/3x274_v1

4. External Dependencies & Key Risks

- **External Dependencies:**

External Dependencies & Risks	Actions	Deadline
Sustainability	Support for the implementation of the Life Science Research Node will come, in part, from current and future EU project funding, as well as institutional support of the RIs involved. With respect to the nature and form of future EU funding, this is not yet currently known to partners, yet will dictate the relative levels of engagement.	Ongoing
Connections to EOSC EU Node	This will be supported by e.g. engagement with the EOSC United Coordination and Support Action commencing during 2025, involving ELIXIR and Instruct-ERIC directly.	Ongoing
Connections with other EOSC Nodes including other thematic Nodes and national nodes	Efforts will be made to find suitable opportunities to collaborate with, in particular, but not exclusively BBMRI, BlueCloud and PANOSC, where the life science research focus of the node has greatest potential for touching points with other thematic Nodes. In addition, institutes involved in national EOSC Nodes will also jointly be involved in the Life Science Research Node, so these will be mapped and interactions developed.	During 2025

Key Risks & Mitigation Measures:

Risk	Mitigation
Lack of adequate funding risks	Partners to work to secure appropriate institutional

delivery of activities (technical, operational, integration of new members)	support or new EU funding to cover actual costs of effective engagement in EOSC Node
Visible success of the EOSC Node obscures the value of the underlying Research Infrastructures involved	Partner's comms work to ensure that the successes of the Node are communicated, but make clear that this is possible due to the importance and availability of the underlying RIs and their services
Additional coordination challenges that come with being a thematic Node	Partners are establishing a light-weight MoU and effective operational processes to ensure the Node operates as efficiently as possible Node will raise issue with EOSC Governing Board/EC about setting of more realistic deadlines
Lack of clarity on obligations of offering a service through EOSC Node (e.g. SLAs, service review process/governance)	Engage closely in the discussions on these topics during 2025 Contribute to the development of the EOSC Node Handbook.
Limited sustainability and scalability	Even if the Node is successful in its early stages, there could be challenges in maintaining momentum, expanding to additional members and ensuring long-term financial and operational sustainability. Planning for scalability from the outset, by designing flexible governance and operational frameworks.
Lack of end-user engagement	Researchers and other end-users might not fully engage with the services offered by the EOSC Node if the services do not align with their needs or are not visible enough. Engaging with end-users early in the process to ensure that the services provided are user-centric. Continuous feedback mechanisms and regular communication efforts to raise awareness about the benefits and available services of the Node.

5. Contributions

Contributions will come from each Research infrastructure of the Life Science Research Node. All of the participating infrastructures are themselves distributed over multiple geographical locations and, in some instances, participating organisations. In the case of Instruct, ELIXIR and Euro-BioImaging each have a central hub and multiple RI nodes [distinct from EOSC Nodes]. Some resources (both human and computational or data assets) are owned and hosted by the hub but others are owned or hosted by one or more nodes or third parties, or some combination thereof. Given that there are multiple resource providers within each of the current four RIs making up the Life Science Research

Node that each have their own capabilities and resource constraints, a sliding scale of contribution needs to be considered when a resource provider onboards to the Life Science Research Node. For example, a simple wiki resource may only need to be added to a common node catalogue, but a more sophisticated tool may require integration with a common AAI.

We list below the initial sets of intended contributions from the four partner RIs. In the coming period more detailed plans will be developed including Deliverables assigned to the partners. This will be incorporated into a full set of Deliverables to include IT Governance documentation for appropriate services; data protection policies and procedures; operational deployment strategies and plans (including quality metrics and support); and stakeholder and community engagement plans.

Instruct-ERIC Hub

- **Role in Project:**
  - Data generator of structural biology data
  - Provider of ARIA access management platform for cataloging infrastructure services and managing access to services.
  - Provider of ARIA IDSS an identity provider enrolled in MyAccess ID (and hence EOSC EU Node). Supports AuthN via LS AAI/LS Login
  - Co-developer of FandanGO data management orchestration software
- **Main Contributions and Deliverables:**
  - Onboard ARIA's catalogue to EOSC Resource Hub
  - Explore integration with EOSC Federating capacities and EU Node services via expansion of integration between EOSC-AAI and ARIA IDSS
  - Support enrollment of Instruct Facility software to ARIA IDSS
  - Support enrollment of organisations with FandanGO
  - Pilot data storage mediation for user research projects through ARIA/FandanGO
- **Resources Provided:**
  - Personnel: 1.2 FTEs for managing contributions listed above.
  - Provision of ARIA access management platform, and joint provider of the automated metadata capture tool FandanGo.

