EOSC Future

# D2.2a

# Mapping of EOSC Readiness of EU MS/AC

The EOSC Future project is co-funded by the European Union Horizon Programme call INFRAEOSC-03-2020, Grant Agreement number 101017536

Version 1.0
September 2021

# D2.2a / Mapping of EOSC Readiness of EU MS/AC

Lead by **CESSDA ERIC**
Authored by Bojana Tasić (CESSDA/FORS), Carsten Thiel (CESSDA)
Reviewed by Rudolf Dimper (ESRF) & Rocío Mantero Cejudo (LIFEWATCH)

## Dissemination Level of the Document

Public

## Abstract

This deliverable provides an overview of exiting and ongoing initiatives collecting indicators for EOSC Readiness. Based on the landscaping work previously carried out, existing frameworks including the SRIA are summarised and available data on EOSC adoption is considered. The resulting overview of alignment directions, concrete action steps and open questions will form the basis for the work in WP2 Activity 2.2.1 of EOSC Future together with the parties identified in the report.

## Version History

| Version | Date | Authors/Contributors | Description |
|---|---|---|---|
| V0.1 | 16/07/2021 | Bojana Tasić (CESSDA-FORS), Carsten Thiel (CESSDA-FORS), | Initial draft |
| V0.2 | 20/07/2021 | Bojana Tasić (CESSDA-FORS), Carsten Thiel (CESSDA), Gareth O'Neil (TGB) | Incorporating feedback |
| V0.3 | 21/07/2021 | Bojana Tasić (CESSDA-FORS), Carsten Thiel (CESSDA), Ivana Ilijašić Veršić (CESSDA) | Incorporating feedback |
| V0.4 | 26/07/2021 | Bojana Tasić (CESSDA-FORS), Carsten Thiel (CESSDA), Franciska de Jong (CLARIN) | Incorporating feedback |
| V0.5 | 29/07/2021 | Bojana Tasić (CESSDA-FORS), Carsten Thiel (CESSDA), Natalia Manola (OpenAIRE) | Incorporating feedback |
| V0.6 | 29/07/2021 | Bojana Tasić (CESSDA-FORS), Carsten Thiel (CESSDA), Ivana Ilijašić Veršić (CESSDA) | Incorporating feedback |
| V0.7 | 30/07/2021 | Bojana Tasić (CESSDA-FORS), Carsten Thiel (CESSDA) | Submitted for QA |
| V0.8 | 13/09/2021 | Bojana Tasić (CESSDA-FORS), Carsten Thiel (CESSDA), Rudolf Dimper (ESRF) | Incorporated Reviewer comments |
| V0.9 | 14/09/2021 | Carsten Thiel (CESSDA), Athanasia Spiliotopoulou (JNP), Rocío Mantero Cejudo (LIFEWATCH) | Incorporated Reviewer comments |
| V1.0 | 30/09/2021 | Bojana Tasić (CESSDA-FORS), Carsten Thiel (CESSDA), Ron Dekker (TGB), Mike Chatzopoulos (ATHENA) | Final Version submitted to EC |

## Copyright Notice

 The EOSC Future project is co-funded by the European Union Horizon Programme call INFRAEOSC-03-2020, Grant Agreement number 101017536.

# Table of Contents



# Table of Tables



# Table of Figures

# List of Abbreviations

| Acronym | Definition |
|---|---|
| **API** | Application Programming Interface |
| **DM** | Data Management |
| **EC** | European Commission |
| **EOSC-A** | EOSC Association |
| **EOSC-SB** | EOSC Steering Board |
| **ESFRI** | European Strategy Forum for Research Infrastructures |
| **EU** | European Union |
| **KPI** | Key Performance Indicator |
| **MS/AC** | Member States and Associated Countries |
| **OS** | Open Science |
| **R&I** | Research and Innovation |
| **RPO** | Research Performing Organisation |
| **SRIA** | Strategic Research and Innovation Agenda |

# 1 Introduction

This document is the first version of this deliverable reporting on the work carried out in WP2 Activity 2.2.1 of the EOSC Future project. The activity aims at raising the profile of EOSC Portal in EU Member States and Associated Countries (MS/AC) with low EOSC uptake, with a particular focus on the Western Balkans. The second version of the deliverable is scheduled for spring 2023.

