

**EGI-Engage**

**ELIXIR Competence Centre**

Life science requirements analysis and driver
use case(s) with implementation roadmap

M6.3

|  |  |
| --- | --- |
| **Date** | Xx xx xxxx |
| **Activity** | SA2 |
| **Lead Partner** | EGI.eu |
| **Document Status** | DRAFT |
| **Document Link** | <https://documents.egi.eu/document/2675>  |

Abstract

The ELIXIR Competence Centre of the EGI-Engage project facilitates collaboration between EGI and ELIXIR service developers and providers by collecting, analysing and comparing life science community needs with EGI technical offerings. The process is driven by scientific use cases that are selected for the 18 month long work. This document is the first milestone of the Competence Centre effort: a selection of those science cases from the ELIXIR community that could benefit from EGI services; an initial analysis of the requirements derived from these science cases; and a roadmap to implement these science cases with the use of EGI services, within the ELIXIR Compute Platform.

**COPYRIGHT NOTICE**



This work by Parties of the EGI-Engage Consortium is licensed under a Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/). The EGI-Engage project is co-funded by the European Union Horizon 2020 programme under grant number 654142.

**DELIVERY SLIP**

|  |  |  |  |
| --- | --- | --- | --- |
|  | ***Name*** | ***Partner/Activity*** | ***Date*** |
| **From:** |  |  |  |
| **Moderated by:** |  |  |  |
| **Reviewed by** |  |  |  |
| **Approved by:** |  |  |  |

**DOCUMENT LOG**

|  |  |  |  |
| --- | --- | --- | --- |
| ***Issue*** | ***Date*** | ***Comment*** | ***Author/Partner*** |
| **v.1** | 12/Jan/2016 | ToC with initial text | G. Sipos / EGI.eu-SZTAKI |
| **v.2** |  |  |  |
| **...** |  |  |  |
| **v.n** |  |  |  |

**TERMINOLOGY**

A complete project glossary is provided at the following page: <http://www.egi.eu/about/glossary/>

**Contents**

[1 Introduction 5](#_Toc440296117)

[2 Scientific use cases 6](#_Toc440296118)

[2.1 Marine metagenomics use case - Kimmo 6](#_Toc440296119)

[2.1.1 Introduction 6](#_Toc440296120)

[2.1.2 Scientific use case description 6](#_Toc440296121)

[2.1.3 E-infrastructure requirements 7](#_Toc440296122)

[2.1.4 Impact 7](#_Toc440296123)

[2.2 Phenomenal project use case - Steven, Enol 8](#_Toc440296124)

[2.2.1 Introduction 8](#_Toc440296125)

[2.2.2 Scientific use case description 8](#_Toc440296126)

[2.2.3 E-infrastructure requirements 8](#_Toc440296127)

[2.2.4 Impact 9](#_Toc440296128)

[2.3 cBioPortal replication use case - Miroslav, Michal 9](#_Toc440296129)

[2.3.1 Scientific use case description 9](#_Toc440296130)

[2.3.2 Scientific use case description 10](#_Toc440296131)

[2.3.3 E-infrastructure requirements 10](#_Toc440296132)

[2.3.4 Impact 10](#_Toc440296133)

[3 Implementation roadmap 12](#_Toc440296134)

[3.1 Introduction 12](#_Toc440296135)

[3.2 The ELIXIR Compute Platform – Steven 12](#_Toc440296136)

[3.3 EGI services in the ELIXIR Compute Platform – Gergely 12](#_Toc440296137)

[3.4 Use case implementation roadmaps 12](#_Toc440296138)

[3.4.1 Marine metagenomics use case - Kimmo 12](#_Toc440296139)

[3.4.2 Phenomenal project use case - Steven 12](#_Toc440296140)

[3.4.3 cBioPortal replication use case - Miroslav, Michal 12](#_Toc440296141)

**Executive summary**

Aaaa

# Introduction

ELIXIR is a pan-European research infrastructure in agreement between 15 European governments to build a sustainable European infrastructure for biological information, supporting life science research and its translation to medicine, agriculture, bioindustries and society.

Life science is a fast moving field. For the EGI services to become relevant and help keep European Life Sciences globally competitive, it is important to develop mechanisms that allow the research infrastructure to flexibly meet new challenges and respond to new scientific and technical developments.

