DOI 10.5281/zenodo.10475460

# Deliverable 2.2 Open Science Assessment Framework

| | |
|---|---|
| Due Date of Deliverable | 31/12/2023 |
| Actual Submission Date | 27/12/2023 |
| Work Package | WP2 |
| Tasks | T2.2 |
| Type | R - Report |
| Approval Status | Submitted |
| Version | 2.0 |
| Number of Pages | 44 |



# Abstract

This deliverable outlines the concept and development of the Open Science Assessment Framework (OSAF) in the GraspOS project. OSAF aims to provide an assessment framework that reflects on OS-aware responsible assessment and OS research practices. The deliverable is an introduction to the OSAF concept and related co-development implementation factors.

This project has received funding from the European Union's Horizon Europe framework programme under grant agreement No. 101095129. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the European Research Executive Agency can be held responsible for them.

# Revision History

| Version | Date | Reason | Revised by |
|---|---|---|---|
| 0.0 | 13.10.2023 | First draft | Clifford Tatum, Zeynep Anli |
| 0.1 | 13.10.2023 | Agreement on structure & References | Clifford Tatum, Zeynep Anli |
| 0.2 | 13.10.2023 | Intermediate version | Clifford Tatum, Zeynep Anli, Anna-Kaisa Hyrkkänen, Janne Pölönen, Josefine Nordling |
| 0.3 | 13.10.2023 | Peer review | Lottie Provost, Suzanne Dumouchel, Fotis Mystakopoulos, Thanasis Vergoulis, Dragan Ivanović |
| 0.4 | 20.10.2023 | Peer review comments addressed | Clifford Tatum, Zeynep Anli |
| 1.0 | 01.12.2023 | Final version after proofreading | Clifford Tatum, Zeynep Anli |
| 2.0 | 08.12.2023 | Second draft | Clifford Tatum, Zeynep Anli, Ludo Waltman, Anna-Kaisa Hyrkkänen, Janne Pölönen, |
| 2.1 | 15.12.2023 | Peer review | Laura Himanen, Andrea Mannocci, Lottie Provost |
| 2.2 | 22.12.2023 | Peer review comments addressed | Clifford Tatum, Zeynep Anli |
| 2.3 | 22.12.2023 | Final version after proofreading | Clifford Tatum, Zeynep Anli |

# Author List

| Organization | Name | Contact Information |
|---|---|---|
| CWTS | Clifford Tatum | c.c.tatum@cwts.leidenuniv.nl |
| CWTS | Zeynep Anli | z.anli@cwts.leidenuniv.nl |
| CWTS | Ludo Waltman | waltmanlr@cwts.leidenuniv.nl |
| TSV | Anna-Kaisa Hyrkkänen | anna-kaisa.hyrkkanen@tsv.fi |
| TSV | Janne Pölönen | janne.polonen@tsv.fi |
| CSC | Josefine Nordling | josefine.nordling@csc.fi |

# Contributor List

| Organization | Name | Contact Information |
|---|---|---|
| CSC | Laura Himanen | laura.himanen@csc.fi |
| CNR | Andrea Mannocci | andrea.mannocci@isti.cnr.it |
| CWTS | Clara Calero-Medina | clara@cwts.leidenuniv.nl |
| CNR | Lottie Provost | lottiemiaprovost@cnr.it |
| OPERAS | Suzanne Dumouchel | suzanne.dumouchel@operas-eu.org |
| OPERAS | Fotis Mystakopoulos | fotis.mystakopoulos@operas-eu.org |
| ARC | Thanasis Vergoulis | vergoulis@athenarc.gr |
| TSV | Dragan Ivanović | dragan.ivanovic@uns.ac.rs |
| TSV | Janne Pölönen | janne.polonen@tsv.fi |

# Table of Contents

# List of Tables



# List of Figures

# Abbreviation List

**ALLEA** - All European Academies (The European Federation of Academies of Sciences and Humanities)
**API** - Application Programming Interface
**ARRA** - Agreement on Reforming Research Assessment
**EOSC** - European Open Science Cloud
**CoARA** - Coalition for Advancing Research Assessment
**FAIR** - Findability, Accessibility, Interoperability, and Reusability
**FAIRCORE4EOSC** - FAIR Convergence for EOSC
**GraspOS** - next Generation Research Assessment to Promote Open Science
**JIF** - Journal Impact Factor
**INORMS -** International Network of Research Management Societies
**ISO** - International Organization for Standardization
**ORCID** - Open Researcher and Contributor ID
**OS** - Open Science
**OSAF** - Open Science Assessment Framework
**PID** - Persistent Identifier
**RAiD** - Research Activity Identifier
**RoR** - Research Organization Registry
**RRA** - Responsible Research Assessment
**SCOPE** - Start with what you value, Context considerations, Options for Evaluating, Probe Deeply, Evaluate for Evaluation
**SCOPE+i** - SCOPE plus infrastructure
**SEP** - Strategy Evaluation Protocol
**SURF** - Samenwerkende Universitaire Reken Faciliteiten (Collaborative University Computing Facilities)
**WG** - Working Group
**WP** - Work Package

# Glossary

**Agreement on Reforming Research Assessment (ARRA):** The agreement establishes a common trajectory for transforming assessment practices within the research community, encompassing researchers and research-performing organizations. The overarching objective is to enhance the quality and impact of research. This agreement outlines key principles, commitments, and a timeframe for implementing reforms. It also sets forth the foundational principles for a coalition of organizations committed to collaborating in the execution of these changes.[1]

**Analytics Infrastructure:** This refers to the comprehensive collection of tools, technologies, processes, services, and resources utilized by an organization to gather, process, analyze, and visualize data with the purpose of making informed business decisions. This infrastructure is specifically designed to facilitate the extraction of insights, identification of patterns, and recognition of trends from extensive datasets.[2]

**Application Programming Interface (API):** It is a set of rules and protocols that allows different software applications to communicate with each other. It defines the methods and data formats that applications can use to request and exchange information. APIs enable developers to access the functionality or data of a software application, service, or platform without needing to understand its internal workings. They serve as intermediaries, allowing applications to interact and share data seamlessly. APIs are crucial for building integrations, enabling interoperability between different software systems, and fostering the development of third-party applications that can leverage the features of a given platform.[3]

**Assessment Protocol:** This is the framework in which the assessment is conducted.

**Assessment Event:** This is the actual assessment.

**Assessment Infrastructure:** This concept includes all the assessment items (such as portfolio and registry, see below) that make up the assessment process.

