D2.4 Technical, Policy and Service Management Integration Report

Lead partner:	EGI Foundation
Version:	1.0
Status:	Final
Dissemination Level:	PUBLIC
Keywords:	Onboarding, Portfolio, Catalogues, SMS, ITSM, FAIR
Document Link:	https://documents.egi.eu/document/3815

Deliverable Abstract
EGI-ACE delivers 'EOSC Compute Platform' services and 'Data Space services' in EOSC. This deliverable reports on the integration and alignment activities between EGI-ACE and EOSC. The document describes the approaches that were taken by the project for the service portfolio integration; technical interoperability; service management system; FAIRness assessment of the data spaces. Based on the findings the project will continue to strengthen the links with EOSC in 2022. An updated version of this deliverable will be produced as D2.8 in December 2022.

EGI-ACE receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 101017567.

go.egi.eu/egi-ace

COPYRIGHT NOTICE

This work by parties of the EGI-ACE consortium is licensed under a Creative Commons Attribution 4.0 International License. (http://creativecommons.org/licenses/by/4.0/).

EGI-ACE receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 101017567.

DELIVERY SLIP

Date	Name	Partner/Activity
From:	Gergely Sipos	EGI Foundation / WP2
Moderated by:	Malgorzata Krakowian	EGI Foundation / WP1
Reviewed by:	Małgorzata Krakowian Andrea Manzi	EGI Foundation / WP1 EGI Foundation / WP6
Approved by:	SDS	

DOCUMENT LOG

Issue	Date	Comment	Authors
v.0.1	11/12/2021	Full draft for external review	Gergely Sipos, Eleonora Testa, Giuseppe La Rocca, Diego Scardaci, Alessandro Paolini, Matthew Viljoen, Gianni Dalla Torre, Andrea Manzi, Enol Fernandez, Yin Chen (all EGI Foundation)
v.1	16/12/2021	Updated draft based on review feedback.	Gergely Sipos (EGI Foundation)

TERMINOLOGY

https://confluence.egi.eu/display/EGIG

Table of contents

Executive summary

EGI-ACE is a 30-month project (Jan 2021 - June 2023) with a mission to empower researchers from all disciplines to collaborate in data- and compute-intensive research through free-at-point-of-use services.

EGI-ACE delivers the 'EOSC Compute Platform', an integrated compute environment that federates compute and storage facilities with various platforms and access layers. The project also contributes to the EOSC Data Commons through the setup and provisioning of 'Data Spaces' that integrate scientific datasets and data analytics tools on top of the Compute Platform and deliver them as 'Thematic services' in EOSC.

EGI-ACE services are made available for users via the EOSC Portal and Marketplaces. High level of usability and quality required the project to integrate services, to align service capabilities and management processes with those required by EOSC. This deliverable reports on these integration and alignment activities. The overall collaboration with EOSC is covered in the multilateral 'Collaboration Agreement that EGI-ACE signed with the other projects of the INFRAEOSC-07 call, and with the EOSC Future project (INFRAEOSC-03). This deliverable focuses on the service integration, covering the following details:

- Onboarding - Making EGI-ACE services available for access in EOSC: The project already made available and delivers 30 services via the EOSC Portal. 2 more generic services (AppDB, Binder), and 5 thematic services (OpenRiskNet/NanoCommons, PROMINENCE, LOFAR Science Products, SeaDataNet and GBIF Cloud data space) are planned for onboarding in the next months.
- Technical integration - Making EGI-ACE services compatible with the EOSC Core and Exchange: EOSC does not have an Interoperability Framework yet. EGI-ACE is actively contributing to the establishment of the framework, providing requirements and feedback to the interoperability guidelines for EOSC Core, with a focus on the interfaces for interconnecting the EOSC Compute platform to the Core platform. In the meantime, EGI-ACE established compatibility with services of some of the other INFRAEOSC-07 projects (DICE, C-SCALE, OpenAIRE-Nexus). When the interoperability guidelines for the EOSC Core reaches an adequate level of maturity, EGI-ACE will make its Compute platform compliant with it.
- Service Management System (SMS) - Ensuring that the services are managed according to the EOSC expectations: All the onboarded EGI-ACE services meet the EOSC requirements for delivery, and most of them are operated according to ISO2000 IT Service Management standard of the EGI Foundation SMS. In 2022 the project will introduce the generic and thematic EGI-ACE services into a service management system that is coherent with the existing EGI Foundation SMS. D7.2 in March 2022 will provide a separate document about the SMS of the project.
- FAIRness - Ensuring that the EGI-ACE data space services provide Findable, Accessible, Interoperable, Reusable research results. We carried out a self-assessment using the 'FAIR Data Maturity Model' from the RDA FAIR Data Maturity Model Working Group. The findings revealed that compliance is quite high in the Findabile, Accessible and Reusable areas. The project should harvest this value through demonstrators and articles. The study also revealed that four of the data

spaces (WeNMR, OPENCoastS, PROMINENCE and VIP) are computational platforms that do not store and manage data for users. It is advised that these are distinguished from the rest of the data spaces in future communication and dissemination activities.

Based on the findings of this report the project will continue with integration and alignment activities in 2022, with the Conclusion section providing the list of actions for each of the 4 areas.

An updated version of this deliverable (D2.8), with the same title will be produced in 12 months from now, in December 2022.

1. Introduction

EGI-ACE is a 30-month project (Jan 2021 - June 2023) with a mission to empower researchers from all disciplines to collaborate in data- and compute-intensive research through free-at-point-of-use services.

EGI-ACE delivers two main results:

1. the 'EOSC Compute Platform', an integrated compute environment that federates compute and storage facilities with various platforms and access layers.
2. contributions to the EOSC Data Commons through the setup and provisioning of 'Data Spaces'. Data spaces integrate scientific datasets and data analytics tools on top of the Compute Platform and deliver them as 'Thematic services' in EOSC.

EGI-ACE is run by the EGI community, an international collaboration that federates the digital capabilities, resources and expertise of hundreds of national and international research communities in Europe and worldwide. With EGI-ACE the EGI Federation and research communities of pan-European relevance are joining efforts to deliver a distributed federated infrastructure that responds to the present and future needs of data-centric scientific computing in Europe through the EOSC.

The project was designed with the following main objectives:

1. Deliver the European Open Science Cloud Compute Platform and expand the supply-side.
2. Contribute to the implementation of the EU Data Strategy[1] and particularly its EOSC Data Commons to support the Green Deal, Health, Fundamental Research and Social Sciences and Humanities.
3. Integrate the EOSC Compute Platform with the EOSC Portal and the EOSC Core.
4. Contribute to the realization of a global Open Science Cloud.
5. Expand the demand-side of EOSC across sectors and disciplines.

After 12 months of operation these deliverable reports about the integration and alignment work that the project carried out to federate and deliver services in EOSC through the EOSC Portal. The overall collaboration with EOSC is covered in the multilateral Collaboration Agreement that EGI-ACE signed with the other projects of the INFRAEOSC-07 call, and with the EOSC Future project[2] (INFRAEOSC-03). The collaboration agreement covers the following joint activity areas:

Activity 1 Technical Activities:

- Resource Onboarding
- Architecture & Technical Interoperability
- Resource Provisioning and Technical Support

Activity 2 Uptake:

- Promotional activities

[1] European Data Strategy and EOSC Data Commons: https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy

[2] EOSC Future project: https://eoscfuture.eu/

- Joint engagement activities through events
- Joint EOSC Training activities

This document is focused on the Technical Activities, and starts with an introduction of the EGI-ACE service portfolio and support for thematic spaces (Section 2), then covers the different technical integration activities that exist between EGI-ACE and the EOSC Core:

- Service portfolio integration (Section 3)
- Technical integration (Section 4)
- Service Management System alignment (Section 5)
- FAIR maturity alignment (Section 6)

2. The EGI-ACE Service Architecture

2.1. EOSC Compute Platform services

The EOSC Compute Platform federates distributed compute and storage facilities to support processing and analytics via a set of services for distributed data and compute use cases. The EOSC Compute Platform architecture is organized in functional blocks as shown in Figure 1.

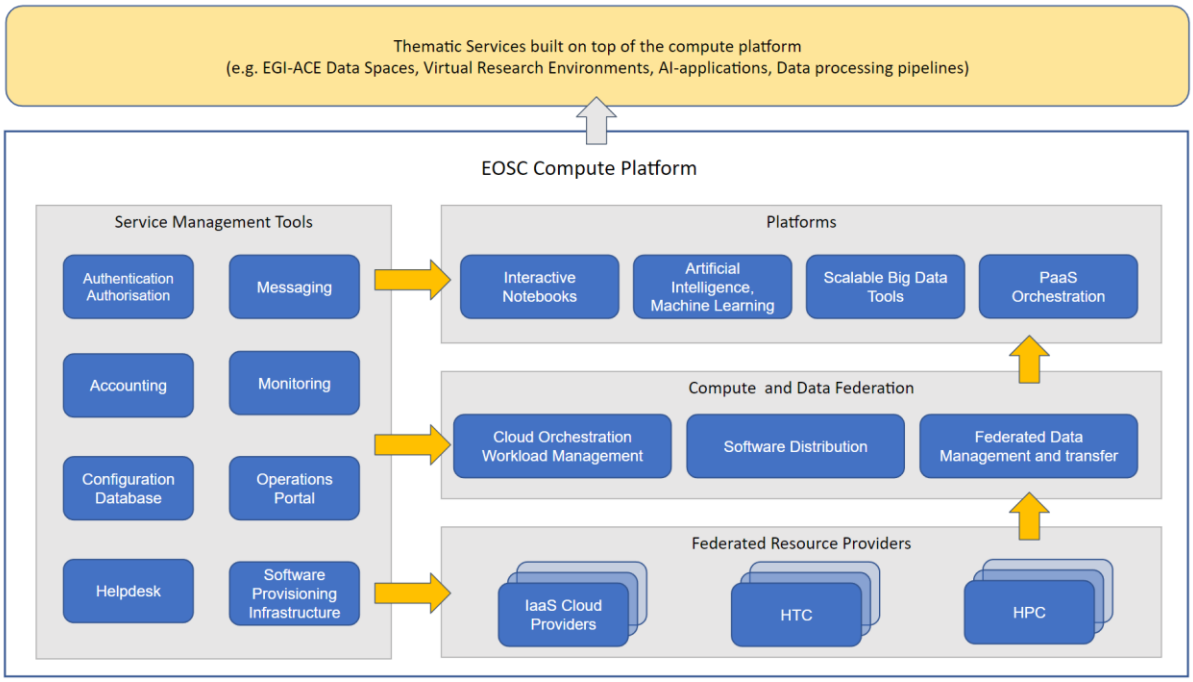

Figure 1. EOSC Compute Platform functional block diagram

Infrastructure layer

At the bottom of the architecture, the Federated Resource Providers deliver a hybrid infrastructure from academic and commercial providers for running/hosting research applications and data. Different types of providers are included in this layer:

- **IaaS Cloud Providers** provide access to Virtual Machine-based computing with associated Object and Block storage. These deliver a very flexible and customisable platform where users have complete control over the software and the supporting compute capacity. This flexibility of the computing platform enables the support of a variety of workloads: user gateways or portals, interactive computing platforms and almost any kind of data- and/or compute-intensive workloads.
- **HTC** (High Throughput Compute) provides access to large, shared computing systems for running computational jobs at scale. These allow researchers to analyse large datasets in an 'embarrassingly parallel' fashion, i.e. by splitting the data into small pieces, and executing thousands, or even more independent computing tasks

simultaneously, each processing one piece of data. HTC means the execution and management of many independent tasks over longer times.