The ELIXIR Competence Centre of the EGI-Engage project evaluates, adopts and promotes technologies and resources from the EGI Federated Cloud towards the wider ELIXIR research community. This is achieved with an iterative approach:

1. Bringing together designated life science experts from ELIXIR and technical experts from EGI’s Federated Cloud collaboration.
2. Identify life science use cases which could benefit from the EGI Federated Cloud services and could make big impact on ELIXIR and EGI communities. Analyse their e-infrastructure requirements, especially with respect to the use of services and technologies adopted within the EGI Federated Cloud.
3. Implement selected life science use cases based on EGI. Collaborate on the implementation with EGI’s and ELIXIR’s partner e-infrastructures, primarily EUDAT.
4. Evaluate the implementations and disseminate the experiences gained with the use cases and with the EGI services towards ELIXIR, EGI and other relevant communities.

This document is a milestone after stage 2 of this process. The document was written by life science and e-infrastructure experts from ELIXIR and EGI, who were brought together within the competence centre. The document captures scientific use cases, derived requirements and envisaged implementation roadmap based on EGI services. Contributors of the report were:

|  |  |  |
| --- | --- | --- |
| Name | Role in ELIXIR/EGI | Contribution to the report |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

# Scientific use cases

This section provides information about the use cases that have been identified by the Competence Centre. These use cases represent scientific workflows that can be ported to the EGI Federated Cloud. The search for use cases was restricted to workflows that require only ‘non-sensitive’ data, because this simplifies complexity, and also make the ELIXIR Competence Centre effort complementary to the BBMRI Competence Centre activities (task SA6.4 of EGI-Engage), where the focus is on handling sensitive data with EGI services. Each of the use cases are described from three perspectives:

1. Scientific
2. E-infrastructure
3. Impact

These aspects together provide a comprehensive view on the use cases and help the Competence Centre focus its limited effort on those cases that would offer the best value vs. implementation and operational cost.

## Marine metagenomics use case - Kimmo

### Introduction

Aaaaa

### Scientific use case description

|  |  |  |
| --- | --- | --- |
| User Story | *Describe here your use case in term of user story:** *Step1: …*
* *Step2: …*
* *…*
 |  |
| (Potential) User base | *Describe the current and potential user base in term of size (number of users), type (from which scientific domain?), characteristics (academic, private company, students, etc…).**Describe how the potential user base could be caught.* |  |

### E-infrastructure requirements

|  |  |  |
| --- | --- | --- |
| HW Resources | *List the amount of HTC and/or cloud resource you need.**List any other HW requirements.* |  |
| SW Resources | *Describe the external SW products you need and the related licenses.**List the technology providers for each of the above SW product.* |  |
| Cost of delivery | *Describe what needs to be developed/integrated and how much this would cost in terms of effort (estimated PMs)* |  |
| Operational aspect | *Describe who would operate the services related to this use case.**Assess the effort needed to maintain these services up & running beyond the lifetime of the CC.**List the NGIs interested on supporting these services.* |  |

### Impact

|  |  |  |
| --- | --- | --- |
| Business plan 1 | *Define one or more business plan taking into the account the (potential) user base, the cost to delivery and maintain the service.**Identify the full market potential (all user types / categories and size) and most likely uptake possible.**Identify what revenue types will the provider obtain in return for the investment described above and possible estimates e.g. direct payment(s); funding; in-kind contribution**What are the organisational, technical, financial, market and/or legal risks associated to the service provider e.g. inability to scale to demand* |  |
| Business plan 2 | *…* |  |

## Phenomenal project use case - Steven, Enol

### Introduction

Aaaaaa

### Scientific use case description

|  |  |  |
| --- | --- | --- |
| User Story | *Describe here your use case in term of user story:** *Step1: …*
* *Step2: …*
* *…*
 |  |
| (Potential) User base | *Describe the current and potential user base in term of size (number of users), type (from which scientific domain?), characteristics (academic, private company, students, etc…).**Describe how the potential user base could be caught.* |  |