**Assessment Portfolio:** These are Responsible Research Assessment (RRA) templates which are specifically crafted to serve as purpose-built frameworks for the systematic collection and organization of both quantitative and qualitative indicators. They are designed to be adaptable and suitable for various assessment needs, ensuring a comprehensive approach to capturing and structuring diverse types of data.

**Assessment Registry:** It enables the publication of an assessment protocol after the completion of an assessment event. “This refers to an online database of OSAF-based Assessment Portfolios and

[1] See https://coara.eu/agreement/the-agreement-full-text/
[2] See https://www.sciencedirect.com/topics/computer-science/analytics-infrastructure
[3] See https://en.wikipedia.org/wiki/API

case studies in a structured and systematic way to promote experience sharing and mutual learning."[4]

**Coalition for Advancing Research Assessment (CoARA):** This coalition unites a diverse array of entities engaged in research assessment and their affiliated associations. This includes research funding organizations, research-performing organizations, national/regional assessment authorities and agencies, learned societies, and researcher organizations. The shared objective is to collaboratively drive systemic reform, guided by the common principles and commitments outlined in the Agreement.[5]

**Community-led approaches:** "Community-led curation refers to the process of managing and organizing information or data by a community of individuals, rather than by a single organization or institution. Community-led curation enables a group of people with a shared interest to collectively curate and validate information, making it more accurate, comprehensive, and accessible. Whereas community-led annotation in this report refers to the process of adding additional information or metadata to existing data or information by members of a community. Community-led annotation can enhance the value and understanding of the information by providing additional context, clarifying meaning, or linking related data."[6]

**CRIS:** Current Research Information System. Also referred to as Research Information System (RIM). A current research information system (CRIS) is typically a database used to store, manage and exchange research information (metadata for the research activity and outputs).[7]

**CrossRef:** This is an organization that provides Digital Object Identifiers (DOIs) for scholarly content. A Digital Object Identifier is a unique alphanumeric string assigned to a document (such as an academic paper, journal article, or book) to provide a permanent link to it, making it easy to locate and access online. CrossRef's primary function is to facilitate the identification and linking of scholarly content on the internet. CrossRef plays a critical role in supporting the infrastructure of scholarly communication by providing a standardized way to identify and link academic publications across various publishers and platforms. Researchers, publishers, and institutions widely use CrossRef services to enhance the accessibility and connectivity of scholarly information.[8]

[4] See Anna-Kaisa Hyrkkänen, Dragan Ivanović, Janne Pölönen, Marita Kari, & Elina Pylvänäinen. (2023). GraspOS Deliverable D2.1 "OS-aware RRA approaches landscape report" (1.0). Zenodo. https://doi.org/10.5281/zenodo.8301792

[5] See https://research-and-innovation.ec.europa.eu/news/all-research-and-innovation-news/coalition-advancing-research-assessment-coara-now-launched-2022-12-02_en#:~:text=The%20CoARA%20brings%20together%20a,researcher%20organisations%2C%20all%20willing%20to

[6] See Anna-Kaisa Hyrkkänen, Dragan Ivanović, Janne Pölönen, Marita Kari, & Elina Pylvänäinen. (2023). GraspOS Deliverable D2.1 "OS-aware RRA approaches landscape report" (1.0). Zenodo. https://doi.org/10.5281/zenodo.8301792

[7] Wikipedia entry for CRIS: https://en.wikipedia.org/wiki/Current_research_information_system

[8] See https://www.crossref.org/about

**DataCite:** This is an international non-profit organization that provides DOIs for research datasets. Similar to how CrossRef assigns DOIs to scholarly articles, DataCite's primary mission is to offer a standardized way to uniquely identify and cite datasets. DOIs assigned by DataCite serve as persistent links to ensure the long-term accessibility and citability of research data.[9]

**Dutch Strategy Evaluation Protocol (SEP):** The primary objective of an SEP evaluation is to assess a research unit in accordance with its own objectives and strategic direction. An independent assessment committee, comprised of experts, evaluates the unit's performance based on both the self-evaluation provided by the unit and a subsequent site visit. The overarching aim of the SEP is to uphold and enhance the quality and societal relevance of research while fostering ongoing discussions about research quality, societal significance, and sustainability within the framework of research quality assurance. To achieve this, the research unit is evaluated in the context of its own goals and strategy.[10]

**European Open Science Cloud (EOSC):** The pan-European project is devised to establish a virtual environment facilitating the sharing and access of research data across borders and scientific disciplines. At the core of this initiative is the EOSC Portal, serving as the primary gateway. It offers a unified access point to a diverse array of research resources and services, streamlining the process for researchers to navigate and leverage the available wealth of information.[11]

**FAIR:** "In 2016, the 'FAIR Guiding Principles for scientific data management and stewardship' were published in Scientific Data. The authors intended to provide guidelines to improve the Findability, Accessibility, Interoperability, and Reuse of digital assets. The principles emphasize machine-actionability (i.e., the capacity of computational systems to find, access, interoperate, and reuse data with none or minimal human intervention) because humans increasingly rely on computational support to deal with data as a result of the increase in volume, complexity, and creation speed of data."[12]

**FAIRCORE4EOSC:** The FAIRCORE4EOSC project is dedicated to advancing the European Open Science Cloud (EOSC) by developing and implementing essential components. Its primary objectives include supporting the creation of a FAIR (Findable, Accessible, Interoperable, and Reusable) EOSC and addressing identified gaps outlined in the Strategic Research and Innovation Agenda (SRIA). By building on existing technologies and services, the project aims to create nine new EOSC-Core components. These components are designed to enhance the discoverability and interoperability of a broader range of research outputs within the EOSC framework.[13]

[9] See https://datacite.org/what-we-do
[10] See https://storage.knaw.nl/2022-06/SEP_2021-2027.pdf
[11] See https://eosc-portal.eu/about
[12] See https://www.go-fair.org/fair-principles
[13] See https://faircore4eosc.eu