For the purpose of this deliverable, the portal that was originally developed by the eInfraCentral, EOSCpilot, EOSC-hub and OpenAIRE-Advance projects and then unified[1] and further improved under the EOSC Enhance[2] project will be referred to as the *EOSC Enhance Portal*[3]. Simultaneously, thematic and regional portals are being developed by the cluster projects (ENVRI-FAIR[4], EOSC-Life[5], ESCAPE[6], PANOSC[7] and SSHOC[8]) and the regional projects (EOSC-Nordic[9], EOSC-Pillar[10], EOSC-Synergy[11], ExPaNDS[12] and NI4OS-Europe[13]) under the INFRAEOSC-04-2018 and INFRAEOSC-05b-2018-2019 calls respectively, referred to as *regional and thematic portals.* Each of these addresses different users and needs. While the EOSC Enhance Portal and some of the others focus primarily on services, the representation of other Open Science resources, such as data, publications, as well as software and software code, are also important in realising the EC Open Science policy and are being collected and visualised in the OpenAIRE Open Science Observatory. For the purpose of this deliverable, the term *resource* is understood to include services, data, publication, software and possibly other types of resources that will be identified as relevant in the future. Part of the further landscaping work will consist of identifying all relevant portals and databases and their target audiences for inclusion in the EOSC Future data collection.

As this deliverable is intended to form the basis for the work to be carried out over the course of the project, its content is limited to an outline of the main existing indicators and available data and their methods of collection. It is intended to start discussions with both the EOSC Association (ESOC-A) and the EOSC Steering Board (ESOC-SB) to ensure optimal alignment in understanding of what is being measured, by whom, why these indicators are being selected and how they will be used.

[1] See 'A brief history': https://eosc-portal.eu/about-eosc-portal, accessed 27 July 2021
[2] https://eosc-portal.eu/enhance, accessed 27 July 2021
[3] Currently available at https://eosc-portal.eu/
[4] https://envri.eu/home-envri-fair/, accessed 27 July 2021
[5] https://www.eosc-life.eu/, accessed 27 July 2021
[6] https://projectescape.eu/, accessed 27 July 2021
[7] https://www.panosc.eu/, accessed 27 July 2021
[8] https://sshopencloud.eu/, accessed 27 July 2021
[9] https://www.eosc-nordic.eu/, accessed 27 July 2021
[10] https://www.eosc-pillar.eu, accessed 27 July 2021
[11] https://www.eosc-synergy.eu/, accessed 27 July 2021
[12] https://expands.eu/, accessed 27 July 2021
[13] https://ni4os.eu/, accessed 27 July 2021

# 2 Layers and Scope

The scope of this deliverable is to identify both known indicators as well as available data to derive these from and to identify what other information is needed to define readiness indicators. Based on this, sources for gathering the missing information will have to be identified and data has to be collected. The natural collection platform will be the EOSC Observatory, to be developed in WP2 Task 2.3 of EOSC Future.

This work is closely related to that of WP2 Task 2.1, which distinguishes four possible layers relevant for the data to be monitored by the EOSC Observatory, shown in Figure 2.1.

- **Layer 1:** Indicators that support monitoring of EOSC Readiness, which should be validated by the MS/AC of the EOSC-SB.
- **Layer 2:** The KPIs defined in the SRIA and validated by the members of the EOSC-A.
- **Layer 3:** The in-kind contributions provided by the member states to the EOSC Partnership and validated by the EC.
- **Layer 4:** Policies relevant for the EOSC ecosystem as shared by the EOSC community.

Contributions to EOSC could also come from outside of member states or be funded through EC grants. Of these layers relevant for the overall EOSC Landscape and monitoring and the work of WP2, is Layer 1 which is the focus of this deliverable. At the outset, EOSC Readiness should be a measure of the ability to fulfil KPIs defined by the EOSC-A, the EOSC-SB, the EC or the member states themselves.