### E-infrastructure requirements

|  |  |  |
| --- | --- | --- |
| HW Resources | *List the amount of HTC and/or cloud resource you need.**List any other HW requirements.* |  |
| SW Resources | *Describe the external SW products you need and the related licenses.**List the technology providers for each of the above SW product.* |  |
| Cost of delivery | *Describe what needs to be developed/integrated and how much this would cost in terms of effort (estimated PMs)* |  |
| Operational aspect | *Describe who would operate the services related to this use case.**Assess the effort needed to maintain these services up & running beyond the lifetime of the CC.**List the NGIs interested on supporting these services.* |  |

### Impact

|  |  |  |
| --- | --- | --- |
| Business plan 1 | *Define one or more business plan taking into the account the (potential) user base, the cost to delivery and maintain the service.**Identify the full market potential (all user types / categories and size) and most likely uptake possible.**Identify what revenue types will the provider obtain in return for the investment described above and possible estimates e.g. direct payment(s); funding; in-kind contribution**What are the organisational, technical, financial, market and/or legal risks associated to the service provider e.g. inability to scale to demand* |  |
| Business plan 2 | *…* |  |

## cBioPortal replication use case - Miroslav, Michal

### Scientific use case description

Aaaaaa

### Scientific use case description

|  |  |  |
| --- | --- | --- |
| User Story | *Describe here your use case in term of user story:** *Step1: …*
* *Step2: …*
* *…*
 |  |
| (Potential) User base | *Describe the current and potential user base in term of size (number of users), type (from which scientific domain?), characteristics (academic, private company, students, etc…).**Describe how the potential user base could be caught.* |  |

### E-infrastructure requirements

|  |  |  |
| --- | --- | --- |
| HW Resources | *List the amount of HTC and/or cloud resource you need.**List any other HW requirements.* |  |
| SW Resources | *Describe the external SW products you need and the related licenses.**List the technology providers for each of the above SW product.* |  |
| Cost of delivery | *Describe what needs to be developed/integrated and how much this would cost in terms of effort (estimated PMs)* |  |
| Operational aspect | *Describe who would operate the services related to this use case.**Assess the effort needed to maintain these services up & running beyond the lifetime of the CC.**List the NGIs interested on supporting these services.* |  |

### Impact

|  |  |  |
| --- | --- | --- |
| Business plan 1 | *Define one or more business plan taking into the account the (potential) user base, the cost to delivery and maintain the service.**Identify the full market potential (all user types / categories and size) and most likely uptake possible.**Identify what revenue types will the provider obtain in return for the investment described above and possible estimates e.g. direct payment(s); funding; in-kind contribution**What are the organisational, technical, financial, market and/or legal risks associated to the service provider e.g. inability to scale to demand* |  |
| Business plan 2 | *…* |  |

# Implementation roadmap

## Introduction

During 2015 the ELIXIR community – in collaboration with various e-infrastructures and other service providers – initiated the development of a reference architecture for ELIXIR, called the ‘ELIXIR Compute Platform’ (ECP). The prime role of the ECP is to support the use cases of the ELIXIR-EXCELERATE H2020 project, however, the platform is expected to serve other ELIXIR-related use cases as well as use cases from other biomedical sciences Research Infrastructures. Because of the fundamental role of the ECP within ELIXIR, EGI-Engage Competence Centre use cases should also consider the ELIXIR ECP as a foundational infrastructure, and support EGI’s effort in integrating EGI services into this platform. The next subsections describe the architecture of the ECP, the ongoing activities in integrating EGI services into the ECP, and the implementation roadmap for the previously described use cases on top of the ECP.

## The ELIXIR Compute Platform – Steven

The capabilities and functional building blocks of the ECP have been discussed and identified in 2015 through a series of workshops and online consultation meetings. ….

## EGI services in the ELIXIR Compute Platform – Gergely

This section provides details about the development activities that are ongoing within various EGI product teams to customise EGI services for inclusion in the ECP. The content directly reflects on the EGI-related recommendations of the ‘The ELIXIR Compute Platform: A Reference Technical Services Architecture for supporting Life Science Research’ document’. …<TO COMPLETE BASED ON CONTENT FROM ONLINE GOOGLE DOC>

## Requirements analysis, implementation roadmaps

### Marine metagenomics use case - Kimmo

Aaaaaa

### Phenomenal project use case - Steven

Aaaaaa

### cBioPortal replication use case - Miroslav, Michal

Aaaaaa