**Framework:** A framework is a basic structure underlying a system, concept, or text (Oxford dictionary). In software development, a framework is a set of pre-established and reusable components, libraries, and tools organized in a specific structure. It provides a foundation for developers to build applications with standardized practices, reducing the need to recreate common functionalities from scratch. Frameworks are often designed to provide a common structure, enhance efficiency, and ensure consistency in different applications or processes. They offer a systematic way to approach complex tasks, enabling easier development, implementation, or analysis within a given domain.[14]

**h-index:** The h-index, also known as the Hirsch index, serves as a metric to gauge the productivity and impact of a researcher's scholarly publications. Physicist Jorge E. Hirsch introduced this metric in 2005, aiming to provide a numerical assessment that considers both the quantity (number of publications) and impact (citation counts) of a researcher's work. Widely employed in academia, the h-index offers a quick evaluation of a researcher's overall influence and productivity within the academic community. However, it is crucial to recognize that the h-index is inherently biased and has severe limitations, and should not be used in the evaluation of a researcher.[15]

**Journal Impact Factor (JIF):** The Impact Factor (IF) or Journal Impact Factor (JIF) is a scientometric index calculated by Clarivate, reflecting the average number of citations received by articles published in a particular journal over the last two years, as indexed by Clarivate's Web of Science. Functioning as a journal-level metric, the Impact Factor is often employed as an indicator of the relative significance of a journal within its field. Journals with higher Impact Factor values are generally perceived as more important or prestigious within their respective disciplines compared to those with lower values.[16]

**Monitoring of Open Science and research:** "Monitoring generates data on an intervention's activity and impact over time in a continuous and systematic way. It helps identify and address any implementation problems of an intervention at the same time as it generates factual data for future evaluation and impact assessment. (European Commission 2015). UNESCO recommends that "Member States should, according to their specific conditions, governing structures and constitutional provisions, monitor policies and mechanisms related to Open Science using a combination of quantitative and qualitative approaches, as appropriate" (UNESCO 2021)."[17]

**Open Access publications:** Open Access is a publishing model for scholarly communication that provides unrestricted access to research information for readers at no cost. This is in contrast to the traditional subscription model, where readers typically gain access to scholarly content by paying a

[14] See https://en.wikipedia.org/wiki/Software_framework
[15] See https://en.wikipedia.org/wiki/H-index
[16] See https://en.wikipedia.org/wiki/Impact_factor
[17] See Anna-Kaisa Hyrkkänen, Dragan Ivanović, Janne Pölönen, Marita Kari, & Elina Pylvänäinen. (2023). GraspOS Deliverable D2.1 "OS-aware RRA approaches landscape report" (1.0). Zenodo. https://doi.org/10.5281/zenodo.8301792

subscription fee, often facilitated through libraries or other institutions. The aim of open access is to remove financial barriers, making research findings freely accessible to a global audience, thereby fostering widespread dissemination of knowledge and encouraging collaboration among researchers.[18]

**Open Science:** "The United Nations Educational, Scientific and Cultural Organization (UNESCO) defines Open Science as "an inclusive construct that combines various movements and practices aiming to make multilingual scientific knowledge openly available, accessible and reusable for everyone, to increase scientific collaborations and sharing of information for the benefits of science and society, and to open the processes of scientific knowledge creation, evaluation and communication to societal actors beyond the traditional scientific community. It comprises all scientific disciplines and aspects of scholarly practices, including basic and applied sciences, natural and social sciences and the humanities, and it builds on the following key pillars: open scientific knowledge, Open Science infrastructures, science communication, open engagement of societal actors and open dialogue with other knowledge systems." (UNESCO 2021.)"[19]

**Open Science Assessment Framework (OSAF):** The Open Science Assessment Framework (OSAF), developed in the GraspOS project, has three elements: the SCOPE+i method (SCOPE plus infrastructure) to help guide the use of SCOPE toward Responsible Research Assessment protocols and to implement the use of assessment-specific infrastructure in the SCOPE process, thereby extending SCOPE; digital Assessment Portfolios to facilitate collecting and sharing of diverse contributions to be included in an assessment event; and an Assessment Registry for publishing the assessment protocol from completed assessment events.

**Open Science aware Responsible Research Assessment (OS-aware RRA):** Responsible Research Assessment (RRA) that takes into account the Open Science paradigm, thus evaluating research practices in a manner that also aligns with the principles of Open Science. This approach emphasizes transparency, collaboration, and accessibility in research, as. In the context of RRA, it means assessing not only the traditional scholarly outputs but also considering practices such as open access, data sharing, and collaborative efforts. The goal is to promote research that adheres to Open Science principles, fostering a more inclusive and impactful research environment.

**Open Researcher and Contributor ID (ORCID):** ORCID, an acronym for Open Researcher and Contributor ID, is a global, non-profit organization that sustains itself through fees collected from its member organizations. It operates as a community-driven initiative with governance provided by a Board of Directors representing a diverse range of stakeholders. ORCID's structure is designed to ensure broad representation and involvement from its membership. The organization is further supported by a dedicated and knowledgeable professional staff, working collaboratively to advance

[18] See https://www.openaccess.nl/en/what-is-open-access

[19] See Anna-Kaisa Hyrkkänen, Dragan Ivanović, Janne Pölönen, Marita Kari, & Elina Pylvänäinen. (2023). GraspOS Deliverable D2.1 "OS-aware RRA approaches landscape report" (1.0). Zenodo. https://doi.org/10.5281/zenodo.8301792

the mission and objectives of ORCID in facilitating unique and persistent identifiers for researchers and contributors in the scholarly community.[20]

**Persistent Identifier (PID):** A persistent identifier is a long-lasting reference to a digital resource.[21]

**Research Activity Identifier (RAiD):** "A Research Activity Identifier (RAiD) is a globally unique, persistent identifier (PID) for research projects and activities. It comprises both a RAiD name containing the unique persistent identifier '10.25' (called a 'DOI RAiD handle'), and a RAiD metadata record. A RAiD links a project with its non-sensitive metadata information (such as contributors, organizations, grants, instruments, publications and datasets), without linking this information between each other or duplicating information that can be found elsewhere."[22]

**Responsible Research Assessment (RRA):** Responsible research evaluation centers around generating research metrics that align with specific principles, including ensuring data accuracy, transparent data collection and analysis, and the utilization of a diverse range of indicators.