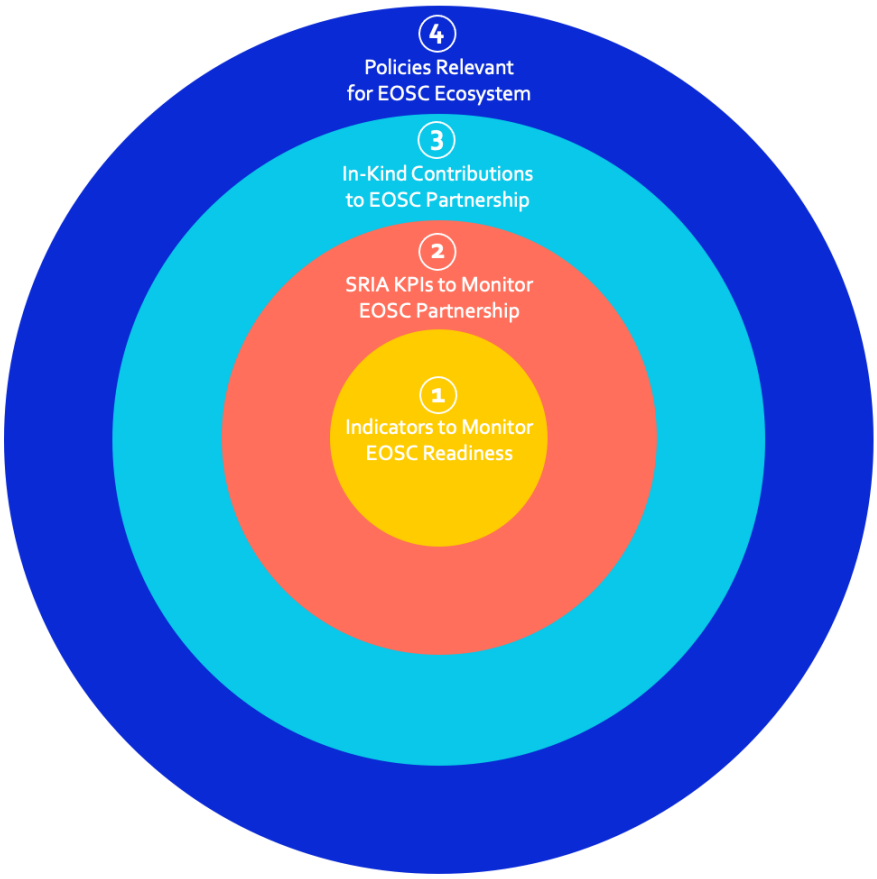


*Figure 2.1: Four possible layers of data that EOSC Observatory could collect, credit: Gareth O'Neil*

# 3 Readiness as the basis for the work carried out under WP2 Activity 2.2.1

Readiness in the context of this deliverable is understood as the technical preparedness for onboarding resources to the EOSC, representing Layer 1 in Figure 2.1. The concrete implementation of any such solution is dependent on the policies and priorities influencing the decision process in the higher layers of the figure. Work on indicators measuring policy implementation has already been carried out to a significant extent by the EOSC Landscape Working Group in their report 'Landscape of EOSC-Related Infrastructures and Initiatives' from November 2020 [3] and the work has been continued by the regional projects in the 'Second Working Proposal for Living Indicators to Monitor MS Progresses Towards EOSC Readiness' [1] with funding from the EOSCsecretariat.eu[14] project in 2021. The background for any work in the direction is set by the Strategic Research and Innovation Agenda (SRIA) [6] , Layer 2 in Figure 2.1. A summary of the indicators defined therein is given in Sections 0 and 0 below.

The foundation for evaluating technical readiness by measuring onboarding is collecting and mapping information about the resources available in EOSC.

The Open Science Observatory[15] already monitors the availability of data, publications and software across Europe. The information collected there will be one of the sources of the EOSC Observatory.

A canonical platform for the monitoring of available services does not yet exist, apart from information already provided to the EOSC Enhance Portal during onboarding. While the portal is an obvious candidate for data collection, its contents is dependent on an action taken to onboard to EOSC, following rules and criteria that affect the data.

The number of resources identified in this manner do however represent those that already are onboarded. In order to support the MS/AC in onboarding, they must be compared to the total available resources to identify gaps in uptake. Section 0 examines possibilities for this.

## 3.1 Other Portals and user adoption

While the onboarding through listing in dedicated EOSC catalogues is the obvious choice to channel information about available resources to researchers, not all catalogues are designed for same workflows and data sources. Regional or thematic portals can form a steppingstone and onboard all their listed services directly into the EOSC Enhance Portal through its API, or they might do the opposite, use the EOSC Enhance Portal as one source and add additional resources that will not be onboarded to the EOSC Enhance Portal, such as SSHOC [8] .