- **HPC** (High Performance Compute) (to be available in 2022) supports very optimised application of machines that have a lot of interconnected processing units, with many dependent tasks that need large amounts of parallel computing along with a low latency and high bandwidth interconnection network.

Federation layer

The Compute Federation services orchestrate the execution of user workloads in the Federated Resource Providers. They exploit data locality by moving computing near data and facilitate application portability with the support of a diverse range of computing platforms (Cloud IaaS, HTC, HPC) and the interaction with software distribution tools (as VM images, container images or binaries directly). There are three services in this layer of the architecture:

- **Hybrid cloud orchestration** for the deployment of custom virtual infrastructure over multiple IaaS cloud backends within academic and commercial clouds;
- **Workload Manager** for the scheduling and execution of jobs in the federated resource providers (both cloud and HTC/HPC);
- **Software distribution**, for making software available at the Federated Resource Providers (e.g., as VM images).

The Federated Data services support exposing discoverable datasets and staging data into/out of the EOSC Compute Platform Cloud. The **Federated Data Management** services control the raw storage capacity offered by the Federated Resource Providers to deliver data products that can be transferred among the EGI-ACE providers, and between EGI-ACE providers and external data repositories. The Federated Data Management function uses the **Data Transfer** service to perform the transfers.

Platforms layer

A Platforms service area provides generic added-value services for scientific communities to build thematic services for end-users (typically for researchers). The platforms rely on the existing Compute Federation and Data Federation services to access the Federated Resource Providers and deliver **Interactive Notebooks**, **PaaS Orchestration** to facilitate the deployment of complex applications, and **Artificial Intelligence and Machine Learning** and **Scalable Big Data Tools** that can be reused in several research disciplines.

Service Management Tools

The Service Management Tools pillar delivers the functionality for services of all other areas to be integrated in the Federation. They support the operation of the EOSC Compute platform and integrate and interoperate with the EOSC Core that is run and is further developed in the EOSC Future project. EGI's **Authentication and Authorisation** service, called Check-in, is a key component of the architecture that enables using a common identity across all the layers and services of the EOSC Compute platform.

Configuration Database, **Monitoring**, **Accounting**, and **Helpdesk** services are also included in this area alongside with other non-technical services and coordination activities like Operations Management, and Security and Incident Response.

2.2. Support to Data spaces

The project contributes to the EOSC Data Commons through the setup and provisioning of 'Data Spaces' that integrates and hosts scientific datasets and data analytics applications on top of the Compute Platform. Data spaces are 'thematic services', i.e. they provide discipline specific capabilities for the end users. However, unlike other types of Thematic Services, Data Spaces host and integrate both data and applications into a single unit, enabling the scalable analysis of big datasets.
In contrast to simple "Publication of Open Data", a Public Data Space manages issues of access and use, as well as provides related tools and infrastructure. The EC's usage of the term "data space" assumes a public data space, so we interpret a public Data space as a "*public collection of FAIR, quality data and related resources consumed, produced and provided by identified participants, each respecting societal values and operating within an explicit framework of trust and governance*".
EGI-ACE Data spaces are built by scientific communities, research infrastructures and projects. The EGI-ACE consortium includes 13 Data spaces (See Figure 2), and supports additional ones that contact EGI-ACE via its open calls with the intention to set up Data spaces on top of the EOSC Compute Platform. Data spaces and other thematic services share the EOSC Compute Platform as a common architecture. The rest of their setup is specific to their scientific domains and target user groups.

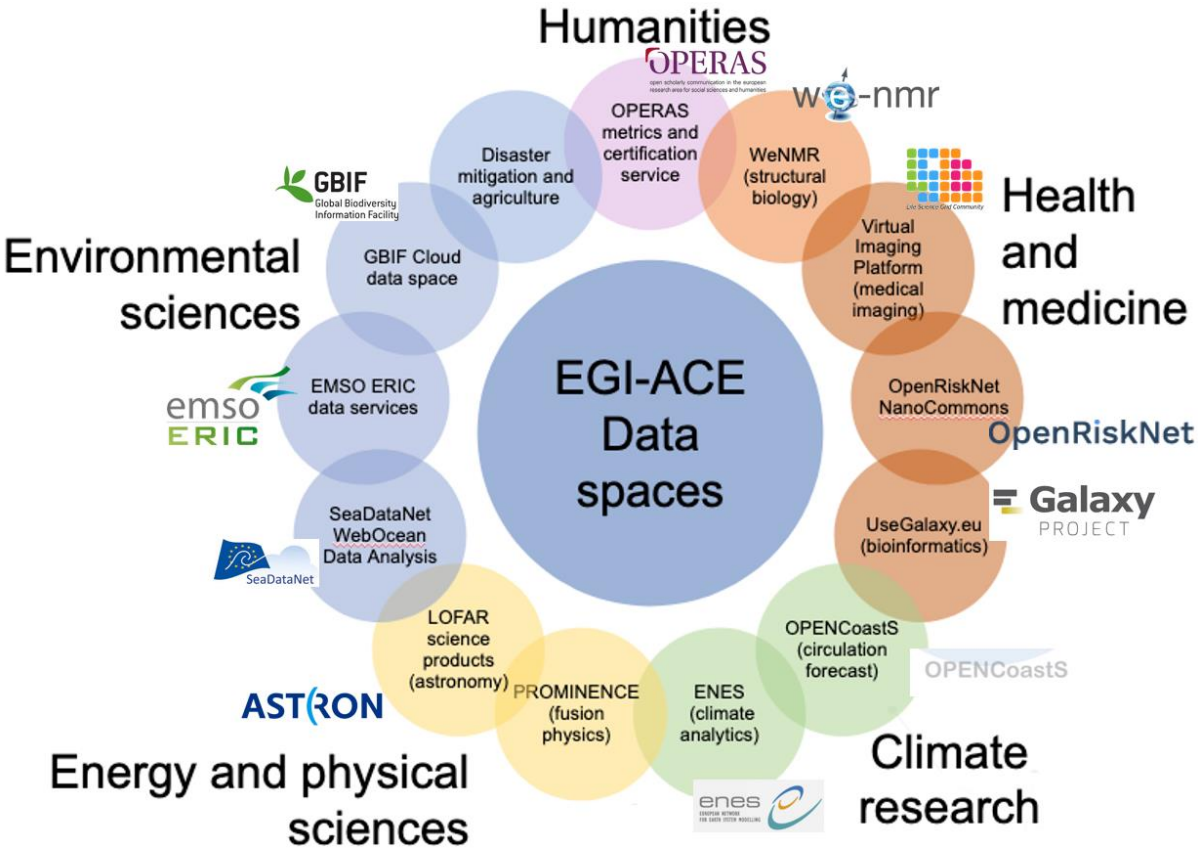

Figure 2. EGI-ACE Data Spaces included in the project consortium

2.3. Support for users and providers

Support for users and providers is an integral part of the project workplan. User support helps individual users and user communities in the uptake and use of the services; provider support helps new providers join the infrastructure, and existing providers in operating according to the EOSC and EGI standards. Without support there would be no (or much less) uptake as experience shows that e-infrastructures use and deployment are quite effort intensive activities where proper support can save a lot of time and effort for the customers.

- User support is based on the following 4 pillars: Shepherds, Training programme, User documentation, Site/service specific support teams.
- Provider support is based on the following 4 pillars: Provider documentations, EGI.eu and NGI operations teams (OMB)[3], Onboarding support (WP2), Integration support (in WP7 and EOSC-Future).

The structured and integrated support that EGI-ACE provides for users and service providers is an important distinctive feature that sets the EOSC Compute Platform apart from commercial compute services.

[3] EGI Operations Management Board: https://confluence.egi.eu/display/EGIBG/Operations+Management+Board

3. Service portfolio integration

3.1. EGI-ACE services in the EOSC Portal

Making a service available in EOSC requires the service to be visible and accessible to users via the EOSC Portal and its Marketplace. Column 2 of Table 1 below provides a service-by-service overview of the status of EGI-ACE services in the EOSC Portal.

Most of the EGI Foundation services, and some of the NGI services of the EOSC Compute Platform have been onboarded to the EOSC Portal during the EOSC-hub project. Task 2.2[4] of EGI-ACE provided assistance to the remaining providers to complete the onboarding task that, in most of the cases, meant filling out the service registration form on the EOSC Portal, in some of the cases meant registering the provider behind an already registered EOSC Service (for example registered GSI as provider for the EGI Cloud Compute service). Only two services are not registered in the EOSC Portal:

- AppDB - which is used as a component within the EGI Cloud Compute Service,
- Binder - which still needs to be integrated with the EGI Cloud Compute and Check-in services to open it for broader usage via EOSC.

A few of the EGI-ACE Data Space services have been onboarded to EOSC Portal during EOSC-hub, however most of these have to be updated in EGI-ACE due to the new functionalities they offer, or on-boarded as new services with the support of Task 2.2. Five Data Space services are not yet in EOSC Portal: OpenRiskNet/NanoCommons (unfunded), PROMINENCE, LOFAR Science Products, SeaDataNet and GBIF Cloud data space. These required substantial integration with the EOSC Compute Platform services in 2021, and are expected to start the onboarding to EOSC in the next months.

All the onboarded services of EGI-ACE are pulled together into a single page[5] within the EOSC Marketplace, using the 'Related platforms'

In the first nine months of the project the EGI-ACE services were used by more than 40,000 users. In particular, the data space services were used by over 37,000 users while the Compute Platform services by ~3,000 users via direct access (i.e. to manage virtual machines and jobs) and by 37,000 users via the hosted data spaces.

The EOSC Compute Platform layer served 56 user communities in this period. These communities are either part of the consortium (data space providers), applied for access via the EOSC Portal, via the EGI-ACE Open Calls, or with EGI Foundation directly. The total capacity requested by EGI-ACE use cases amounts to 105 million CPU hours. More detailed analysis of the usage of the EOSC Compute Platform is provided in EGI-ACE D2.2[6].

[4] EGI-ACE task 2.2: EOSC Portal alignment and coordination

[5] EGI-ACE services in the EOSC Marketplace (both delivered and supported services are included): https://marketplace.eosc-portal.eu/services?related_platforms=52

[6] EGI-ACE D2.2 EGI-ACE Strategic Plan: https://zenodo.org/record/5745168

The EGI-ACE project is leveraging the decade-long service delivery experiences of the EGI Foundation. The services of the EGI Foundation are governed by the EGI Council and are grouped into two service portfolios:

- External services[7] (or EGI services in short) target scientists, multinational projects and research infrastructures and are provided by EGI's federated cloud providers and data centres. The services can be requested by everyone involved in academic research and businesses via the EGI Marketplace and recently via the EOSC Marketplace. The External services are part of the 'Federated resource providers', the 'Compute and data federation' and the 'Platforms' layers of the EOSC Compute Platform (See Figure 1). EGI external services are sustained from a mix of national funds and EGI Council membership fees.
- Internal services[8] are provided for the benefit of the EGI Council members and affiliated organisations. The internal services complement the EGI Services for academia and business with tools designed to facilitate coordination and improve how the EGI Federation works together. The EGI internal services form the 'Service Management tools' pillar of the EOSC Compute Platform (See Figure 1). The EGI Internal services receive funding from the EGI Council membership fees, thus they are sustainable outside EOSC.