**Research Organization Registry (ROR):** The Research Organization Registry (ROR) is a global initiative, led by the community, that serves as a registry for open and persistent identifiers assigned to research organizations. ROR plays a vital role in facilitating the unambiguous identification of institution names, enabling seamless connections between research organizations, researchers, and research outputs. This registry is utilized across various systems in journal publishing, data repositories, funder and grant management platforms, open access workflows, and other components of research infrastructure. Its primary functions include disambiguating institutional affiliations, enhancing the discovery and tracking of research outputs based on affiliations, and supporting open access publishing workflows, among other important use cases.[23]

**SCOPE:** (Start with what you value, Context considerations, Options for Evaluating, Probe Deeply, Evaluate for Evaluation) – "The SCOPE framework for research evaluation is a five-stage model for evaluating responsibly. It is a practical step-by-step process designed to help research managers, or anyone involved in conducting research evaluations, in planning new evaluations as well as check existing evaluations. SCOPE is an acronym, where S stands for START with what you value, C for CONTEXT considerations, O for OPTIONS for evaluating, P for PROBE deeply, and E for EVALUATE your evaluation."[24]

**Software Infrastructure:** Infrastructure refers to the fundamental software components, tools, frameworks, and resources that deliver crucial support and services for the entire lifecycle of software applications, including development, deployment, and operation. This infrastructure establishes the underlying structure essential for the smooth functioning of software systems and

[20] See https://info.orcid.org/what-is-orcid
[21] See https://en.wikipedia.org/wiki/Persistent_identifier
[22] See https://raid.org/overview
[23] See https://ror.org/about
[24] See https://inorms.net/scope-framework-for-research-evaluation

applications. Software infrastructure encompasses a broad spectrum of elements that collectively contribute to the overall software ecosystem, ensuring the robustness and efficiency of software development and deployment processes.[25]



[25] See https://en.wikipedia.org/wiki/Federated_architecture

# Executive Summary

The Open Science Assessment Framework is an initiative for research assessment infrastructure by integrating principles from both the Open Science (OS) and Responsible Research Assessment (RRA) movements. Aligned with the Coalition for Advancing Research Assessment (CoARA) agreement and the SCOPE framework, OSAF seeks to support the commitments to transforming research evaluation practices.

OSAF consists of three essential elements: the SCOPE+i Method, Assessment Portfolios, and the Assessment Registry. The SCOPE+i Method enhances the SCOPE approach by incorporating contextual factors, guiding the use of SCOPE toward RRA protocols and integrating assessment-specific infrastructure. Assessment Portfolios play a pivotal role in collecting and sharing diverse contributions for assessment. The Assessment Registry serves as a repository for publishing assessment protocols, providing transparency and contributing to a collective understanding of evaluation practices.

Emphasizing Responsible Research Assessment (RRA), the OSAF framework prioritizes responsible and inclusive evaluation practices. In the early release, the focus is on guiding and enabling the assessment of OS contributions within the broader context of RRA. The next development phase outlines key priorities, including the continued development of assessment guidelines, mapping contributions to OS-aware RRA, and integrating OSAF into the GraspOS federated infrastructure.

The development phase seeks to refine the OSAF concept by leveraging insights from pilot analyses, advancing assessment guidelines and resources and mapping OS contributions using the Research Activity Identifier (RAiD) metadata schema. The proposed integration meeting format aims to align end-to-end requirements across different work packages, ensuring integration and interoperability. Overall, these priorities signify a strategic and collaborative approach to advancing OSAF, incorporating lessons learned, and ensuring compatibility with existing standards while contributing to the evolution of responsible and inclusive research evaluation practices.

# 1. Introduction

The Open Science Assessment Framework (OSAF) is being developed to facilitate the use of research *assessment -specific* infrastructure informed by sensibilities of both the Open Science (OS) and Responsible Research Assessment (RRA) movements. This effort is guided by principles articulated in the Coalition for Advancing Research Assessment (CoARA)[26] agreement and by the SCOPE framework.[27] Furthermore, we identify Agreement on Reforming Research Assessment[28] (ARRA) signatories as key beneficiaries for project outcomes.

The aim here is to facilitate signatories' efforts in complying with the ARRA commitments toward reforming research evaluation practices. To facilitate this, the GraspOS pilots serve as co-production partners in developing the OSAF. Operationally, GraspOS has adopted the SCOPE evaluation framework as the common approach for piloting Open Science aware responsible research assessments. As such, each of the nine pilot assessments follows the SCOPE process in their respective projects.

The CoARA agreement, or ARRA, articulates four priority commitments:

1. *Recognize the diversity of contributions to, and careers in, research in accordance with the needs and nature of the research.*
2. *Base research assessment primarily on qualitative evaluation for which peer review is central, supported by responsible use of quantitative indicators.*
3. *Abandon inappropriate uses in research assessment of journal- and publication-based metrics, in particular inappropriate uses of Journal Impact Factor (JIF) and h-index.*
4. *Avoid the use of rankings of research organizations in research assessment.*

The SCOPE framework, developed by the INORMS Research Evaluation Group[29] aligns well with the ARRA and importantly, its use has supported a move towards implementing RRA. The SCOPE framework is guided by three main principles:

1. *Evaluate only where necessary. Evaluation is not always the right strategy. When it comes to incentivizing behaviors, for example, it may be more fruitful to enable them than to evaluate them.*
2. *Evaluate with the evaluated. Any evaluation should be co-designed and co-interpreted by the communities being evaluated.*
3. *Draw on evaluation expertise. We should apply the same rigor to our evaluations that we apply to our academic research.*

These CoARA and SCOPE principles highlight three key objectives for the OSAF: First, accommodating the **diversity of contributions and roles**. Second, facilitating collaboration in development of the assessment protocol, especially together **with those who are being**

[26] CoARA website https://coara.eu
[27] SCOPE framework https://inorms.net/scope-framework-for-research-evaluation
[28] ARRA https://coara.eu/app/uploads/2022/09/2022_07_19_rra_agreement_final.pdf
[29] INORMS Research Evaluation Group https://inorms.net/research-evaluation-group

**evaluated**. And third, facilitating **contextual factors** related to research assessment, which are also relevant to both the CoARA and SCOPE.

