

8.1 ELIXIR (CC)

- [Ambition](#)
- [Use cases](#)
- [Requirements for EOSC-hub providers](#)
 - [Technical Requirements](#)
 - [Capacity Requirements](#)
- [Validation plan](#)

| | |
|-------------------|-------------------------------|
| Short description | ELIXIR CC (Task 8.1) |
| Type of community | Competence Center |
| Community contact | Susheel Varma |
| Meetings | |
| Supporters | |

Ambition

The CC will enable ELIXIR to establish an ELIXIR Compute Platform (ECP) which allows ELIXIR cloud and data providers to share cloud compute and storage capacity to replicate and share reference datasets with each other and with their users. The platform aims to enable researchers to combine technical components of the ELIXIR Compute Platform services into a seamless ecosystem, thereby creating a science ready, standardised interface to the key resources and technological capabilities that are available for life sciences. The ECP aims to leverage the EOSC Service Catalogue to enable two related yet distinct activities for ELIXIR.

User stories

| No. | User stories |
|-----|--|
| US1 | <p>ELIXIR wants to establish a federation of cloud sites, each providing storage and compute capacity for researchers. The federated clouds should be connected to a data replication service (Reference Data Set Distribution Service with the ELIXIR terminology - RDSDS) that enables ELIXIR to stage 'ELIXIR Core Data Resources' to the cloud sites on-demand. As a result, the cloud sites become data hosting nodes which are equipped with CPUs/GPUs and are suited for large-scale data analysis and analytics.</p> <p>Centrally provided and curated datasets can ensure high-quality research in any of the partner states/regions. Researchers can go to their 'local' ELIXIR cloud provider, choose an already pre-staged ELIXIR dataset or request the staging of an ELIXIR dataset, choose an application of their choice (from a VM catalogue or container catalogue), maybe upload some additional data and then perform data analysis/analytics.</p> <p>Different conditions of access may apply at the different cloud sites, but it is expected that the cloud compute resources would be free at the point of use for national/local researchers, while pay-for-use or other special conditions apply for foreigners.</p> <p>The replication of community assets to national cloud providers maximises the utilisation of national funding and lowers the total cost of access for researchers.</p> <p>The services in the setup should recognise users via their ELIXIR identity, therefore ELIXIR AAI (Life science AAI) should be integrated with the RDSDS as well as with the national clouds.</p> |
| US2 | <p>The cloud federation can be also equipped with a 'container replication and orchestration service' that enables application providers to deploy containerised community/reference applications to any of the federated cloud sites, and users to instantiate and use the applications on those sites.</p> <p>Having a centrally managed or self-managed container orchestration service will allow users who do not have access to cloud resources of their own to deploy containerised workflows co-located with existing datasets in cloud locations.</p> <p>The services in the setup should recognise users via their ELIXIR identity, therefore ELIXIR AAI (Life science AAI) should be integrated with the RDSDS as well as with the national clouds.</p> |

Use cases



Instruction

A use case is a list of actions or event steps typically defining the interactions between a role (known in the Unified Modeling Language as an actor) and a system to achieve a goal.

Include in this section any diagrams that could facilitate the understanding of the use cases and their relationships.

| Step | Description of action | Dependency on 3rd party services (EOSC-hub or other) |
|------|---|--|
| UC1 | Joining the cloud federation with a cloud site (cloud provider perspective): <ol style="list-style-type: none"> AAI integration between ELIXIR and EOSC Connect ELIXIR cloud compute or data storage location with EOSC Policy compliance with policies (ELIXIR, EGI cloud, and EOSC-hub policies) | ELIXIR AAI EOSC AAI ELIXIR Cloud Services EOSC Compute / Storage Catalogue Policy compliance between ELIXIR and EOSC |
| UC2 | Making reference/core datasets available for replication to the federated cloud providers (data provider perspective): <ol style="list-style-type: none"> Data provider publishes reference dataset to RDSDS using ELIXIR AAI Based on a defined dataset replication policy RDSDS triggers one or more data transfer/synchronisation process via a central EOSC FTS service EOSC FTS manages the transfer process between source and location and notifies RDSDS on completion RDSDS notifies Data Provider that the data has been synchronised and/or any errors | Assumption: User is allowed to perform this action EOSC-hub centrally provided data distribution service (a new requirement to EOSC-hub!) EOSC-hub centrally provided application distribution and orchestration service (a new requirement to EOSC-hub) ELIXIR cloud federation policies, protocols and interfaces. (under definition) |
| UC3 | Requesting the replication of a reference/core dataset to my local cloud (researcher perspective): <ol style="list-style-type: none"> User searches and finds for a dataset with the RDSDS catalogue User initiates a transfer of a dataset to their local cloud resource using the EOSC central FTS EOSC FTS notifies the completion of the data transfer and/or errors | Same as above. |
| UC4 | Making virtualised, reference/core applications available for replication and orchestration on the federated cloud providers (data provider perspective): <ol style="list-style-type: none"> User searches the EOSC Service Catalogue to identify a container orchestration service The user uses their ELIXIR credentials to instantiate a container orchestration service User deploys a containerised application or workload to the container orchestration service | Assumption: User is allowed to perform this action ELIXIR AAI EOSC AAI Kubernetes as a Service |

Requirements for EOSC-hub providers

Technical Requirements

| Requirement number | Requirement title | Link to Requirement JIRA ticket | Source Use Case |
|--------------------|--|--|-----------------|
| RQ1 | EOSC-hub to provide an FTS data transfer service | EOSCWP10-21 - ... <input type="text"/> | UC1, 2, 3 |
| RQ2 | EOSC-hub to provide Kubernetes as a service | EOSCWP10-22 - ... <input type="text"/> | UC1, 4, 5 |

Capacity Requirements

The cloud capacity will initially come from the ELIXIR CC members (EBI, CESNET, CSC). Others will join in a second stage.

Capacity requirements for the centrally provided EOSC-hub FTS:

- 100 concurrent users
- 100 TB throughput / month
- Simultaneous staging of data to 4 sites

Capacity requirements for the centrally provided Kubernetes as a service:

- 100 concurrent users
- 500 container throughput / month
- Simultaneous staging of containers to 4 sites

Validation plan

| # | Task | Description | Expected Outcome |
|-----|---|--|---|
| UC1 | Integration of Cloud Provider with EOSC | ELIXIR service provider connects their service with EOSC | End-user is able to search and discover ELIXIR Compute and Data storage locations within EOSC |
| UC2 | Data Provider: Data Replication /Synchronisation | Data provider publishes data to RDSDS | Datasets are synchronised to multiple data storage location with ELIXIR and/or EOSC |
| UC3 | User: Data Transfer | User searches and initiates transfer for data to their local cloud | Dataset is made available to the user-spcified cloud location |
| UC4 | User: Container Orchestration Service | User initiates a containerised workload deployment | Users containerised workload was successfully deployed. |