The second possible line of investigation for assessing the national uptake of EOSC is an analysis of the concrete use of the EOSC provided resources by the country of the user. As this can only be done for each resource type or resource provider individually, this constitutes a challenge to implement broadly for every resource. It may be possible to focus on the use of central EOSC core services to assess their uptake in a geographical distribution. However, consistent measuring of user activity across a wide variety of services and service providers in multiple countries poses both technical and legal challenges as it effectively constitutes user tracking. Efforts are being undertaken in WP4 to establish a solution for usage measurement that could serve to solve this issue, as long as broad adoption of central services are not required for the measurement, excluding services that do not rely on specific EOSC implementations, yet still provide real valuable end user services.

## 3.2 Interoperability

Independently from the listing of a resource on any given portal, its integration into EOSC is also dependent on its interoperability. This concerns the composability of services with other services in the EOSC ecosystem and the technical compatibility in terms of the standards of data, software and other resources available. This should be evaluated in both its design and readiness as well as actual implementation and concrete use. The points to be considered are:

[14] https://www.eoscsecretariat.eu/, accessed 27 July 2021
[15] Developed and provided by OpenAIRE at https://osobservatory.openaire.eu/home, accessed 27 July 2021

- Use of common and recommended standards (metadata, data formats, protocols),
- Availability of APIs and their documentation,
- Access and usage mechanisms in compliance with EOSC norms, e.g., AAI, usage monitoring.

Thus, building on available indicator data as outlined in Section 0 below, the work to be done in WP2 Activity 2.2.1 regarding objectives of raising the profile and adoption of EOSC is threefold:

- Support national infrastructures to onboard their resources to EOSC in compliance with interoperability standards and Rules of Participation (RoP).
- Support national infrastructures in adopting resources available from the EOSC Enhance Portal or any regional or thematic portal for national use.
- Support connecting national resources with those provided by e-Infrastructures as well as regional and thematic infrastructures through standardised APIs and formats, and ensure this composability enriches the relevant portal listings.

The adoption ultimately has to be measured by the acceptance and consequent use of the individual users of EOSC and the resources provided, in collaboration with WP10.

# 4 Background of EOSC landscaping efforts

The bases for the alignment in this deliverable for assessing EOSC Readiness are the reports 'Landscape of EOSC-Related Infrastructures and Initiatives' [3] and 'Country sheets analysis' [4] from November 2020 by the EOSC Landscape Working Group, as well as the 'Second Working Proposal for Living Indicators to Monitor MS Progresses Towards EOSC Readiness' from January 2021 by the Landscaping Task Force [1] .

While the Landscape Report summarises the existing policies and investments based on inputs provided by the MS/AC[16], the Landscape Analysis goes several steps further. It reflects the rapid development in the EOSC and goes deeper into the country sheet data as well as complements the work being carried out by the INFRAEOSC 5B Projects, which are assessing the landscape at the local level. The country sheets and resulting analysis offer a snapshot of the current state in 2020. In the report, EOSC Readiness is loosely defined as having relevant policies in place or in the planning stage (European Commission - Directorate-General for Research and Innovation, 2020, p. 3) [3] . Using this loose definition, all European countries show a fairly high degree of EOSC Readiness. In the majority of countries, many of the relevant policies are still in the planning stage and there is an opportunity to support better harmonisation across such countries during the early stages of policy development. With regard to available infrastructures that may be federated or made available, for many EU MS/AC, there is a very large number of infrastructures that could potentially be made available. As the emerging EOSC related indicators are further elaborated and agreed upon, the level of detail sought in future iterations of the country sheets, in relation to both policies and infrastructures, will need to be refined to better reflect a more comparable and qualitative approach to EOSC Readiness.

The Landscaping Task Force, consisting of representatives of EOSC regional projects and FAIRsFAIR project in collaboration with EOSCsecretariat.eu, has proposed an initial list of potential EOSC Readiness indicators in EU MS and AC in the 'Second Working Proposal for Living Indicators to Monitor MS Progresses Towards EOSC Readiness' document [1] . An overview of the main indicators which gained more than 80% each of respondents' votes, prioritized according to the highest rankings of respondents is represented in Table 4-1 below.