Table 1: Status of EGI-ACE services in EOSC Portal and in the EGI Catalogues

EGI-ACE service	Status in EOSC Portal Catalogue	Status in EGI Catalogue
EOSC Compute Platform: Federated resource provider services		
EGI Cloud Compute	Onboarded as EGI Cloud Compute	In external catalogue
SURF HTC	Registered as provider of the previously onboarded EGI HTC Compute	Provider of the HTC service in the external catalogue
EGI Container	Onboarded as EGI Cloud Container Compute	In external catalogue
Dynamic DNS Service	Onboarded as Dynamic DNS Service	Component of the EGI Cloud service
EGI Online Storage	Onboarded as EGI Online Storage	In external catalogue
EOSC Compute Platform: Compute and data federation services		
AppDB	Not in EOSC Catalogue	Component of the EGI Cloud service
DataHub	Onboarded as EGI DataHub	In external catalogue
FTS	Onboarded as EGI Data Transfer	In external catalogue (as Data Transfer)
Rucio	Onboarded as SCD STFC Rucio Data Management Service	Not in EGI catalogues (to be added to External portfolio as new service 'Data Orchestrator')

[7] EGI External services: https://www.egi.eu/services/
[8] EGI Internal services: https://www.egi.eu/internal-services/

OpenRDM	**Onboarded as openRDM EU**	**Not in EGI catalogues**
CVMFS	**Onboarded as STFC CVMFS Content Distribution Service**	**Not in EGI catalogue, but is in Alpha state in the external service portfolio as 'Content distribution' service**
EOSC Compute Platform: Platform services		
EC3	**Onboarded as Elastic Cloud Compute Cluster (EC3)**	**Not in EGI catalogues**
Infrastructure Manager	**Onboarded as Infrastructure Manager (IM)**	**Not in EGI catalogues**
DODAS	**Onboarded as Dynamic On Demand Analysis Service (DODAS Portal)**	**Not in EGI catalogues**
EGI Workload Manager (DIRAC)	**Onboarded as EGI Workload Manager**	**In external catalogue**
EGI Notebooks	**Onboarded as EGI Notebooks**	**In external catalogue**
Binder	**Not yet**	**Not in EGI catalogues**
Indigo PaaS Orchestrator (TOSCA)	**Onboarded as PaaS Orchestrator**	**Not in EGI catalogues**
DEEP training solution	**Onboarded as DEEP training facility**	**Not in EGI catalogues**
Service Management Tools		
Check-in	**Onboarded as EGI Check-in**	**In external and internal catalogues**
ARGO Messaging Service (AMS)	**Onboarded as EGI Operational Tools and delivered by EOSC Future**	**In internal catalogue**
ARGO Monitoring Service	**Onboarded as EGI Service Monitoring and delivered by EOSC Future**	**In internal catalogue**
GGUS Helpdesk Service	**Onboarded as EGI Helpdesk and delivered by EOSC Future**	**In internal catalogue**
GOCDB Configuration database	**Onboarded as EGI Configuration Database and delivered by EOSC Future**	**In internal catalogue**
Operations Portal	**Onboarded as EGI Operational Tools and delivered by EOSC Future**	**In internal catalogue**
Software Provisioning Infrastructure	**Onboarded as EGI Validated Software and Repository by EOSC Future**	**In internal catalogue**

Thematic services (Data Spaces)		
Health and medicine		
WeNMR	Onboarded as HADDOCK2.4 web portal	Not in EGI catalogues and no plan to include them in the existing catalogues
	Onboarded as DisVis web portal	Not in EGI catalogues and no plan to include them in the existing catalogues
	Onboarded as PowerFit web portal	Not in EGI catalogues and no plan to include them in the existing catalogues
	Onboarded as SpotOn web portal	Not in EGI catalogues and no plan to include them in the existing catalogues
	Onboarded as AMBER-based Portal Server for NMR structures (AMPS-NMR)	Not in EGI catalogues and no plan to include them in the existing catalogues
Virtual Imaging Platform	Onboarded as Virtual Imaging Platform	Not in EGI catalogues and no plan to include them in the existing catalogues
OpenRiskNet/ NanoCommons	Not yet	Not in EGI catalogues and no plan to include them in the existing catalogues
UseGalaxy.eu	Onboarded as European Galaxy Server	Not in EGI catalogues and no plan to include them in the existing catalogues
Climate research		
OpenCoastS	Onboarded as OPENCoastS Portal	Not in EGI catalogues and no plan to include them in the existing catalogues
ENES Data Space	Onboarded as ENES Data Space	Not in EGI catalogues and no plan to include them in the existing catalogues
Energy and physical sciences		
PROMINENCE	Not yet	Not in EGI catalogues and no plan to include them in the existing catalogues
LOFAR Science Products	Not yet	Not in EGI catalogues and no plan to include them in the existing catalogues
Environmental sciences		
SeaDataNet WebOcean Data Analysis	Not yet	Not in EGI catalogues and no plan to include them in the existing catalogues
EMSO ERIC data services	Onboarded as EMSO ERIC Data Portal	Not in EGI catalogues and no plan to include them in the existing catalogues
GBIF Cloud data space	Not yet	Not in EGI catalogues and no plan to include them in the existing catalogues

Disaster mitigation and agriculture	**Onboarding started for iCOMCOT portal**	**Not in EGI catalogues and no plan to include them in the existing catalogues**
Social sciences and humanities		
OPERAS Certification (DOAB)	**Onboarded as OPERAS Certification (DOAB)**	**Not in EGI catalogues and no plan to include them in the existing catalogues**

3.2. EGI-ACE services in EGI

The EOSC Compute Platform service portfolio was assembled during the EGI-ACE proposal preparation time based on the EGI Foundation services (internal and external), and based on additional services that EGI Council members wished to offer for pan-European access via the EOSC Portal.

Column 3 of Table 1 provides a service-by-service overview of the status of the EGI-ACE services in the EGI Service External/Internal portfolios. As can be seen, most of the EOSC Compute Platform services are already included in the EGI Catalogues (they are green in column 3):

- 5/5 federated resource providers are in the EGI Catalogue (external)
- 4/6 compute & data federation services are in the EGI Catalogue (external)
- 2/8 platform services are in the EGI Catalogue (external)
- 6/6 federation management tools are in the EGI Catalogue (internal)

EGI Catalogue membership offers better sustainability for a service outside the project. One of the focus for the service portfolio management task of EGI-ACE in 2022 has to be the clarification of the relationship between the EOSC Compute Platform and the EGI Federation governance and portfolios. This is expected to result in the onboarding of additional services from the EOSC Compute Platform to the EGI Service Portfolios, and in open and transparent processes for new providers and new services to join the EOSC Compute Platform. This work has been stated as strategic recommendation #2 in the recently published 'EGI-ACE Strategic Plan' D2.2 deliverable. EGI-ACE Task 2.2 and WP7 will work on this with the EGI-ACE Project Management Board (PMB), the EGI Executive Board (EB), and the EGI Services and Solutions Board (SSB).

Services of the EGI-ACE Data Spaces are not part of the EGI Service Portfolios. A subset of these Data Space services is delivered by institutes that are represented in the EGI Council directly (e.g. EMSO ERIC), or indirectly (e.g. University of Utrecht represented by SURF). Some of the Data Space providers are outside the network of EGI Council members. The Data Space services are all thematic services, i.e. are relevant to specific scientific disciplines (which is narrower or broader depending on the thematic service). Given that the EGI Portfolios currently include only services that are crosscutting across all disciplines and represent the 'common denominator' for big data science on e-infrastructures, the inclusion of EGI-ACE thematic services in the EGI Portfolios probably does not make sense. Thematic Services are better supported and promoted by the EGI community in alternative ways, such

as ‘Service Level Agreements’ and use case stories[9]. However, some thematic services that are appealing for several and broad disciplines (e.g. Galaxy for life sciences and environmental sciences) could be considered for stronger support by EGI beyond EGI-ACE. This needs to be discussed and decided in 2022 alongside the EOSC Compute Platform service action mentioned above.

[9] EGI Use cases: https://www.egi.eu/use-cases/

4. Technical interoperability

The EOSC Compute platform delivered by EGI-ACE will guarantee technical interoperability supporting the standards and interfaces of the EOSC Interoperability Framework (EIF), that is currently being built in the EOSC Future project in collaboration with other relevant initiatives, including the EGI-ACE project.

At the time of writing EOSC does not have an Interoperability Framework. The 'onboarding' process implements the compliance between EOSC Core and Exchange services.

Once developed, the EIF will facilitate interdisciplinary research and foster service/resource integration and composability. It will be made of a wide library of policies and interoperability guidelines describing standards and API's. The EIF will provide guidelines for providers to connect resources to EOSC-Exchange but will also provide guidelines to be adopted within services made available through EOSC-Core, supporting the composability and integration of resources across boundaries. EGI-ACE is an active actor in shaping the EIF and is collaborating with EOSC Future to define the interoperability guidelines for EOSC Core, providing requirements and feedback with a focus on the interfaces for interconnecting the EOSC Compute platform to the Core platform.

When the interoperability guidelines for the EOSC Core reaches an adequate level of maturity, EGI-ACE will update its services to be compliant as it is needed. However, the required effort is not expected to be major since many of the technologies adopted by EOSC Future to implement the EOSC Core (e.g. monitoring, AAI, accounting, etc) are also used by EGI to implement its service management tools. A clear example is AAI, the interoperability guidelines for AAI will be derived from the work delivered by the AARC project series[10] and further enhanced by AEGIS[11]. The AAI solution adopted by EGI-ACE, Check-in, is already compliant with many of these interoperability standards and joining the future EOSC AAI Federation will be a simple task for EGI-ACE.

Furthermore, EGI-ACE will contribute to the interoperability guidelines for EOSC Exchange developing guidelines in the technical areas of its interest (e.g. Computing, Data Platform for processing, etc.). As a first step in this direction, EGI-ACE submitted a proposal for a compute continuum Working Group in EOSC Future, which aims at defining a metadata schema as extension of the EOSC profile to describe the compute resources in the EOSC resource catalogue. This metadata schema will be a flexible and extensible specification for describing services providing access to generic computing resources covering as much as possible the complete compute continuum: cloud, HTC and HPC and potentially the edge, including access to hardware accelerators (e.g. GPUs) in all these systems whenever available. Having such specifications will mitigate the lack of standards in the compute service area and provide the basis to interoperation by enabling the discovery and potentially automated usage of compute services by the user communities. Thanks to this metadata schema, user communities and single users/LToS with computing needs for a specific scientific aim can be triaged and dispatched to the most appropriate kind of compute

[10] https://aarc-project.eu/
[11] https://aarc-project.eu/about/aegis/

platform according to their requirements. The WG was approved by the EOSC Future Technical Coordination Board and will start its operations in January 2022.

Finally, EGI-ACE will contribute to the validation of the EIF developing resource composability demonstrators and early adopter pilots in collaboration with EOSC Future and the other INFRAEOSC-07 projects. These pilots will foresee the combined usage of services from multiple providers with a level of automation expected to increase during the project lifetime, when more services and resources will adhere to the EOSC interoperability framework and its guidelines. EGI-ACE already contributes to the development of a demonstrator using the EGI Notebook service that was presented during the first EOSC Future interim review. As has been shown in the demo, the EGI-ACE Notebooks service is already interoperable with the B2DROP service[12] of DICE, and the Zenodo service[13] of OpenAIRE-Nexus.