# 2. Open Science Assessment Framework

The Open Science Assessment Framework (OSAF) has three elements: the **SCOPE+i method** (SCOPE plus infrastructure) to help guide the use of SCOPE toward Responsible Research Assessment protocols and to implement the use of assessment-specific infrastructure in the SCOPE process, thereby extending SCOPE; digital **Assessment Portfolios** to facilitate collecting and sharing of diverse contributions to be included in an assessment event; and an **Assessment Registry** for publishing the assessment protocol from completed assessment events.

The OSAF framework is focused on enabling Responsible Assessment (RRA), as it forms the basis of assessing Open Science. From a development sequence perspective, RRA is the larger, overarching concept, so in this early release of the OSAF we prioritized RRA to guide and enable assessment of Open Science contributions.

The three OSAF elements are described in more detail below. In this section, we first discuss the SCOPE approach with further elaboration of contextual factors, then bring these together in an outline of the SCOPE+i method. In Section 3, we develop the Assessment Portfolio concept, and its role in collecting assessment content as a machine-readable evidence package for distribution among assessment stakeholders and for use in downstream assessment analytics. While not all evidence will be suitable for downstream computational analysis, the intent is to co-locate the broad diversity of evidential formats along with relevant assessment documents, e.g., value statement, context, etc. We also introduce the Research Activity Identifier (RAiD), which provides the technological underpinning for both the Assessment Portfolios and Assessment Registry. Section 4 introduces assessment resources, which are key elements of the SCOPE+i method. Section 5 concludes with a brief account of the next steps of OSAF development.

## Adopting SCOPE

Our approach to research assessment begins with the premise that context, purpose and values inform assessment protocol, and that each context is different. And that research assessment should be oriented toward mutual learning valued by both researchers and institutions. The SCOPE+i (SCOPE plus infrastructure) method builds upon the existing SCOPE framework by integrating state-of-the-art open infrastructures into the assessment process. In this section, we first outline the original SCOPE framework, we then elaborate on the context dimension of this model, and finally we extend the SCOPE model by introducing the SCOPE+i method.

## SCOPE

The SCOPE Framework, a method that enables designing and conducting research assessment, provides a holistic and participatory approach to research evaluation. The model is based on its five key stages:

START with what you value emphasizes beginning the evaluation process by identifying and articulating the core values pertinent to the research or entity being evaluated. It involves a comprehensive understanding of what is intrinsically important to the stakeholders involved,

ensuring that the evaluation aligns with values rather than relying on the available indicators and data sources.

CONTEXT considerations address the specific context in which the evaluation is being conducted. It includes understanding the organizational setting (e.g., size, location, discipline), the reasons for the evaluation, and the potential effects of the evaluation. Identifying contextual factors ensures the evaluation is appropriate for the local circumstances, thus avoiding a one-size-fits-all approach.

OPTIONS for evaluating explores both quantitative and qualitative methods for evaluation. This stage encourages evaluators to think broadly about the tools and methods available, to ensure a good fit for the approach to evaluation.

PROBE deeply involves a critical examination of the chosen evaluation approach. This effort includes assessing the potential for creating or perpetuating inequalities, gaming the assessment criteria, and other unintended consequences, as well as the cost-benefit of conducting the evaluation. This proactive probing is intended to identify and mitigate any negative impacts of the evaluation process.

EVALUATE your evaluation is a reflective process and the final stage of the assessment, where the effectiveness and impact of the evaluation itself are assessed. This involves reviewing whether the evaluation met its aims, was sufficiently formative and/or summative, and opportunities for improvement in future evaluations.

Before moving on, we build on the context dimension in the SCOPE model.

### Elaboration on Context

For contextual factors we draw on an expert report, Indicator Frameworks for Fostering Open Knowledge, commissioned by the European Commission.[30] This report elaborates further on multiple layers of context that together enable a more granular account of local practices. Following is a summary of context in relation to research assessment, which also includes open science as a dimension of context.

Disciplinary Variations: Different academic disciplines have unique research cultures, publication norms, and impact measures. Contextual factors like the field of study, prevalent methodologies, and the nature of scholarly communication within that field are crucial for fair and relevant evaluation.

Institutional Settings: The type of institution (e.g., research-intensive university, liberal arts college, industry research lab) can influence the goals, resources, and expectations for research. Evaluations need to account for these institutional differences to avoid one-size-fits-all assessments.

Geographical and Cultural Context: Research impact and relevance can vary greatly across different geographical and cultural contexts. What is considered significant or innovative in one region or culture might not hold the same value in another.

[30] Indicator frameworks for fostering open knowledge practices in science and scholarship: https://data.europa.eu/doi/10.2777/445286

Societal Relevance: The societal needs and challenges of a particular time and place are important contextual factors. Evaluations should consider how research addresses local or global societal issues, which is especially pertinent in fields like public health, environmental science, and social policy.

Economic and Political Climate: The broader economic and political environment can influence research priorities, funding availability, and the feasibility of certain types of research. This context is crucial for understanding the constraints and opportunities researchers face.

Technological Advancements: The state of technology and its availability can greatly impact research methods, dissemination, and impact. Evaluations should consider the technological context in which research is conducted.

Open Science and Accessibility: In the context of Open Science, evaluations might consider how research contributes to making scientific knowledge more accessible and reusable. This includes the use of Open Access publications, open data, and open-source tools.

Research Ethics and Integrity: The ethical standards and practices prevalent in a given research context are critical. Evaluations should consider how researchers adhere to ethical guidelines and contribute to the integrity of their field.

Incorporating contextual factors in research assessment not only enables fair and equitable approaches but also ensures that assessments are relevant and meaningful within the specific environment in which research is conducted. Context provides a basis for selecting an assessment approach. For example, a learning evaluation would suggest a formative (or developmental) approach. The evaluation would be more closely linked to context, such as competencies and the details of doing research, and would call for more qualitative input. Whereas an assessment to inform resource allocation might suggest a summative approach, for example when comparing research institutes or groups.

## From SCOPE to SCOPE+i

The SCOPE method provides a comprehensive, value-driven, and inclusive approach to research evaluation. It encourages a participatory process, ensuring that evaluations are context-sensitive, nuanced, and reflective, aiming to enhance rather than hinder the research ecosystem.