*Table 4-1: List of proposed indicators by the Landscaping Task Force*

| | |
|---|---|
| **Architecture** | |
| **1** | National initiative in place/planned |
| **2** | National (regional) registry or other federation mechanisms for data in place/planned:<br>• Number of enrolled services |
| **3** | National (regional) dataset catalogue(s) in place/ planned<br>• Number of enrolled datasets<br>• Integration with data catalogues |
| **4** | National PID Policy |
| **5** | Interoperability |
| **Organisation & Governance** | |
| **6** | National initiative in place/planned<br>• Funding<br>• Stakeholders involved |
| **7** | Strategic roadmap |
| **8** | Specific funding programmes |
| **9** | EU Initiatives |
| **10** | Interoperability with trans-national initiatives |
| **Policies** | |
| **11** | OS/FAIR policies supported/ monitored/ planned<br>• National<br>• Incentives |

[16] Some of the reports distinguish also Other Countries, such as the UK, which we include in the understanding of MS/AC.

| | |
|---|---|
| **12** | DM policies in e/supported/ monitored/ planned<br>• National<br>• Organisation level<br>• Incentives |
| **13** | Source of funding |
| **Infrastructure** | |
| **14** | Resources<br>• Infrastructure availability<br>• Availability of certain types of services |
| **15** | Access policies publicly available |
| **Training and Skills** | |
| **16** | National/regional curricula in place/planned (compliance with international?)<br>• Data scientists<br>• Data stewards<br>• University courses & graduates |
| **17** | Basic training available for researchers & research support staff<br>• National competence centres |
| **18** | People trained per year |

# 5 SRIA Indicators

In the document 'Strategic Research and Innovation Agenda (SRIA) of the European Open Science Cloud' Key Performance Indicators (KPI) for the European Partnership (EOSC Executive Board, 2021, p.160) [6] have been proposed. They are listed in Table 5-1 below.

*Table 5-1: Key Performance Indicators defined for the European Partnership by SRIA*

| SRIA Objective | KPI | Target | Measure Direct (D) / Survey (S) |
|---|---|---|---|
| Ensure that Open Science practices and skills are rewarded and taught, becoming the 'new normal' | Percentage of publications from EOSC Association research-performing members that become immediate open access. | 70% by 2023 | D |
| | Number of national education systems that recognise European curricula for data stewardship. | 5 by 2025 | S |
| | Percentage of RPOs that are EOSC Association members that have data stewards to support their research. | 50% by 2025 | S |
| | Percentage of EOSC Association members that recognise Open Science activities in research career assessments. | 50% by 2025 | S |
| | Percentage of research-funding members of the EOSC Association that require data sharing and incentivise reuse. | 70% by 2025 | S |
| Enable the definition of standards, and the development of tools and services, to allow researchers to find, access, reuse and combine results | Percentage of EOSC Association members that have policies which require FAIR to be implemented in project design via Data Management Plans. | 70% by 2023 | D |
| | Percentage of research data from EOSC Association members which is deposited in repositories that is made as open as possible. | 50% by 2025 | D |
| | A first generation of pan-European federation of infrastructures for preservation, management and sharing of research software is available. | 1 by 2025 | S |
| | Percentage of the active data spaces that take up data management practices, including the FAIR data principles, and provide into the EOSC ecosystem. | At least 50% by 2027 | D |
| Establish a sustainable and federated infrastructure enabling open sharing of scientific results | Number of core functions of Minimum Viable EOSC that are developed to make the EOSC ecosystem accessible to researchers across disciplines and countries. | 4 by 2025 | S |
| | Number of additional functionalities and services dedicated to the requirements of end users from the public sector in the EOSC-Core and EOSC-Exchange. | 10 by 2025 | D |
| | Percentage of the repositories in EOSC that will have a certification such as CoreTrustSeal. | 30% by 2025 | D |
| | Percentage of research disciplines that have documented standards and protocols for data sharing and reuse. | 60% by 2023 | S |
| | Percentage of the metadata related to publicly funded research datasets which are | 70% by 2025 | D |

| | | | |
|---|---|---|---|
| | defined as Open Data that are discoverable through EOSC federated infrastructure. | | |
| | Number of geographically spread observer organisations that have joined EOSC from outside EU MS/AC. | At least 10 by 2025 | D |
| | Number of non-EU initiatives with which EOSC establishes connections, offering additional resources to the EOSC ecosystem. | At least 3 by 2027 | D |

# 6 Other factors to be considered

The measurement of national readiness indicators is embedded in the wider EOSC landscape, that includes the other layers of Figure 2.1, as well as the ESFRI infrastructures that follow an independent monitoring approach. This section describes the relevant background.