In addition to the collaboration with EOSC Future, EGI-ACE will also contribute to the EOSC Association (EOSC-A) task forces with its own representatives (See Table 2). Such EOSC-A TFs are expected to steer the EOSC evolution and, in particular, the technical ones will also perform a key role in defining the future directions of the EIF. In this context, the experience gained dealing with real use cases will make the contributions of the EGI-ACE representatives valuable for the activities of these TFs.

Table 2: EGI-ACE participation in the technical Task Forces of the EOSC Association.

EOSC Task Force	Relevance for EGI-ACE	WPs/Tasks
Technical interoperability of data and services	The experience in implementing interoperability guidelines will be shared	T2.2
AAI Architecture	Ensuring the continued compatibility of the EGI Check-in service with the EOSC AAI requirements	T7.6
Researcher engagement and adoption	The uptake of EGI-ACE services, the understanding of community needs and areas of unmet demands can be fed into it	T2.3
FAIR metrics and data quality	The experience in conducting FAIR assessment of data and services can be shared with the TF	T2.2
Semantic interoperability	EGI-ACE data spaces can provide input, or take outputs for implementation	WP5

[12] B2DROP service: https://marketplace.eosc-portal.eu/services/b2drop
[13] Zenodo service: https://marketplace.eosc-portal.eu/services/zenodo

Infrastructure for quality research software	Alignment with the EGI software provisioning infrastructure	T7.6
PID policy and implementation	N/A	
Long-term data preservation	The EGI-ACE data spaces can provide input, or take outputs for implementation	WP5

5. Service Management System alignment

5.1. The EOSC Service Management System

EOSC Future operates the EOSC IT Management System (ITSM), covering the EOSC Core, and demanding some level of ITSM readiness from providers of the EOSC-Exchange. The EOSC Future ITSM builds on the ITSM that was laid down by EOSC-hub[14] to ensure a robust yet pragmatic service delivery in the EOSC federated infrastructure with different types of many-to-many relationships between users, providers and clients.

> **What is an IT Service Management System?**
> The key idea behind IT service management could be summarized like this: by following a service-oriented approach, an IT organisation (which may be everything from an internal IT department over a shared IT unit up to an external IT provider) is able to better understand what they do and offer, and how this is aligned to the needs of their customers and users. A Service Management System is the overall management system that controls and supports management of services within an organisation or federation. The SMS can be regarded as the entirety of interconnected policies, processes, procedures, roles, agreements, plans, related resources and other elements needed and used by a service provider to effectively manage the delivery of services to customers. By following the processes of the SMS the activities carried out to plan, deliver, operate and control the services become more structured and repeatable, with clearly defined responsibilities. All this helps an IT organisation to increase its level of professionalism and organisational maturity.

The EOSC Future Service Management System (SMS) is structured and organised into processes and procedures according to the FitSM IT Management standard[15], i.e. the same standard that is used by the EGI Foundation for the EGI External service[16] and Internal services[17]. FitSM is a free, pragmatic, lightweight and achievable standard aimed at facilitating service management in IT service provision, including federated scenarios. By defining requirements, the 14 processes of FitSM help EOSC service providers as is shown in Table 3.

Table 3: The 14 processes of FitSM, and their benefits for EOSC providers.

Process	Objective
Service portfolio management (SPM)	To define and maintain a service portfolio
Service level management (SLM)	To maintain a service catalogue, and to define, agree and

[14] https://www.eosc-hub.eu/eosc-hub-key-exploitable-results/#KER2
[15] FitSM IT Service Management standard: https://www.fitsm.eu/
[16] https://www.egi.eu/services/
[17] https://www.egi.eu/internal-services/

	monitor service levels with customers by establishing meaningful service level agreements (SLAs) and supportive operational level agreements (OLAs) and underpinning agreements (UAs) with suppliers
Service reporting management (SRM)	To specify all service reports and ensure they are produced according to specifications in a timely manner to support decision-making
Service availability and continuity management (SACM)	To ensure sufficient service availability to meet agreed requirements and adequate service continuity
Capacity management (CAPM)	To ensure sufficient capacities are provided to meet agreed service capacity and performance requirements
Information security management (ISM)	To manage information security effectively through all activities performed to deliver and manage services, so that the confidentiality, integrity and accessibility of relevant information are preserved
Customer relationship management (CRM)	To establish and maintain a good relationship with customers receiving services
Supplier relationship management (SUPPM)	To establish and maintain a healthy relationship with suppliers supporting the service provider in delivering services to customers, and monitor their performance
Incident and service request management (ISRM)	To restore normal / agreed service operation within the agreed time after the occurrence of an incident, and to respond to user service requests
Problem management (PM)	To investigate the root causes of (recurring) incidents in order to avoid future recurrence of incidents by resolving the underlying cause, or to ensure workarounds/temporary fixes are available
Configuration management (CONFM)	To provide and maintain a logical model of all configuration items (CIs) and their relationships and dependencies
Change management (CHM)	To ensure changes to CIs are planned, approved, implemented and reviewed in a controlled manner to avoid adverse impact of changes to services or the customers receiving services
Release and deployment management (RDM)	To bundle changes of one or more CIs to releases, so that these changes can be tested and deployed to the live environment together
Continual service improvement management (CSI)	To identify, prioritize, plan, implement and review improvements to services and service management

For each of these processes, as well as for a number of general aspects in the context of ITSM, FitSM (within the FitSM-1 document[18]) defines a small number of implementation

[18] FitSM-1 document - Requirements: https://www.fitsm.eu/downloads

requirements, while the FitSM-2 document[19] provides guidelines on the activities to set up and implement ITSM using these processes. The FitSM-3 document[20] describes the proposed roles to be assigned to execute the ITSM processes as part of a service management system.

At a base level, all onboarded services become in the scope of EOSC SPM when they are included into the EOSC Exchange Service Portfolio, and then publicly exposed in a Service Catalogue (the EOSC Portal and its Marketplace). How the scope of other EOSC SMS processes impacts on new onboarded services depends on the choices the service providers make for integrating with other EOSC Core services. For example, enabling 'ordering' (i.e. users have to request access to the service via the EOSC Marketplace) will bring the Exchange service partially into the scope of CRM, using the Helpdesk involves the Exchange service in the ISRM process, and so on. Additional integration activities may bring the services within the scope of other SMS processes.

5.2 EGI-ACE services in the EOSC and EGI SMSs

The minimum requirements of the EOSC SMS are met by any provider who successfully onboard services to the EOSC Portal. The EGI-ACE services that are onboarded to EOSC therefore already meet the EOSC Criteria. This section provides an overview of the SMS maturity of the EGI-ACE services independently from the EOSC requirements. As the section shows, most of the EGI-ACE services operate with a very mature SMS, and the project puts emphasis on lifting the SMS maturity of its whole portfolio.

The EGI Foundation has established a Management System for its IT-Services. This Service Management System (SMS) holds an ISO/IEC 20000 certification-1:2018. This is an international standard that outlines the requirements for design, transition, delivery and improvement of IT services that fulfil service requirements and provide value for both the customer and the service provider. The ISO/IEC 20000-1:2018 standard allows to demonstrate excellence and prove best practice in IT service management.

The EGI-ACE services relate to the EGI SMS in one of these three ways:

1. Some of the EGI-ACE EOSC Compute Platform services are already governed by the EGI Council (i.e. they are in the EGI External or Internal portfolios) therefore are covered by the EGI SMS. (See these services listed with green background in column 3 under the 'EOSC Compute Platform...' sections of Table 1.)
2. Some of the EGI-ACE EOSC Compute Platform services are not (yet) included in the EGI Service portfolios, therefore their SMS is not covered by the EGI Foundation SMS. (See these services listed with red or yellow background in column 3 under the 'EOSC Compute Platform...' sections of Table 1.)
3. EGI-ACE thematic services (Data space services) are not covered by the EGI SMS at all. (See these services listed under the 'Thematic services...' section of Table 1., with red background in column 3.)

[19] FitSM-2 document - Objectives and Activities: https://www.fitsm.eu/downloads
[20] FitSM-3 document - Role model: https://www.fitsm.eu/downloads

The project intends to raise the maturity of the services in group 2 by bringing them under the EGI SMS. The existing level of service management of these services is an important consideration for this work. The project therefore performed a preparatory activity, the maturity assessment of these services. From the operational perspective the important aspects of an SMS are to ensure that the services are monitored (to ensure high availability), they have a helpdesk (to ensure users and the monitoring can open trouble tickets), they areregistered in the Configuration DB (so changes can be tracked, and status information can be obtained for monitoring), and that Capacity plans and Availability and Continuity plans are available for them. The findings of this assessment are summarised in Table 4 below.

Table 4: Maturity assessment of the EGI-ACE Compute Platform services that are outside of the EGI Foundation SMS.

EGI-ACE service	Monitored?	Has a Helpdesk support unit?	Has an entry in the GOCDB configuration DB?	Has a capacity plan?	Has an availability and continuity plan?
EOSC Compute Platform: Compute and data federation services					
DynamicDNS	YES	YES	YES	NO	IN PROGRESS
EOSC Compute Platform: Compute and data federation services					
AppDB	YES	YES	YES	NO	YES
Rucio	YES	YES	YES	NO	IN PROGRESS
OpenRDM	YES	YES	YES	NO	NO
CVMFS	YES	YES	YES	YES	YES
EOSC Compute Platform: Platform services					
EC3	YES	YES	YES	NO	IN PROGRESS
Infrastructure Manager	YES	YES	YES	NO	IN PROGRESS
DODAS	NO	YES	YES	NO	NO
Binder	YES	YES	YES	YES	YES
Indigo PaaS Orchestrator (TOSCA)	NO	NO	NO	NO	NO
DEEP training solution	NO	NO	NO	NO	NO

Based on the findings we can estimate the amount of work required to bring these services under the EGI Foundation SMS. If these services would like to receive EGI Core funding (i.e. the funds collected by EGI Foundation from its members and redistributed to service providers), then we need to bring the services into the existing EGI governance and SMS.

If Core funding is not needed for these new services, then we can go for a 'lightweight' SMS, with e.g. requiring maturity in the areas covered in Table 4 above, as well as maturity in some user-facing activities especially Customer Relationship Management and Service Level Management. These requirements could be formulated in a new, 'lightweight SMS' that would apply to the services that EGI includes in its portfolio but does not consider as Core service. Deciding on the direction to make with these services is work in 2022 for Task 2.2, the affected service providers, and the EGI Services and Solutions Board (SSB).

5.3 Service Delivery Maturity Improvement

The EGI-ACE project is planning to implement a new initiative to drive up the level of service delivery maturity across all the EGI-ACE services. This will start with the EGI Internal services that are co-funded by EGI Foundation and will then extend to other services of the EGI-ACE Compute Platform, then to the data space/thematic services.

SMS maturity is a cornerstone of service delivery maturity. EGI has been delivering FitSM/ISO27k courses[21] for 6 years. The project will actively promote the increased update of these courses to key staff involved in service delivery for services delivered as part of EGI ACE. It will also identify areas of the SMSes of suppliers of services with which integration with the EGI SMS is required and ensure that these areas are integrated.