The success of the SCOPE method relies in part on its simplicity. It was developed as a high-level framework to ensure its utility across a wide range of contexts. Our aim is to extend SCOPE by providing systematic guidance on the infrastructures needed in the implementation of the various stages of research assessment. Here, we introduce the SCOPE+i method. In addition to the key principles outlined above, the SCOPE+i method pays special attention to the openness of infrastructures and data sources to ensure assessments can be performed in transparent and responsible ways.

The following table summarizes the SCOPE+i method and assessment infrastructures within the different phases of an assessment event. In the left column, the assessment event is divided into four phases: 1) assessment readiness, 2) assessment protocol, 3) assessment execution, and 4) assessment evaluation & dissemination. We then index the SCOPE process (column 2) and OSAF

elements (columns 3 & 4) to these phases according to their relevance in the respective phase. The aim is to provide an overview of how SCOPE and OSAF work together.

### Table 1 Open Science Assessment Framework (OSAF)

| | Open Science Assessment Framework (OSAF) | | |
|---|---|---|---|
| Assessment event phases | Scope | SCOPE+i Method | Assessment Infrastructure |
| Assessment readiness | **1-Start with what you value**<br>**2-Context & purpose** | - OS assessment guidelines<br>- assessment team guidelines<br>- template, assessment readiness<br>- template, stakeholder mapping<br>- template, value statement<br>- template, purpose statement<br>- template, contextual factors | **Assessment Portfolio**<br>- assessment team<br>- readiness report<br>- stakeholder map<br>- value statement<br>- purpose statement<br>- relevant contextual factors |
| Assessment design | **3-Options for evaluation**<br>**4-Probe deeply** | - translating values, purpose and context into an assessment protocol<br>- narrative template<br>- strategy template<br>- evaluator/evaluand guide<br>- RRA obstacles guide<br>- diversity of OS contributions guide<br>- equity, diversity, inclusion guide<br>- responsible assessment checklist<br>- assessment protocol guide/template<br>- indicator toolbox guidelines<br>- open research information sources<br>- GraspOS services catalog | **Assessment Portfolio**<br>- collaborative evidence selection<br>- evaluand(s) narrative<br>- indicators and data sources<br>- assessment protocol document |
| Assessment execution | | | **Assessment Portfolio**<br>- distribute portfolio to stakeholders |
| Assessment evaluation & dissemination | **5-Evaluate the evaluation** | **-** Evaluate the evaluation guidelines | **Assessment Registry**<br>- assessment team<br>- readiness report<br>- stakeholder map<br>- value statement |

| | | | |
|---|---|---|---|
| | | | - purpose statement<br>- relevant contextual factors<br>- assessment protocol |

The assessment infrastructures in the first four stages of the SCOPE+i method together constitute a so-called Assessment Portfolio. The Assessment Portfolio and Assessment Registry are outlined below and discussed further in Section 3. The assessment resources that comprise the method are elaborated in Section 4.

### Outline: Assessment Portfolio

Assessment Portfolios offer an assessment infrastructure that brings together the key information in the first four stages of the SCOPE+i method, providing an account of the evidence to be assessed and a shared digital resource for conducting the assessment. As shown in the above table, this information for instance includes value statements, purpose statements, and contextual factors. Below, we discuss the key elements of an Assessment Portfolio.

### Outline: Assessment Registry

The Assessment Registry facilitates the publication of an assessment protocol after the completion of an assessment event. Registration of assessment protocols facilitates transparency and mutual learning. An assessment protocol would include a description of the assessment, contextual factors, data sources, and indicators (and how they were calculated). Not included are individual identities and the specific evidence used. The collection of registered assessment protocols will provide a searchable resource for others looking for inspiration in designing assessment approaches for OS contributions and/or RRA more broadly.

# 3. Implementation

In this section we delve deeper into the OSAF elements, and we provide further background on the adopted technology (RAiD) underpinning both the Assessment Portfolio and Assessment Registry.

## SCOPE+i Method

SCOPE will be the backbone of the OSAF method, upon which we incorporate assessment infrastructure as a bridge to conducting assessment analytics. This SCOPE+i (SCOPE plus infrastructure) method provides practical resources (guidelines, templates and checklists) derived from the D2.1 landscape analysis.[31] The aim of these resources is twofold. First, to enable OS-aware RRA planning and design, and second, to facilitate the digital collection of assessment-related documents and diversity of evidence formats. These planned resources are individually described in Section 4, Assessment Resources.

## Assessment Portfolio

Assessment Portfolios facilitate the collection of inputs for research assessment, serving both as an account of the agreed evidence for a given assessment event and as a shared resource for conducting the assessment. In a general sense, this means building a collection of contributions by transferring object records from research databases (e.g., CRISs and/or ORCID records) into a unique portfolio allocated for each assessment event.

Incorporating Assessment Portfolios into the assessment protocol and execution phases brings together key information about the assessment, e.g., values, contextual factors, and purpose, with the decision-making tasks of expanding the diversity of what counts as evidence in assessment and the formulation of an assessment data strategy. The portfolio serves not only as a means for digital distribution of assessment materials to stakeholders, but also as a means to encode these elements in a machine-readable format for further analysis. In this way, the SCOPE+i method is focused on supporting local decision-making associated with assessment reform.

The following diagram provides a high-level view of the Assessment Portfolio. As the OSAF is engaging at the level of assessment practice, this diagram aims to clarify terminology in relation to usage. The diagram therefore includes a rudimentary level of functionality (RAiD) to help illustrate how the Assessment Portfolios can be used.



[31] Deliverable 2.1 OS-aware RRA approaches landscape report: https://doi.org/10.5281/zenodo.8301792