## 6.1 In-kind contributions by EOSC members

In the document 'Memorandum of Understanding for the Co-programmed European Partnership on the European Open Science Cloud' [5] all the contributions and activities by the Partners have been defined. The European Commission envisages to dedicate up to EUR 490 million to actions within the scope of the European Partnership. Furthermore, the Partners other than the Union envisage to dedicate up to **EUR 500 million** for the period 202[1/2] - 2030 in research, innovation and other activities in the area of the European Partnership (EOSC Association, 2020, p.4) [5]

Contributions by the Partners other than the Union will take the form of:

- **In-kind contributions to the actions funded by the Union**, consisting of eligible costs in accordance with the Horizon Europe rules minus the Union contribution.
- **In-kind contributions to Additional Activities** that are in the scope of the SRIA and identified in the annual Additional Activities Plan, which is approved by the Partnership Board. In-kind Additional Activities are those activities which contribute to achieving the objectives of the European Partnership, including R&I at higher Technology Readiness Level and/or to ensure demonstration, market, regulatory and societal uptake, which are in the scope of the SRIA but are not covered by Union funding (EOSC Association, 2020, p.5) [5]

The scope of In-kind Additional Activities is listed in Table 6-1 below.

*Table 6-1: The scope of In-kind Additional Activities by the Partners*

| Item | Activity |
|---|---|
| 1 | Implementing policy commitments to further the mainstreaming of Open Science practices across the member organisations of the EOSC Association. |
| 2 | Supporting research communities to develop and adopt open science practices and standards. |
| 3 | Adopting standards and tools for preserving, referencing, describing and citing research outputs, including data, software and publications. |
| 4 | Encouraging and incentivising use of European infrastructure for sharing of research software. |
| 5 | Coordinating and aligning relevant curricula on skills for FAIR and Open Science, and training frameworks. |
| 6 | Promoting EOSC at all levels by engaging with relevant communities and stakeholders. |
| 7 | Contributing to a rewards and recognition framework that incentivises FAIR data and Open Science. |
| 8 | Integrating in EOSC recognised domain-specific activities that support Open Science. |
| 9 | Contributing to the definition of models for availability and costing of services across borders. |
| 10 | Providing research communities with resources for data curation. |

The scope of investments in operational activities may include:

- Personnel and infrastructure contributing to the upgrade of existing Research Infrastructures and e-Infrastructures so that they may be federated through EOSC.
- Additional capacity to support the expanding use of research output related to the adoption of Open Science policies.

The continuous monitoring and periodic reporting by the Partners will be carried out at least annually. In addition to that, there will be a simplified reporting over one year and a full reporting every second year. The outcomes of monitoring and reporting from all the in-kind initiatives by the MS/AC should be considered as well while evaluating the EOSC Readiness.

## 6.2 ESFRI

In the EOSC ecosystem, the ESFRI Research Infrastructures - organised in five thematic science clusters - are critical for the provision of data and services to their research communities and integrating these into the emerging EOSC infrastructure. While these infrastructures are by design international, the majority of their funding usually comes from national membership contributions and thus they and their membership structure are relevant for understanding national EOSC readiness.

In the 'Report of the ESFRI Working Group on monitoring of research infrastructures performance' the ESFRI Working Group [7] has defined a set of KPIs to address the most commonly held objectives of pan-European Research Infrastructures. The KPIs are listed in Table 6-2 below:

*Table 6-2: ESFRI numerical KPIs per objective for monitoring of research infrastructures performance*

| Objective | KPIs |
|---|---|
| Enabling scientific excellence | 1. Number of user requests for access<br>2. Number of users served<br>3. Number of publications<br>4. Percentage of top (10%) cited publications |
| Delivery of education and training | 5. Number of master and PhD students using the RI<br>6. Training of people who are not RI staff |
| Enhancing collaboration in Europe | 7. Number of members of the RI from ESFRI countries<br>8. Share of users and publications per ESFRI member country |
| Facilitating economic activities | 9. Share of users associated with industry and publications with industry<br>10. Income from commercial activities and the number of entities paying for service |
| Outreach to the public | 11. Engagement achieved by direct contact<br>12. Outreach through media<br>13. Outreach via the RI's own web and social media |
| Optimising data use | 14. Number of publicly available data sets used externally |
| Provision of scientific advice | 15. Participation by RIs in policy related activities<br>16. Citations in policy related publications |
| Facilitating international cooperation | 17. Share of users and publications per non-ESFRI member country<br>18. International trainees<br>19. Number of members of the RI from non-ESFRI countries |
| Optimising management | 20. Revenues<br>21. Extent of resources made available |

## 6.3 Policies

Monitoring of the implementation of EOSC relevant policies on EU and national levels is expected to be taken up by the EOSC-A, once the recently announced position[17] of Policy Officer is filled.