The requirements of SMS integration can mean different things for different processes and services. For example, for Information Security Management this implies providing details of the security contact responsible for service delivery at an organisation providing services within the EGI Federation and an ability to follow the correct EGI procedures when dealing with a security incident. For Incident and Service Request Management this implies being able to react to tickets raised against a service or resource. For internal services funded by EGI Foundation, the SMS integration requirements are more stringent; for example, extending to using the Change Management service run by EGI Foundation for changes which have the potential to directly affect other services; producing and periodically updating Capacity Management plans and Availability and Continuity plans.

Within the project we plan to drive improvements of service delivery by a process including three main parts:

1. Review, improvement and updating of documentation and training materials for all areas deemed to be within scope of this work for the target services, depending on the considerations outlined above.
2. Conduct self evaluation by service suppliers on themselves to verify whether the requirements are being met.
3. Conduct external evaluation to validate the self evaluation. This may be done with an external party, e.g. EGI Foundation.

[21] FitSM courses by EGI: https://indico.egi.eu/category/327/

Improvements to service availability may be made sustainable by incorporating the expectations outlined within this work in future Operational Level Agreements as part of subsequent phases of service delivery.

6. FAIR maturity of the EGI-ACE data spaces

6.1. The approach

Supporting 'Findable, Accessible, Interoperable, Reusable' (FAIR) research is one of the main goals of EOSC. Reaching FAIRness of research objects in EOSC is a shared responsibility of the EOSC-Core, and the services in the EOSC-Exchange. The most important contribution of EGI-ACE to FAIR research objects are the Data Spaces. We therefore carried out a FAIRness assessment of the EGI-ACE Data Spaces with the goals to understand their current level of FAIR maturity, and to identify areas for improvement to reach 'more FAIRness' in the remaining 18 months.

The project used the "FAIR Data Maturity Model: specification and guidelines"[22] from the RDA FAIR Data Maturity Model Working Group for the assessment work. This Model, developed between January 2019 and June 2020, was assembled by over 200 experts of research data from more than 20 countries. The main objectives of the RDA FAIR Data Maturity Model are to:

1. Define a sort of lingua franca for the evaluation of FAIRness on a general level, and
2. make results of FAIR assessment approaches comparable.

From the technical perspective, the RDA model proposes a set of:

- **Indicators**: the individual aspects of FAIRness that are evaluated on a service/objects/platform,
- **Priorities**: the relative importance of the indicators to achieve FAIRness, and
- **Maturity levels**: the way that the results of the evaluation of the indicators can be given a value. The model uses only two levels: True or False.

The Model includes **47 indicators**: 7 for the Findable, 12 for the Accessible, 12 for the Interoperable, and 16 for the Reusable aspect. Each indicator can be evaluated as TRUE, or FALSE for the object that is in the focus of the assessment. True means that the object satisfies the aspect of the indicator, false means that the object does not satisfy it.

The Model assigns one **of the following 3 Priorities for each indicator**:

- **Essential**: such an indicator addresses an aspect that is of the utmost importance to achieve the given aspect (F or A or I or R) under most circumstances, or, conversely, the given part of FAIRness would be practically impossible to achieve if the indicator were not satisfied.
- **Important**: such an indicator addresses an aspect that might not be of the utmost importance under specific circumstances, but its satisfaction, if at all possible, would substantially increase the F/A/I/R characteristic.
- **Useful**: such an indicator addresses an aspect that is nice-to-have but is not necessarily indispensable.

[22] FAIR Data Maturity Model Working Group: FAIR Data Maturity Model Specification and Guidelines: https://zenodo.org/record/3909563

We turned the Model into a **self-assessment survey** which was filled by the EGI-ACE Data Spaces. The survey itself can be found in the Appendix 1, with the same colours as above indicating the Priorities of the indicators (Red/Orange/Blue):

- **Red indicators** are Essential
- **Orange indicators** are Important
- **Blue indicators** are Useful

Appendix 2 includes all the received responses from all the data spaces to all the indicators. The Priority of the indicators are expressed with the same colour coding there too. The next subsections break down the responses by area:

- Section 6.3 details the responses for the “Findable” indicators.
- Section 6.4 details the responses for the “Accessible” indicators.
- Section 6.5 details the responses for the “Interoperable” indicators.
- Section 6.6 details the responses for the “Reusable” indicators.

To make the responses easier to understand, we came up with a summarising methodology that is used in Sections 6.3-6.6. for each Data Space we measure how many percent of the Essential and Important indicators are met within a given area. These summaries give a good feeling about ‘how close’ (or how far) is a specific Data Space to meet an F/A/I/R area. The left sides of the tables in Sections 6.2-6.5 list the Indicators of the given areas using the same Red/Orange/Blue colour coding for Priority as earlier. The right side of the tables provide the percentage values for satisfying the Essential and Important indicators. These numbers have coloured background:

- If **at least 90% of the Essential and Important indicators are met** then the background colour is **GREEN**. These Data Spaces require no/minimal further improvement in the respective F/A/I/R area.
- If **60-90% of the Essential and Important indicators are met** then the background colour is **ORANGE**. These Data Spaces require further work to achieve compliance in the respective F/A/I/R area.
- If **less than 60% of the Essential and Important indicators are met** the the background colour is **RED**. These data spaces require significant work to achieve compliance in the respective F/A/I/R area.

Section 6.6 provides a summary of the assessment findings across all the 4 areas, using the same **GREEN** / **ORANGE** / **RED** background colouring scheme.

6.2. Which data spaces were covered, which were not

We aimed to cover all the 13 data spaces of WP5 with the assessment. However, we found that two of the data spaces (OpenRiskNet/NanoCommons and OPERAS) are not ready yet for the assessment because their services are not mature enough. These two communities are self-funded in the project.

The assessment was not carried out to the full extent to the WeNMR, OPENCoastS, PROMINENCE and VIP Data Spaces, because they act as a computational platform but do not store and manage data for users. Within these platforms the users can upload (or define) input data, define/select data analysis/transformation steps, then download the results after

the computation is finished. Making the computational results FAIR is entirely the users' responsibility, outside these platforms. The FAIRness assessment therefore does not cover these 4 Data Spaces.

The 7 remaining data spaces have been fully assessed with the methodology and their results are included in the sections and tables below: useGalaxy.eu, GBIF, ENES, LOFAR Science Project, EMSO ERIC, SeaDataNet, Disaster Mitigation and Agriculture. These are already in production, or close to this stage so their FAIRness features are established.

6.3. Findable

Table 5 shows the summary of the compliance with the Essential and Important indicators of the Finable area. Observations and explanations:

- GBIF, LOFAR, SeaDataNet, useGalaxy.eu and EMSO ERIC Data Spaces are already fully compliant.
    - useGalaxy.eu does not provide unique identifiers - on purpose, but supports (meta)data exports to public archives, where users can get a global persistent identifier. The PID stored in the Galaxy database can be used to identify each dataset.
    - In LOFAR Data Space (meta)data deposited in the advanced science data repository will have a Handle PID associated with it. In addition, also the (Meta)data in the LOFAR instrument archive is assigned a custom persistent identifier maintained by the LOFAR Observatory. A rich set of metadata is supported to facilitate the discovery of the data by the astronomical Virtual Observatory registries. All the data deposited in the advanced data repository will have a globally unique Handle assigned to it.
- The ENES Data Space satisfies 5 indicators out of 7. In this Data Space metadata is not identified by a PID; it relies on CMIP6 Data Reference Syntax and controlled vocabularies. Every CMIP6 record is assigned to a universally unique identifier across the federation and each dataset/file is published with a PID.
- The Disaster Mitigation and Agriculture Data Space is still intensively working on developments and expects improvements in the area during the second part of the project[23].

Table 5: The outcome of the "Findable" indicators assessment.

Findable Indicators	**Data Space**	**% of Essential and Important indicators**

[23] The FAIRness maturity of the EGI-ACE Data Spaces will be reassessed during the course of the project at M24 in December 2022.

		satisfied[24]
RDA-F1-01M Metadata identified by a persistent identifier **RDA-F1-01D** Data identified by a persistent identifier **RDA-F1-02M** Metadata is identified by a globally unique identifier **RDA-F1-02D** Data is identified by a globally unique identifier **RDA-F2-01M** Rich metadata is provided to allow discovery **RDA-F3-01M** Metadata includes the identifier for the data **RDA-F4-01M** Metadata is offered in such a way that it can be harvested and indexed	**useGalaxy.eu**	**100%**
	GBIF	**100%**
	ENES	**71%**
	LOFAR Science Project	**100%**
	EMSO ERIC	**100%**
	SeaDataNet	**100%**
	Disaster Mitigation and Agriculture	**0%**

6.4. Accessible

Table 6 shows the summary of the compliance with the Essential and Important indicators of the Accessible area. Observations and explanations:

- GBIF, ENES, LOFAR, EMSO, SeaDataNet, Disaster Mitigation and Agriculture are already 90-100% compliant with the indicators:
    - GBIF metadata contains information to enable users to get access to the data. The GBIF metadata is uniquely referenced with DOI, the Ecological Metadata Language standard protocol[25] is used to expose the metadata in human readable format. GBIF data is uniquely identified via GBIF_id and can be accessed either using the Darwin Core standard[26] and RESTful JSON APIs[27].
    - In the ENES Data Space, PIDs are not officially used to reference metadata however, PIDs assigned to dataset/file and APIs allow to get a metadata

[24] There are only Essential indicators in the Findable area.
[25] https://eml.ecoinformatics.org/
[26] https://dwc.tdwg.org/
[27] https://www.gbif.org/pt/developer/summary

landing page and data hosts. Overall, the ENES data can be accessed either through a free access protocol (e.g.: HTTP, GridFTP, DAP), and other authN and AuthZ protocols. From a technical perspective, also the ENES Data Space is not fully compliant with these indicators (only 1 indicator is failing) but this will improve during the second half of the project.
    - In the LOFAR Data Space the metadata schema contains information to enable users to get access to the data using HTTP protocol. Also the data can be accessible with standard protocols provided by the SURF Data Repository.
    - The EMSO ERIC Data Space partially relies on the metadata of the Regional facilities to get access to the data. One of the goals of the Data Space is to provide an harmonization of the different metadata. Metadata and datasets can be accessed via standard protocols and REST APIs. Currently the Data Space is only failing with one indicator. The level of compliance of this Data Space with the Accessible indicator will improve during the second part of the project.
    - In the Disaster Mitigation and Agriculture Data Space data and publications are accessible in open access (over research data management tools such as Zenodo or Depositar) under the terms and conditions of the CCBY4.0 licence[28].
- useGalaxy.eu fails to meet 1 Essential (RDA-A1-03M) and 1 important indicator in this area (RDA-A1-01M).

Table 6: The outcome of the "Accessible" indicators assessment.