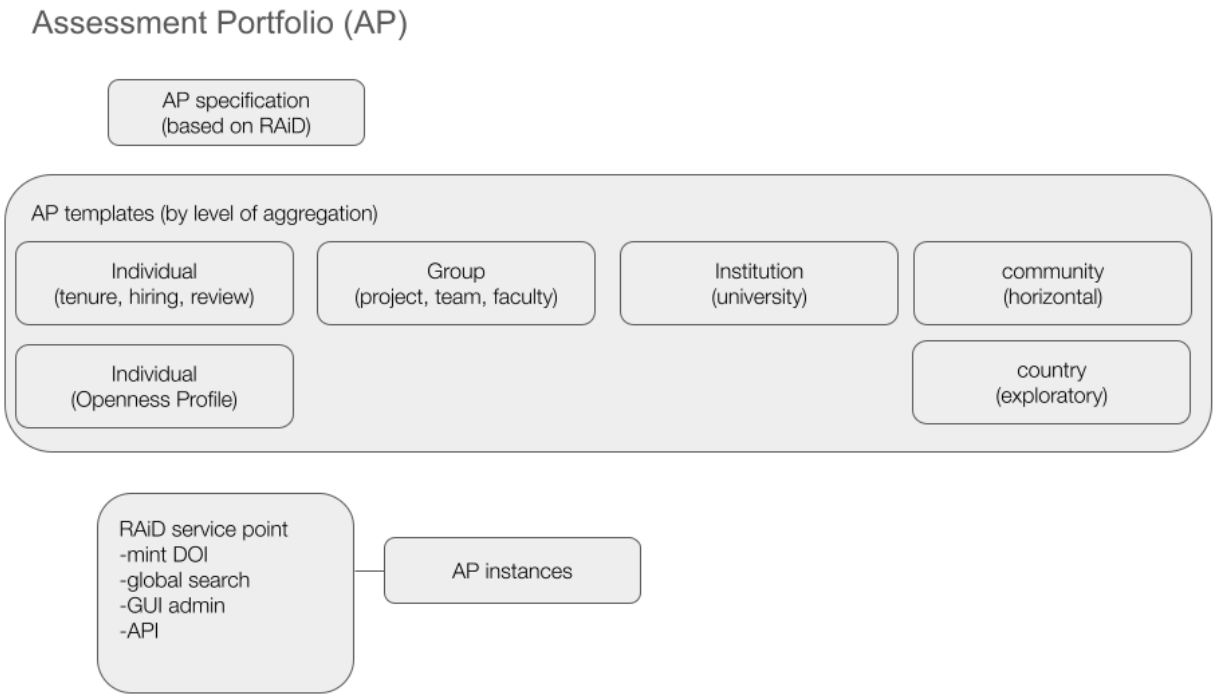


**Figure 1 Assessment Portfolio Architecture**

The top-level specification for Assessment Portfolios is shaped by the capacities of the RAiD, especially the metadata schema and the multi-actor interface. We then identify assessment-specific requirements (such as unique evidence types) not covered by RAiD metadata that together will be addressed as a RAiD metadata extension. From this top-level specification, we then develop differentiated assessment templates.

The templates, identified in Figure 1 above, represent Assessment Portfolios at different levels of aggregation, for which we anticipate the need to create differentiated templates. For example, to accommodate different information needs for individual and group assessments. In addition, there are two versions of individual portfolios. One is for assessment events and the other, Openness Profile, is an updatable display of an individual's Open Science activities.

For the highest level of aggregation, there are also two versions; one for research communities of practice (e.g., computer science) and the other for the country level of aggregation. Given the large scale of information anticipated for these two assessment use cases, we will explore the possibility of using multiple linked portfolios.

The templates will provide a generalized information model for each level of aggregation. An Assessment Portfolio instance occurs when a RAiD is registered, which mints a RAiD DOI and is then available to use.

## Assessment Registry

As the Assessment Registry design will be informed by the SCOPE+i method and Assessment Portfolios, it will be addressed in this next phase of the project. Accordingly, the Assessment Registry will be described in detail in the next version of this deliverable.

## Research Activity Identifier (RAiD)

Both the Assessment Portfolio and the Assessment Registry utilize the Research Activity Identifier (RAiD)[32] service, which is being implemented as an EOSC core component in the FAIRCORE4EOSC[33] project. Fundamentally, the RAiD is an editable information record, which is issued a persistent identifier (DOI). This combination enables curation of digital objects, by multiple actors, over time, whereby changes to the information record do not require versioning of the DOI.

> *The RAiD provides persistent, unique and resolvable information for research projects. The EOSC RAiD will mint Persistent Identifiers for research projects, which will allow users and services to manage information about project-related participants, services, and outcomes. RAiD also collects related identifiers (for, e.g., contributors, organizations, inputs, outputs, etc.) plus descriptive information about the project (e.g., title, description, subject, etc.) and stores them in a metadata record associated with the identifier. The EOSC RAiD implementation will allow authorized EOSC users and services to manage information about project-related participants, inputs, services, and outcomes.*[34]

In addition, the RAID was recently approved as an ISO standard.[35] While the EOSC RAiD service will be new, RAiD has been a service in Australia for several years. Albeit at a lower level of functionality. In joining the FAIRCORE4EOSC project, the Australian Research Data Commons (ARDC) realized a boost in their internationalization plan.

### EOSC RAiD

GraspOS is coordinating with FAIRCORE4EOSC regarding the use of the forthcoming EOSC RAiD production beta (March 2024), which enables early experimentation of the GraspOS pilot assessments. RAiD is a new persistent identifier developed by the Australian Research Data Commons (ARDC). The ARDC is the global RAiD authority, which provides global coordination of RAiD policy, a common metadata scheme, core functionality, and both API and GUI interfaces across RAiD Registration Agencies. Within the FAIRCORE4EOSC project the first RAiD Registration Agency outside of Australia will be implemented in the Netherlands, at the Dutch

[32] https://raid.org/
[33] FAIRCORE4EOSC website https://faircore4eosc.eu/
[34] EOSC Research Activity Identifier Service (RAiD), accessed 10 October 2023
[35] https://www.iso.org/standard/75931.html

collaborative organization for IT in education and research (SURF).[36] The SURF-based RAiD Registration Agency will support the EOSC RAiD service as well as the broader European region.

In the next phase of the GraspOS project (year 2), our focus is on the OSAF proof of concept, which is in part to mint and administer RAiDs to serve as Assessment Portfolios for the GraspOS pilots. We will use the RAiD beta, managed by ARDC and SURF as part of the FAIRCORE4EOSC project. Integration with GraspOS federated infrastructure will also begin in this phase, as we further develop the OSAF on pace with the pilot assessments. In the following we outline the rationale for using the RAiD.