[17] https://www.eosc.eu/work-with-us/policy-officer-monitoring, accessed 27 July 2021

# 7 Collection of indicator data

With a view towards the EOSC Observatory to be developed in WP2 Task 2.3, the objective is to define statistical indicators that can be collected in a regular and automated fashion, ideally from existing sources.

Acknowledging that not all data is readily available in machine-readable fashion, the aim is therefore to collect measurements for readiness in two ways:

1. Available resources, including services, data, software and publications will be measured by national distribution directly from the EOSC Enhance Portal and the Open Science Observatory and feed into the EOSC Observatory.
2. For indicators not measurable automatically, a common approach with the EOSC-SB will be pursued. As the EOSC-SB plans to execute an annual member survey, that survey should include all relevant indicators for measuring EOSC Readiness. Furthermore, the survey results should be fed directly into or collected through the EOSC Observatory.

## 7.1 Indicators from available resources

The Open Science Observatory has a representation of national and thus regional distribution of open access publications, datasets and software. In addition, the EOSC Enhance Portal contains all the services that have been onboarded to EOSC by the respective service providers. They can also be mapped to the country of the service providing organisation.

These numbers can form the baseline for creating an overview of current adoption of EOSC across Europe. They are, however, not comparable across countries in absolute form and must be seen in relation to the respective country's capacity. Defining a method to compare these numbers across countries will be one of the tasks that needs to be defined in close collaboration with the EOSC-SB.

Furthermore, the number of current resources provided to EOSC does not necessarily present a measure of readiness, but a measure of adoption. One possibility could be to define target adoption rates for each type of resource and calculate the percentage achieved, the *relative adoption rate.* This could be done in a generic form, using a comparative indicator, as explained in the previous paragraph or by defining national targets for each country. This will need agreement by the EOSC-SB and all MS/AC.

## 7.2 Surveyed indicators

In addition to data collected from services operated by EOSC, data will also need to be collected manually from each country. The EOSC-SB currently plans to survey its members on an annual basis to collect the information needed to influence their decision-making process. With multiple organisations and projects working on country surveys, duplication of any form must be avoided, and the surveying should be unified.

Of course, the survey and its questions must be well designed to avoid any ambiguity following the common standards and practices in Social Sciences. Both the data that is expected and how it is to be collected must be defined as explicitly as possible to ensure that responses are comparable across countries and in fact useable.

## 7.3 Data validation

Both data collected from portals and those provided by the MS/AC through the EOSC-SB annual survey must be validated. A manual process to validate the information must be established and tested in a first pilot before applying a potentially time-intensive method to all countries. For this process, the roles and responsibilities of the organisations and individuals involved must be clearly defined and carefully implemented to ensure proper knowledge and balance while having ensured commitment.

## 7.4 Expected limitations and challenges

Resources onboarded to the EOSC Enhance Portal do not necessarily represent a country's full catalogue of available resources. The data from other EOSC Portals such as regional or thematic portals should be included as well. This is dependent on de-duplication of entries present in multiple portals.

Different portals and catalogues have different dataset representation. Merely combining these numbers will lead to highly incorrect data, as duplicates must first be removed. This is only possible by direct comparison of all entries across portals, but still requires unique identifiers for services used consistently throughout.

Services of international institutions such as ESFRI Landmarks, ERICs and e-Infrastructures are often counted towards the country of the seat of the coordination office or the coordinating institution in the case of federated infrastructures. This can be seen as skewing the numbers to show a higher number of services provided by a country even though the services are not provided by national efforts. Conversely, the establishment of the seat of an international infrastructure in a country can in fact be the direct consequence of the country's policy in support of EOSC. This must be investigated carefully.

## 7.5 Open Questions

While the number of resources collected in this way will include those provided by ESFRI Landmarks, ERICs and e-Infrastructures, they will be counted as national contributions, as explained above. Simultaneously, ESFRI monitoring methods and KPIs have been developed by the EC and the contributions of these infrastructures must become visible in the EOSC Observatory in their own right. What data can be collected, from where and how that can be disambiguated from national data is yet to be examined, and is part of the work of WP2 Task 2.3.

Measurement of the SRIA indicators is a central aspect in monitoring the EOSC. Where and how this data is collected and measured has to be clarified and its inclusion in the EOSC Observatory must be considered.