Accessible Indicators	**Data Space**	**% of Essential and Important indicators satisfied**
RDA-A1-01M Metadata contains information to enable the user to get access to the data **RDA-A1-02M** Metadata can be accessed manually (i.e. with human intervention)	**useGalaxy.eu**	**82%**
RDA-A1-02D Data can be accessed manually (i.e. with human intervention) **RDA-A1-03M** Metadata identifier resolves to a metadata record	**GBIF**	**100%**
RDA-A1-03D Data identifier resolves to a digital object **RDA-A1-04M** Metadata is accessed through standardised protocol	**ENES**	**91%**

[28] https://creativecommons.org/licenses/by/4.0/

RDA-A1-04D Data is accessible through standardised protocol **RDA-A1-05D** Data can be accessed automatically (i.e. by a computer program) **RDA-A1.1-01M** Metadata is accessible through a free access protocol **RDA-A1.1-01D** Data is accessible through a free access protocol **RDA-A1.2-01D** Data is accessible through an access protocol that supports authentication and authorisation **RDA-A2-01M** Metadata is guaranteed to remain available after data is no longer available	**LOFAR Science Project**	**100%**
	EMSO ERIC	**91%**
	SeaDataNet	**100%**
	Disaster Mitigation and Agriculture	**100%**

6.5. Interoperable

Table 7 shows the summary of the compliance with the Essential and Important indicators of the Interoperable area. Observations and explanations:

- This is the area with the lowest level of compliance overall. Only two data spaces, ENES and SeaDataNet meet all the Essential and Important criteria in this area. However note that this area does not have any Essential indicators, only Important and Useful ones. No-compliance in this area is therefore not as critical as in any of the other three areas.
- useGalaxy.eu, GBIF, Disaster Mitigation and Agriculture are above 80%.
- LOFAR meets 57%, EMSO ERIC 14%.

Table 7: The outcome of the “Interoperability” indicators assessment.

Interoperability Indicators	Data Space	% of Essential and Important indicators satisfied
RDA-I1-01M Metadata uses knowledge representation expressed in standardised format **RDA-I1-01D** Data uses knowledge representation expressed in standardised format	**useGalaxy.eu**	**86%**
RDA-I1-02M Metadata uses machine-understandable knowledge representation **RDA-I1-02D** Data uses machine-understandable	**GBIF**	**86%**

knowledge representation **RDA-I2-01M** Metadata uses FAIR-compliant vocabularies **RDA-I2-01D** Data uses FAIR-compliant vocabularies **RDA-I3-01M** Metadata includes references to other metadata **RDA-I3-01D** Data includes references to other data **RDA-I3-02M** Metadata includes references to other data **RDA-I3-02D** Data includes qualified references to other data **RDA-I3-03M** Metadata includes qualified references to other metadata **RDA-I3-04M** Metadata includes qualified references to other data	**ENES**	**100%**
	LOFAR Science Project	**57%**
	EMSO ERIC	**14%**
	SeaDataNet	**100%**
	Disaster Mitigation and Agriculture	**86%**

6.6. Reusable

Table 8 shows the summary of the compliance with the Essential and Important indicators of the Interoperable area. Observations and explanations:

- 4 of the Data spaces, GBIF, ENES, SeaDataNet, Disaster Mitigation and Agriculture are fully compliant with this area.
- useGalaxy is missing only one important indicator (RDA-R1.2-01M), but meeting all essential ones, reaching 89% overall compliance.
- LOFAR misses two important indicators (RDA-R1.1-03M, RDA-R1.3-02D) but meets all essential ones, reaching 78%.
- EMSO is missing 3 important indicators (RDA-R1.1-02M, RDA-R1.1-03M, RDA-R1.2-01M) and met all the essential ones, reaching 67% compliance.

Table 8: The outcome of the “Reusable” indicators assessment.

Reusable indicators	Data Space	% of Essential and Important indicators satisfied
RDA-R1-01M Plurality of accurate and relevant attributes are provided to allow reuse **RDA-R1.1-01M** Metadata includes information about the licence under which the data can be reused **RDA-R1.1-02M** Metadata refers to a standard reuse licence **RDA-R1.1-03M** Metadata refers to a machine-understandable reuse licence **RDA-R1.2-01M** Metadata includes provenance information according to community-specific standards **RDA-R1.2-02M** Metadata includes provenance information according to a cross-community language **RDA-R1.3-01M** Metadata complies with a community standard **RDA-R1.3-01D** Data complies with a community standard **RDA-R1.3-02M** Metadata is expressed in compliance with a machine-understandable community standard **RDA-R1.3-02D** Data is expressed in compliance with a machine-understandable community standard	**useGalaxy.eu**	**89%**
	GBIF	**100%**
	ENES	**100%**
	LOFAR Science Project	**78%**
	EMSO ERIC	**66%**
	SeaDataNet	**100%**
	Disaster Mitigation and Agriculture	**100%**

6.7. Overall findings

Table 9 brings together the first FAIR assessments of the EGI-ACE Data Spaces into a single view.

Findability, Accessibility and Reusability features are overall quite high, nearly all the studied data spaces are compliant or near-compliant. The project and the communities should harvest this value through demonstrators and articles. SeaDataNet is 100% compliant with the FAIR indicators and FAIRness success stories should be harvested from it.

The EMSO-ERIC Data Space is compliant with the Findable and Reusable indicators, further work on improving and harmonsing metadata schemes could bring the Interoperability area also to compliance, resulting in another data space that is compliant with all the four F-A-I-R areas.

Concerning future improvements, the Findable area of the Disaster Mitigation and Agriculture (one of the unfunded tasks) is the most critical development area. While Interoperability is the least satisfied area overall, this area does not have any 'essential' conformance indicator, has only important and 'useful' ones, non-conformance is not as critical here as for other areas. At the same time the studied data spaces have the biggest space for improvement overall in this area. Note however, that EGI-ACE does not include development effort specifically for FAIRness.

The findings about the 4 computational platforms in WP5 give a warning for the project to treat these differently in similar activities, and probably also distinguish them from the rest of the data spaces in project communication and dissemination activities.

Overall, the FAIRness maturity of the EGI-ACE Data Spaces will be re-assessed again at month 24 (December 2022), for the D2.8 deliverable.

Table 9: Overall summary of the FAIRness assessment of the EGI-ACE Data Spaces.

Data Space	Findability (F)	Accessible (A)	Interoperable (I)	Reusability (R)
useGalaxy.eu	**100%** Compliant	**82%** Compliant	**86%** Compliant	**89%** Compliant
GBIF	**100%** Compliant	**100%** Complaint	**86%** Compliant	**100%** Compliant
ENES	**71%** Compliant	**91%** Compliant	**100%** Compliant	**100%** Compliant
LOFAR Science Project	**100%** Compliant	**100%** Compliant	**57%** Compliant, but working on it	**78%** Compliant
EMSO ERIC	**100%** Compliant	**91%** Compliant	**14%** Compliant, but working on it	**66%** Compliant
SeaDataNet	**100%** Compliant	**100%** Compliant	**100%** Compliant	**100%** Compliant
Disaster Mitigation and Agriculture	**0%** Compliant, but working on it	**100%** Compliant	**86%** Compliant	**100%** Compliant
WeNMR	**Computational platforms that do not manage data. FAIRness cannot be interpreted for these.**			
Virtual Imaging Platform				

PROMINENCE	
OpenCoastS	
OpenRiskNet/ NanoCommons	**Unfunded data spaces that are not ready yet for FAIRness assessment.**
OPERAS DOAB	

7. Conclusions

The document provided details on the 4 main EOSC integration areas: Service portfolio integration, Technical integration, Service Management System alignment, FAIR maturity alignment.

The level of integration on all of these areas are already very high, but there are a number of actions remaining in each area for the next 18 months of the project:

1. Service portfolio integration: Reach full integration into the EOSC Portal, and increase integration into the EGI Service portfolios:
    a. 2 more generic services (AppDB, Binder), and 5 thematic services (OpenRiskNet/NanoCommons, PROMINENCE, LOFAR Science Products, SeaDataNet and GBIF Cloud data space) to finish the onboarding in the EOSC Portal.
    b. Finish the integration of the 'Data Orchestrator' and 'Content distribution' services to the EGI External Catalogue.
    c. Decide on and implement the integration of the OpenRDM, EC3, Infrastructure Manager, DODAS, Binder, Indigo PaaS Orchestrator, DEEP training solution services to the EGI External Catalogue.
2. Technical integration: Increase technical interoperability with the EOSC Core and with relevant services of EOSC-Exchange:
    a. Contribute to the establishment of the EOSC Interoperability Framework with a focus on the interfaces for interconnecting the EOSC Compute platform to the Core.
    b. Develop additional compatibility demonstrators across EGI-ACE services and non-EGI-ACE services of the EOSC-Exchange, particularly with the services of the INFRAEOSC-07 projects.
3. Service Management System: Improve the service delivery maturity of EGI-ACE providers by
    a. Review, improve and update the documentation and training materials.
    b. Conduct self and external evaluations to verify whether the service management requirements are being met.
    c. Provide training and support for service providers to increase the maturity of their service management system.
4. FAIR maturity: Increase FAIRness of the data spaces with the primary focus on:
    a. Improving the Findability features of the Disaster Mitigation and Agriculture data space.
    b. Increasing the Reusability features of the EMSO ERIC data space.
    c. Increasing the Interoperability features of the LOFAR Science Project and EMSO ERIC data spaces.

Appendix 1 - FAIR maturity assessment sheet

The template used in EGI-ACE to assess the FAIRness maturity of the Data Spaces is the following:

Indicators for Findable	Priority	Results
RDA-F1-01M Metadata identified by a persistent identifier. ● Principle: F1 (meta)data are assigned a globally unique and eternally persistent identifier. ● Description: This indicator evaluates whether or not the metadata is identified by a persistent identifier. A persistent identifier ensures that the metadata will remain findable over time and reduces the risk of broken links. ● Assessment details: The persistence of an identifier is determined by the commitment of the organisation that assigns and manages the identifier, so the evaluation of this indicator needs to take into account the persistence policy of that organisation. Such a commitment could be expressed by a university o r research institute, by a research infrastructure or by an organisation that issues formal identifiers, such as the International DOI Foundation. A possible way to evaluate this indicator is to verify that the identifier used for the metadata is listed in a registry service like the RDA-endorsed FAIRsharing.	**Essential**	TRUE/ FALSE
RDA-F1-01D Data identified by a persistent identifier. ● Principle: F1 (meta)data are assigned a globally unique and eternally persistent identifier. ● Description: This indicator evaluates whether or not the metadata is identified by a persistent identifier. A persistent identifier ensures that the metadata will remain findable over time and reduces the risk of broken links. ● Assessment details: The persistence of an identifier is determined by the commitment of the organisation that assigns and manages the identifier, so the evaluation of this indicator needs to take into account the persistence policy of that organisation. Such a commitment could be expressed by a university or research institute, by a research infrastructure or by an organisation that issues formal identifiers, such as the International DOI Foundation. A possible way to evaluate this indicator is to verify that the identifier used for the data is listed in a registry service like the RDA-endorsed FAIRsharing.	**Essential**	TRUE/ FALSE
RDA-F1-02M Metadata is identified by a globally unique identifier. ● Principle: F1 (meta)data are assigned a globally unique and eternally persistent identifier. ● Description: The indicator serves to evaluate whether the identifier of the metadata is globally unique, i.e. that there are no two identical identifiers that identify different metadata records. ● Assessment details: Global uniqueness of identifiers should be evaluated based on a description of how identifiers are assigned. Such a description should make it clear that the mechanism for assigning identifiers cannot possibly assign the same identifier to dif ferent resources, or assign an identifier that has already been assigned via	**Essential**	TRUE/ FALSE