Instruct-ERIC Centres

- **Role in Project:**
  - Structural biology data generator
  - Resource managers: Instruct-ERIC Centres manage many of the tools to be incorporated into the life science research node.
- **Main Contributions and Deliverables:**
  - Collaboration with the Instruct hub to onboard relevant services to EOSC via ARIA IDSS where appropriate
  - Inclusion of services within Life science research node service portfolio
- **Resources Provided:**
  - Personnel: 0.4 FTE split across centre's contributing resources for managing contributions listed above
  - Instruct facility software

ELIXIR Hub

- **Role in Project:**
  - Provider of Galaxy Community infrastructure
- **Main Contributions and Deliverables:**
  - coordinate ELIXIR Galaxy Community who provide services, hosting

    - coordinate community events, meetings
- **Resources Provided (if any):**
    - appropriate coordination and effort

EMBL-EBI Data Resources

- **Role in Project:**
    - Provision of relevant databases or data slices
- **Main Contributions and Deliverables:**
    - Data from PDBe, ChEMBL, ENA, MGnify, BIA
- **Resources Provided (if any):**
    - APIs supporting data provision and support

EMBL-Grenoble/Heidelberg Compute Infrastructure

- **Role in Project:**
    - Provision of compute infrastructure for processing datasets, where appropriate
- **Main Contributions and Deliverables:**
    - Compute resource sufficient for demonstration of Use Cases.
- **Resources Provided (if any):**
    - resourcing and support expertise

Euro-BioImaging Hub

- **Role in Project:**
    - Generation of biological and biomedical image data
    - Provider of services connected to imaging data: Consultation services for image analysis, image data curation, development of data standards and general tools for large file handling
    - Connection and outreach to a wide community of Users and Service providers
- **Main Contributions and Deliverables:**
    - Generate cross domain Image analysis workflows, including those using AI methods, as well as relevant standards and training materials.
    - Access to image analysis virtual desktop BAND
- **Resources Provided (if any):**
    - Personnel: 0.7 FTEs for strategic, technical, legal support and FAIR data expertise for integration of Euro-BioImaging image data services.

6. Timing and Milestones

A complete set of Milestones will be developed in the coming months, but we envisage the following tasks to be important for the Node to function effectively.

ID	Milestone Description	Target Delivery Date
1	**Establish membership of governance bodies including Steering Committee and Operations team**	**April 2025**
2	**Node communication and collaborative working environments and processes established**	**April 2025**

ID	Milestone Description	Target Delivery Date
3	**Establishment of links with other relevant nodes**	**April/May 2025**
4	**Onboarding of Instruct services to ARIA for inclusion in LSR Node service portfolio**	**May/June 2025**
5	**Monitor service integration process in the EOSC federation and update LSR Node service portfolio, if needed**	**Late May/early June**
6	**Expand on integration of EOSC AAI with ARIA-IDSS, and integrate Instruct centre services via ARIA IDSS.**	**Q3/4 2025**
7	**Assessment of relevant Use Cases' maturity level ahead of presentation at the EOSC Symposium**	**Late Sep 2025**
8	**Presentation of the final Use Cases at the Symposium**	**November 2025**

7. Contact & Submission

The EOSC Life Science Research Node is currently developing its governance and operational structure and will provide updated information to the Federation in due time on these contacts. We currently foresee two structures:

- An Executive Board, with each of the 4 RIs represented and a rotating Chair (to be appointed)
- Operational team

With respect to the operational team, the role of Coordinator will be tasked by and report to the Executive Board. Those we envisage will form the operational team are shown in the table. Additional roles will be finalised in the coming period.

Role	Name	Email
Coordinator	Peter Maccallum	peter.maccallum@elixir-europe.org
Operation Manager	Aastha Mathur	aastha.mathur@eurobioimaging.eu
Security Officer	Jonathan Tedds	jonathan.tedds@elixir-europe.org
Scientific Officer	Stephane Pesant	pesant@ebi.ac.uk
Technical Officer	Marcus Povey	marcus@instruct-eric.org
Legal Officer	TBC	TBC