Finally, the in-kind contributions provided by the MS/AC as part of the EOSC Partnership agreement should be considered as a second source of data that can serve to validate the results. This should be done in parallel to the counting of resources onboarded to EOSC and can serve as a two-sided validation approach. Ideally, EOSC services provided by the MS/AC as part of their in-kind contributions should match those provided to the researchers through the portals.

# 8 Conclusion & Upcoming work towards D2.2b

A major part of work allocated to Activity 2.2.1 will be spent in preparation of the final deliverable D2.2b. That report will present the result of the alignment process with the EOSC-A and the EOSC-SB that is expected to start following the publication of this deliverable. The deliverable will also summarise the list of indicators agreed upon and their values at the time.

To this end, the alignment process with the EOSC-A and the EOSC-SB will be pursued with the following three objectives in mind:

1. Define the indicators.
2. Define the collection method for each indicator and support their implementation through the EOSC Observatory.
3. Validate the data collected through these methods and their usefulness as indicators.

This will require the understanding of the requirements for EOSC participation as examined through EOSC Future project WP2 Activity 2.2.4 and must be aligned with the development of the EOSC Observatory in WP2 Task 2.3.

In summary, going forward the activities will be to:

- Liaise with the EOSC-A to define the sources for measuring resource distribution across Europe.
- Liaise with the ESOC-SB and EOSC-A to define the reporting methods for country data not available through APIs from existing EOSC services.
- Define a reporting standard and mechanism for reporting of national indicators, potentially levering mandated members.
- Define the indicator for low uptake and identify MS/AC to support in onboarding resources to EOSC.
- Support the interested countries with low uptake in identifying the best strategy for onboarding the resources offered by its national service providers, setting a particular focus on countries from the Western Balkan.
- Liaise with the EOSC-A and the EOSC-SB to define the list of indicators to be measured and their dissemination level to be collected and presented in the EOSC Observatory.
- Align with WP 10 to ensure the outreach activies and stakeholder engagement are interlinked with the aims of increasing uptake of EOSC in the countries as measured by the indicators.

# References

[1] Arvola, M., Beckmann V., Budroni, P., Castelli, D., Cavalli, V., Davidson, J., Di Giorgio, S., Dovey, M., Fischer, L., Hönegger, L., Kosanovic, B., Lazzeri, E., Matyska, L., Papastamatiou, I., Prnjat, O., Rasmussen, T., Robertson, D., Stoy, L., Tanlongo, F., Toli, E. & Van Wezel, J. (2021) *Second Working Proposal for Living Indicators to Monitor MS Progresses Towards EOSC Readiness*. Available at: https://doi.org/10.5281/zenodo.4452799 (Accessed: 16 July 2021)

[2] Boccali, T., Sølsnes, A. E., Thorley, M., Winkler-Nees, S. & Timmermann M. (2021) *Practical Guide to Sustainable Research Data*. Available at: https://doi.org/10.5281/zenodo.4769703 (Accessed: 16 July 2021)

[3] European Commission - Directorate-General for Research and Innovation (2020) *Landscape of EOSC-related infrastructures and initiatives – Report from the EOSC executive board working group (WG) landscape,* version 2. Available at: https://doi.org/10.2777/132181 (Accessed: 16 July 2021)

[4] European Commission - Directorate-General for Research and Innovation (2020) *Country Sheets Analysis.* Available at: https://doi.org/10.2777/568900 (Accessed: 16 July 2021)

[5] EOSC Association (2020) *The Memorandum of Understanding for the Co-programmed European Partnership on the European Open Science Cloud.* Available at: https://eosc.eu/sites/default/files/20210215_EOSC_MoU_FinalDraft.pdf (Accessed: 21 July 2021)

[6] EOSC Executive Board (2021) *Strategic Research and Innovation Agenda (SRIA) of the European Open Science Cloud*. Available at: https://www.eosc.eu/sites/default/files/EOSC-SRIA-V1.0_15Feb2021.pdf (Accessed: 06 July 2021)

[7] ESFRI Working Group (2019) *Report on Monitoring of Research Infrastructures Performance*. Available at: https://www.esfri.eu/sites/default/files/ESFRI_WG_Monitoring_Report.pdf (Accessed: 23 July 2021)

[8] Thiel, C., (2020) *SSHOC Services for the EOSC Portal*, presentation at the EOSC Symposium. Available at: https://eosc-portal.eu/news/introducing-eosc-portal-onboarding-procedure-realising-eosc-conference, minutes 8-20, (Accessed: 27 July 2021)