some other mechanism/organisation. A possible way to evaluate this indicator is to verify that the identifier used for the data is listed in a registry service like the RDA-endorsed FAIRsharing.		
RDA-F1-02D Data is identified by a globally unique identifier. ● Principle: F1 (meta)data are assigned a globally unique and eternally persistent identifier. ● Description: The indicator serves to evaluate whether the identifier of the data is globally unique, i.e. that there are no two people that would use that same identifier for two different digital objects. ● Assessment details: Global uniqueness of identifiers should be evaluated based on a description of how identifiers are assigned. Such a description should make it clear that the mechanism for assigning identifiers cannot possibly assign the same identifier to different resources or assign an identifier that has already been assigned via some other mechanism/organisation. A possible way to evaluate this indicator is to verify that the identifier used for the data is listed in a registry service like the RDA-endorsed FAIRsharing.	**Essential**	TRUE/ FALS E
RDA-F2-01M Rich metadata is provided to allow discovery. ● Principle: F2: Data are described with rich metadata. ● Description: The indicator is about the presence of metadata, but also about how much metadata is provided and how well the provided metadata supports discovery. ● Assessment details: This indicator can be evaluated by verifying that metadata is provided. The amount of metadata to be provided may also be part of the metadata policy of the repository where the data is published.	**Essential**	TRUE/ FALS E
RDA-F3-01M Metadata includes the identifier for the data. ● Principle: F3: Metadata clearly and explicitly include the identifier of the data they describe. ● Description: The indicator deals with the inclusion of the reference (i.e. the identifier) of the digital object in the metadata so that the digital object can be accessed. ● Assessment details: This indicator can be evaluated by verifying that the identifier of the data is included in the metadata element that is specified for that purpose in the metadata standard used, for example in an "about" or "describes" predicate, or a Link Relation 16 such as "describes"/"describedBy".	**Essential**	TRUE/ FALS E
RDA-F4-01M Metadata is offered in such a way that it can be harvested and indexed. ● Principle: F4: (Meta)data are registered or indexed in a searchable resource. ● Description: The indicator tests whether the metadata is offered in such a way that it can be indexed. In some cases, metadata could be provided together with the data to a local institutional repository or to a domain-specific or regional portal, or metadata could be included in a landing page where it can be harvested by a search engine. The indicator remains broad enough on purpose not to limit the way how and by whom the harvesting and indexing of the data might be done. ● Assessment details: This indicator can be evaluated by verifying that the metadata is made available of or indexing. This is the case when the	**Essential**	TRUE/ FALS E

metadata is in fact harvested and indexed, for example in a general search engine or in a more restricted index, such as an institutional repository or a discipline-specific portal.		

Indicators for Accessible	Priority	Results
RDA-A1-01M Metadata contains information to enable the user to get access to the data. ● Principle: A1: (Meta)data are retrievable by their identifier using a standardised communication protocol. ● Description: The indicator refers to the information that is necessary to allow the requester to gain access to the digital object. It is about (i) restrictions to access the data (i.e. access to the data may be open, restricted or closed), (ii) the actions to be taken by a person who is interested to access the data, in particular when the data has not been published on the Web or (iii) specifications that the resources are available through specified authentication/authorisation system including single sign-on providers such as eduGAIN or through specialised solutions.	**Important**	TRUE/ FALSE
RDA-A1-02M Metadata can be accessed manually (i.e. with human intervention). ● Principle: A1: (Meta)data are retrievable by their identifier using a standardised communication protocol. ● Description: The indicator refers to any human interactions that are needed if the requester wants to access metadata. The FAIR principle refers mostly to automated interactions where a machine is able to access the metadata, but there may also be metadata that require human interactions. This may be important in cases where the metadata itself contains sensitive information. Human interaction might involve sending an e-mail to the metadata owner, or calling by telephone to receive instructions. ● Assessment details: The indicator can be evaluated by looking for information about the way that metadata can be accessed with human intervention, either in documentation, for example in a landing page, or in metadata about the metadata in cases where there is multi-layered metadata, for example using CatalogRecord in DCAT.	**Essential**	TRUE/ FALSE
RDA-A1-02D Data can be accessed manually (i.e. with human intervention). ● Principle: A1: (Meta)data are retrievable by their identifier using a standardised communication protocol. ● Description: The indicator refers to any human interactions that are needed if the requester wants to access the digital object. The FAIR principle refers mostly to automated interactions where a machine is able to access the digital object, but there may also be digital objects that require human interactions, such as clicking on a link on a landing page, sending an e-mail to the data owner, or even calling by telephone. ● Assessment details: The indicator can be evaluated by looking for information in the metadata that describes how access to the digital object can be obtained through human intervention.	**Essential**	TRUE/ FALSE

RDA-A1-03M Metadata identifier resolves to a metadata record. ● Principle: A1: (Meta)data are retrievable by their identifier using a standardised communication protocol. ● Description: This indicator is about the resolution of the metadata identifier. The identifier assigned to the metadata should be associated with a resolution service that enables access to the metadata record. ● Assessment details: The indicator can be evaluated by checking that the metadata can be accessed using its identifier. The evaluator or evaluation tool may also want to verify that the resolution delivers the correct metadata record.	**Essential**	TRUE/ FALSE
RDA-A1-03D Data identifier resolves to a digital object. ● Principle: A1: (Meta)data are retrievable by their identifier using a standardised communication protocol. ● Description: This indicator is about the resolution of the identifier that identifies the digital object. The identifier assigned to the data should be associated with a formally defined retrieval/resolution mechanism that enables access to the digital object or provides access instructions for access in the case of human-mediated access. The FAIR principle and this indicator do not say anything about the mutability or immutability of the digital object that is identified by the data identifier -- this is an aspect that should be governed by a persistence policy of the data provider. ● Assessment details: The indicator can be evaluated by invoking the mechanism specific to the protocol (e.g. GET for HTTP) and verifying that this delivers the digital object.	**Essential**	TRUE/ FALSE
RDA-A1-04M Metadata is accessed through standardised protocol. ● Principle: A1: (Meta)data are retrievable by their identifier using a standardised communication protocol. ● Description: The indicator concerns the protocol through which the metadata is accessed and requires the protocol to be defined in a standard. ● Assessment details: This indicator can be evaluated by looking at the way the metadata can be accessed. Common metadata access protocols are HTTP and FTP, Atom, 0AI-PMH and Web Services Metadata Exchange.	**Essential**	TRUE/ FALSE
RDA-A1-04D Data is accessible through standardised protocol. ● Principle: A1: (Meta)data are retrievable by their identifier using a standardised communication protocol. ● Description: The indicator concerns the protocol through which the digital object is accessed and requires the protocol to be defined in a standard. ● Assessment details: This indicator can be evaluated by looking at the way the data can be accessed. Common data access protocols are HTTP and FTP, DAP and JSON-RPC.	**Essential**	TRUE/ FALSE
RDA-A1-05D Data can be accessed automatically (i.e. by a computer program). ● Principle: A1: (Meta)data are retrievable by their identifier using a standardised communication protocol. ● Description: The indicator refers to automated interactions between machines to access digital objects. The way machines interact and grant access to the digital object. ● Assessment details: This indicator can be evaluated by resolving the	**Important**	TRUE/ FALSE

link to the data, e.g. by resolving the persistent identifier and verifying that the data is reached. In the common case that the identifier is an HTTP URI, this can be done using the HTTP GET method. The evaluator or evaluation tool may also want t o verify that the resolution delivers the correct data.		
RDA-A1.1-01M Metadata is accessible through a free access protocol. ● Principle: A1.1: The protocol is open, free and universally implementable. ● Description: The indicator tests that the protocol that enables the requester to access metadata can be freely used. Such free use of the protocol enhances data reusability. ● Assessment details: The indicator can be evaluated on the basis of information provided about whether the use of the protocol is free of charge. Common examples are HTTP and FTP.	**Essential**	TRUE/ FALSE
RDA-A1.1-01D Data is accessible through a free access protocol. ● Principle: A2: Metadata should be accessible even when the data is no longer available. ● Description: The indicator intends to verify that information about a digital object is still available after the object has been deleted or otherwise has become unavailable. If possible, the metadata that remains available should also indicate why the object is no longer available. ● Assessment details: The indicator can be evaluated by assessing whether an authentication and authorisation process is present in the protocol (e.g. HMAC).	**Important**	TRUE/ FALSE
RDA-A1.2-01D Data is accessible through an access protocol that supports authentication and authorisation. ● Principle: A2: Metadata should be accessible even when the data is no longer available. ● Description: The indicator intends to verify that information about a digital object is still available after the object has been deleted or otherwise has become unavailable. If possible, the metadata that remains available should also indicate why the object is no longer available. ● Assessment details: The indicator can be evaluated by assessing whether an authentication and authorisation process is present in the protocol (e.g. HMAC).	**Useful**	TRUE/ FALSE
RDA-A2-01M Metadata is guaranteed to remain available after data is no longer available. ● Principle: A2: Metadata should be accessible even when the data is no longer available ● Description: The indicator intends to verify that information about a digital object is still available after the object has been deleted or otherwise has become unavailable. If possible, the metadata that remains available should also indicate why the object is no longer available. ● Assessment details: The indicator can be evaluated on the basis of information provided about the life cycle of metadata and data, which should indicate that the metadata will remain available if the data is no longer available. This information is likely to be available from the repository where the metadata and data are stored.	**Essential**	TRUE/ FALSE

Indicators for Interoperable	Priority	Results
RDA-I1-01M Metadata uses knowledge representation expressed in standardised format. ● Principle: I1: (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. ● Description: The indicator serves to determine that an appropriate standard is used to express knowledge, for example, controlled vocabularies for subject classifications. ● Assessment details: The indicator can be evaluated by looking at information describing the way metadata values are expressed using controlled vocabularies, verifying that the standard used is appropriate for the domain and the type of digital object. Deciding on the appropriateness of the knowledge representation may be based on its inclusion in a registry like the one developed by FAIRsharing.	**Important**	TRUE/FALSE
RDA-I1-01D Data uses knowledge representation expressed in standardised format. ● Principle: I1: (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. ● Description: The indicator serves to determine that an appropriate standard is used to express knowledge, in particular the data model and format. ● Assessment details: The e indicator can be evaluated by looking at information about the data model and format, verifying that the standard used is appropriate for the domain and the type of digital object. Deciding on the appropriateness of the knowledge representation may be based on its inclusion in a registry like the one developed by FAIRsharing.	**Important**	TRUE/FALSE
RDA-I1-02M Metadata uses machine-understandable knowledge representation. ● Principle: I1: (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. ● Description: This indicator focuses on the machine-understandability aspect of the metadata. This means that metadata should be readable and thus interoperable for machines without any requirements such as specific translators or mappings. ● Assessment details: This indicator can be evaluated by looking at the knowledge representation model used for the expression of the metadata. Examples are RDF, OWL, JSON-LD and SKOS. Information about models and formats can be looked up in a registry like the RDA-endorsed FAIRsharing (see for example : https://fairsharing.org/standards/?q=&selected_facets=type_exact:model/format)	**Important**	TRUE/FALSE
RDA-I1-02D Data uses machine-understandable	**Important**	TRUE/F

knowledge representation. ● Principle: I1: (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. ● Description: This indicator focuses on the machine-understandability aspect of the metadata. This means that metadata should be readable and thus interoperable for machines without any requirements such as specific translators or mappings. ● Assessment details: This indicator can be evaluated by looking at the knowledge representation model used for the expression of the metadata. Examples are RDF, OWL, JSON-LD and SKOS. Information about models and formats can be looked up in a registry like the RDA-endorsed FAIRsharing (see for example : https://fairsharing.org/standards/?q=&selected_facets=type_exact:model/format).		ALSE
RDA-I2-01M Metadata uses FAIR-compliant vocabularies. ● Principle: : I2: (Meta)data use vocabularies that follow the FAIR principles ● Description: The indicator requires the vocabulary used for the metadata to conform to the FAIR principles, and at least be documented and resolvable using globally unique and persistent identifiers. The documentation needs to be easily findable and accessible. ● Assessment details: The indicator can be evaluated by verifying that each of the vocabularies used in the metadata is documented and resolvable using globally unique and persistent identifiers, with the documentation being easily findable and accessible. Typically, the reference to the specification of the vocabularies used will be included in the documentation of the digital object or the repository where it is kept.	**Important**	TRUE/FALSE
RDA-I2-01D Data uses FAIR-compliant vocabularies. ● Principle: I2: (Meta)data use vocabularies that follow the FAIR principles. ● Description: The indicator requires the controlled vocabulary used for the data to conform to the FAIR principles, and at least be documented and resolvable using globally unique and persistent identifiers. The documentation needs to be easily findable and accessible. ● Assessment details: The indicator can be evaluated by verifying that each of the vocabularies used in the data is documented and resolvable using globally unique and persistent identifiers, with the documentation being easily findable and accessible. Typically, the reference to the specification of the vocabularies used will be included in the documentation of the digital object or the repository where it is kept.	**Useful**	TRUE/FALSE
RDA-I3-01M Metadata includes references to other metadata. ● Principle: I3: (Meta)data include qualified references to other (meta)data.	**Important**	TRUE/FALSE

● Description: The indicator is about the way that metadata is connected to other metadata, for example through links to information about organisations, people, places, projects or time periods that are related to the digital object that the metadata describes. ● Assessment details: The indicator can be evaluated by looking at the occurrence of references to other metadata, for example ORCID for people or Geonames for places.		
RDA-I3-01D Data includes references to other data. ● Principle: I3: (Meta)data include qualified references to other (meta)data ● Description: This indicator is about the way data is connected to other data, for example linking to previous or related research data that provides additional context to the data. ● Assessment details: The indicator can be evaluated by looking at the presence of references to other data in the data. For example, there may be links to other resources in cells in a spreadsheet, or in RDF-based data.	**Useful**	TRUE/FALSE
RDA-I3-02M Metadata includes references to other data. ● Principle: I3: (Meta)data include qualified references to other (meta)data. ● Description: This indicator is about the way metadata is connected to other data, for example linking to previous or related research data that provides additional context to the data. Please note that this is not about the link from the metadata to the data it describes; that link is considered in principle F3 and in indicator RDA-F3-01M. ● Assessment details: The indicator can be evaluated by looking at the presence of references to other data in the metadata.	**Useful**	TRUE/FALSE
RDA-I3-02D Data includes qualified references to other data. ● Principle: I3: (Meta)data include qualified references to other (meta)data. ● Description: This indicator is about the way data is connected to other data. The references need to be qualified which means that the relationship role of the related resource is specified, for example that a particular link is a specification of a unit of measurement, or the identification of the sensor with which the measurement was done. ● Assessment details: The indicator can be evaluated by validating the presence of references with specification of the relationship role that the related resource has with the data object.	**Useful**	TRUE/FALSE
RDA-I3-03M Metadata includes qualified references to other metadata. ● Principle: I3: (Meta)data include qualified references to other (meta)data ● Description: This indicator is about the way metadata is connected to other metadata, for example to descriptions of related resources that provide additional context to the data.	**Important**	TRUE/FALSE

The references need to be qualified which means that the relationship of the related resource is specified, for example person Y is the author of dataset X. ● Assessment details: This indicator can be evaluated by looking at the presence of references with specification of the relationship that the related resource has to the described resource.		
RDA-I3-04M Metadata includes qualified references to other data. ● Principle: : I3: (Meta)data include qualified references to other (meta)data. ● Description: This indicator is about the way metadata is connected to other data. The references need to be qualified which means that the relationship role of the related resource is specified, for example dataset X is derived from dataset Y. ● Assessment details: This indicator can be evaluated by looking at the presence of references with specification of the relationship role that the related resource has with the described resource.	**Useful**	TRUE/FALSE

Indicators for Reusable	**Priority**	**Results**
RDA-R1-01M Plurality of accurate and relevant attributes are provided to allow reuse. ● Principle: R1: (Meta)data are richly described with a plurality of accurate and relevant attributes. ● Description: The indicator concerns the quantity but also the quality of metadata provided in order to enhance data reusability. ● Assessment details: This indicator can be evaluated with the help of standards registries such as the RDA-endorsed FAIRsharing (see for example: https://fairsharing.org/standards/?q=/format&selected_facets=type_exact:reporting%20guideline).	**Essential**	TRUE/FALSE
RDA-R1.1-01M Metadata includes information about the licence under which the data can be reused. ● Principle: R1.1: (Meta)data are released with a clear and accessible data usage license. More information about that principle can be found here. ● Description: This indicator is about the information that is provided in the metadata related to the conditions (e.g. obligations, restrictions) under which data can be reused. In the absence of licence information, data cannot be reused. ● Assessment details: This indicator can be evaluated by looking in the metadata for licence information. This information may be in human-readable text; machine-understandability of the information is covered in indicator RDA-R1.1-03M.	**Essential**	TRUE/FALSE
RDA-R1.1-02M Metadata refers to a standard reuse licence. ● Principle: R1.1: (Meta)data are released with a clear and accessible data usage license. ● Description: This indicator requires the reference to the conditions of reuse to be a standard licence, rather than a locally defined licence.	**Important**	TRUE/FALSE

● Assessment details: The indicator can be evaluated by verifying that the licence is indeed a standard licence. Examples of standard licences are: Creative Commons licences, Open Data Commons.		
RDA-R1.1-03M Metadata refers to a machine-understandable reuse licence. ● Principle: R1.1: (Meta)data are released with a clear and accessible data usage license. ● Description: This indicator is about the way that the reuse licence is expressed. Rather than being a human-readable text, the licence should be expressed in such a way that it can be processed by machines, without human intervention, for example in automated searches. ● Assessment details: The indicator can be evaluated by verifying that the link to the licence resolves to a machine-understandable expression of the conditions. An example of such a machine-understandable expression is the RDF expression of Creative Commons licences, or the various serialisations of the Open Data Rights Language (ODRL).	**Important**	TRUE/FALSE
RDA-R1.2-01M Metadata includes provenance information according to community-specific standards. ● Principle: R1.2: (Meta)data are associated with detailed provenance. ● Description: This indicator requires the metadata to include information about the provenance of the data, i.e. information about the origin, history or workflow that generated the data, in a way that is compliant with the standards that are used in the community for which the data is curated. ● Assessment details: The indicator can be evaluated by verifying that the licence is indeed a standard licence. Examples of standard licences are: Creative Commons licences, Open Data Commons.	**Important**	TRUE/FALSE
RDA-R1.2-02M Metadata includes provenance information according to a cross-community language. ● Principle: R1.2: (Meta)data are associated with detailed provenance. ● Description: This indicator requires that the metadata provides provenance information according to a cross-domain language. ● Assessment details: The indicator can be evaluated by assessing whether a cross-domain language is used for provenance information (such as PROV-O).	**Useful**	TRUE/FALSE
RDA-R1.3-01M Metadata complies with a community standard. ● Principle: R1.3: (Meta)data meet domain-relevant community standards. ● Description: This indicator requires that metadata complies with community standards. ● Assessment details: The indicator can be evaluated by verifying that the metadata follows a community standard. A service like the RDA-endorsed FAIRsharing or the Metadata Standards Catalog could be helpful to identify the relevant standards.	**Essential**	TRUE/FALSE
RDA-R1.3-01D Data complies with a community standard. ● Principle: R1.3: (Meta)data meet domain-relevant community standards. ● Description: This indicator requires that data complies with community standards. ● Assessment details: The indicator can be evaluated by verifying that the data follows a community standard. A service like the RDA-	**Essential**	TRUE/FALSE

endorsed FAIRsharing could be helpful to identify the relevant standards.		
RDA-R1.3-02M Metadata is expressed in compliance with a machine-understandable community standard. ● Principle: R1.3: (Meta)data meet domain-relevant community standards. ● Description: This indicator requires that the metadata follows a community standard that has a machine-understandable expression. ● Assessment details: This indicator can be evaluated by verifying that the community standard used for the metadata has a machine-understandable expression.	**Essential**	TRUE/FALSE
RDA-R1.3-02D Data is expressed in compliance with a machine-understandable community standard. ● Principle: R1.3: (Meta)data meet domain-relevant community standards. ● Description: This indicator requires that the data follows a community standard that has a machine-understandable expression. ● Assessment details: This indicator can be evaluated by verifying that the community standard used for the data has a machine-understandable expression.	**Important**	TRUE/FALSE

Appendix 2 - FAIR maturity assessment responses

		GBIF UseGalaxy	ENES	LOFAR	EMSO ERIC	SeaDataNet	Dis.Mitigation	
Findable	RDA-F1-01M	T	T	F	T	T	T	F
	RDA-F1-01D	T	T	T	T	T	T	F
	RDA-F1-02M	T	T	F	T	T	T	F
	RDA-F1-02D	T	T	T	T	T	T	F
	RDA-F2-01M	T	T	T	T	T	T	F
	RDA-F3-01M	T	T	T	T	T	T	F
	RDA-F4-01M	T	T	T	T	T	T	F
		100%	100%	71%	100%	100%	100%	0%
Accessible	RDA-A1-02M	T	T	T	T	T	T	T
	RDA-A1-02D	T	T	T	T	T	T	T
	RDA-A1-03M	F	T	F	T	T	T	T
	RDA-A1-03D	T	T	T	T	T	T	T
	RDA-A1-04M	T	T	T	T	T	T	T
	RDA-A1-04D	T	T	T	T	T	T	T
	RDA-A1-05D	T	T	T	T	T	T	T
	RDA-A1.1-01M	T	T	T	T	T	T	T
	RDA-A1.1-01D	T	T	T	T	T	T	T
	RDA-A1.2-01D	T	F	T	T	T	T	F
	RDA-A2-01M	T	T	T	T	F	T	T
		82%	100%	91%	100%	91%	100%	100%
Interoperable	RDA-I1-01D	T	T	T	T	F	T	T
	RDA-I1-02M	T	T	T	F	F	T	T
	RDA-I1-02D	T	T	T	F	F	T	T
	RDA-I2-01M	F	T	T	F	T	T	T
	RDA-I2-01D	F	T	T	F	F	T	T
	RDA-I3-01M	T	T	T	T	F	T	T
	RDA-I3-01D	T	T	F	F	F	F	F
	RDA-I3-02M	T	T	T	T	F	T	F
	RDA-I3-02D	T	F	F	F	F	F	F
	RDA-I3-03M	T	F	T	T	F	T	F
	RDA-I3-04M	T	F	T	T	F	F	F
		86%	86%	100%	57%	14%	100%	86%
Reusable	RDA-R1.1-01M	T	T	T	T	T	T	T
	RDA-R1.1-02M	T	T	T	T	F	T	T
	RDA-R1.1-03M	T	T	T	F	F	T	T
	RDA-R1.2-01M	F	T	T	T	F	T	T
	RDA-R1.2-02M	T	F	F	F	F	F	F
	RDA-R1.3-01M	T	T	T	T	T	T	T
	RDA-R1.3-01D	T	T	T	T	T	T	T
	RDA-R1.3-02M	T	T	T	T	T	T	T
	RDA-R1.3-02D	T	T	T	F	T	T	T
		89%	100%	100%	78%	67%	100%	100%