### Affordances of RAiD

In this section, we outline the rationale for selecting the RAiD service, a research project identifier plus information record, for use with research assessment events. We also note its potential limitations. Drawing on the ARDC scope and approach document for the RAiD service,[37] a research project is defined as:

- An individual or collaborative enterprise initiated to undertake research (adapted from the OED)
- A planned or proposed research undertaking (adapted from the OED)
- A piece of work that is undertaken or attempted, with a start and end date and defined objectives (ARDC Vocabularies for Registry Schema 1.6.5).
- A privately or publicly funded project on a research topic. (EOSC RDM)
- An enterprise (potentially individual but typically collaborative), planned to achieve a particular aim. (schema.org)
- A temporary endeavor undertaken to achieve defined objectives. (DBPedia ontology).
- A Project is a planned activity with a budget, a sponsor, and a leader. (Simon Cox's Project Ontology)
- An administrative entity that enables an endeavor such as a research investigation. (FRAPO)

Research assessment can be included in this definition. Moreover, use of RAiD as a portfolio began a few years ago with collaboration among SURF, ARDC, and ORCID.[38] Which is to say, the portfolio use-case has been a topic of development discussion since 2018 and is accommodated in the present RAiD configuration. Recently, the GraspOS use case was presented to the RAiD Advisory Group meeting in Salzburg.[39]

Operational aspects of RAiD include a hierarchical, multi-actor administration framework that enables different levels of content administration. Administrators of a RAiD can manage access to its content.

[36] SURF is the collaborative organization for IT in Dutch education and research https://www.surf.nl/en
[37] RAiD Scope and Approach Shawn Ross | Updated 15 December 2022 | Draft (link)
[38] Tatum, McCafferty, and Brown. 2019. Openness Profile: mobilizing PIDs to increase visibility of open scholarship https://zenodo.org/doi/10.5281/zenodo.2549269
[39] Tatum, C. (2023). [RAiD advisory group meeting] GraspOS overview. RAiD advisory group meeting, Salzburg, Austria. Zenodo. https://doi.org/10.5281/zenodo.10374861

Employing RAiD as an Assessment Portfolio offers key affordances to the process of conducting research assessment events. To the extent that research assessment entails collecting evidence, RAiD has a persistent identifier that provides globally unique identification for assessment content that is both machine-readable and resolvable to a webpage. The inclusion of an editable metadata record[40] combined with a robust research activity metadata schema[41] provides the possibility of systematic cataloging of assessment actors, documents and evidence. The RAiD metadata scheme supports a narrative component along with accommodating a wide range of contributions to be considered in an assessment context, along with the ability to document associated contextual factors.

Using the RAiD for compiling evidence for research assessment purposes facilitates a collaborative approach (RAiD is a multi-actor device), while also providing interoperability with contemporary research information systems via an application programing interface (API) and data portability in situations where compatible systems are not readily available.

Like most common PID systems (e.g., Crossref, RoR, DataCite, and ORCID), the RAiD system provides an API[42] for publishing, distributing, and/or ingesting the contents of RAiDs—in this instance, the contents of Assessment Portfolios. In this context, Assessment Portfolios provide the basis for agreeing on, collecting, and distributing the content to be considered in the evaluation event.

Using the FAIR principles as a reference point, being findable is “arguably the most important [principle] because it will be hard to achieve other aspects of FAIR without globally unique and persistent identifiers”.[43]

Implementation of EOSC RAiD service establishes RAiD as an EOSC core service, and thus aligned with EOSC interoperability and PID policies. Via the FAIRCORE4EOSC project, RAiD will be integrated in the EOSC platform/market, in the EOSC (DataCite) PID Graph, and in the EOSC (OpenAIRE) Research Discovery Graph. This points to a desirable level of compatibility for operationalizing OSAF.

However, as EOSC itself could entail operational overhead and possibly some schedule uncertainty, it is presently unclear whether the EOSC RAiD service will be the most suitable platform for piloting the Assessment Portfolio and Assessment Registry. To minimize this risk, we have the option of using the RAiD Registration Agency at SURF, which is also being implemented through the FAIRCORE4EOSC project.

In the coming GraspOS phase, we will work closely with WP3 and WP4 colleagues to assess these and other options for deployment of the OSAF.

[40] this combination is similar to ORCID (persistent identifier (PID) and editable content record)
[41] RAiD metadata schema: https://metadata.raid.org/en/latest/
[42] RAiD API documentation https://api.demo.raid.org.au/swagger-ui/index.html
[43] GoFAIR website, accessed 11 October 2023: https://www.go-fair.org/fair-principles/f1-meta-data-assigned-globally-unique-persistent-identifiers/

# 4. Assessment Resources (SCOPE+i method)

In section 3, we described the SCOPE+i method in general terms. In this section, we describe the planned content of the method; the resources aimed at facilitating the assessment process. At present, these resources are organized in relation to corresponding assessment phase(s). See table 1. for the correlation between a) assessment phases, b) SCOPE+i method, and c) assessment infrastructure.

Following is a short description for each of the planned resources, which are organized into three categories: 1) templates, 2) guidelines & checklists, and 3) open indicators & data sources. As we expect the co-development process to introduce new resource needs, the present listing is considered non-exhaustive. And while there are in some cases existing related templates and guidelines that may be useful (see D2.1 OS-aware RRA approaches landscape report), the resources below will be tailored specifically to Open Science and/or Responsible Research Assessment.

## Templates

### 1. Assessment readiness template

The aim of the assessment readiness template is to describe the current status of the entity's research evaluation aims, context, and resources. The template will help to establish the level of maturity towards research assessment reform, collect contextually relevant information to facilitate shaping of the evaluation event focused on Open Science practices, and provide information on the initial requirements and the current best practices for indicators, data, tools and services. This template will build on and update the Pilot analysis template developed for the GraspOS pilots.

### 2. Stakeholder mapping template

The aim of the stakeholder template is to identify the relevant stakeholders of the entity in question. The stakeholder can refer to individuals or institutions, and everything in between, depending on the entity in question. The important questions to keep in mind are who determines what is valued, and who defines the purpose of the evaluation. This template will build on and update the Stakeholder mapping template developed for the GraspOS pilots.

### 3. Value(s) statement template

Value statements help to identify what is valued about the entity under evaluation. Based on the definition in the SCOPE framework[2], a value is a judgment made about what is important. Value judgments can be done at different granularity levels. This template will build on and update the value statement template developed for the GraspOS pilots.

### 4. Purpose statement template

The aim of the purpose statement template is to help describe the purposes of the entity's research evaluation processes. The template will help to consider the specific needs and requirements for evaluation events (indicators, methods, data, tools and services) according to the purpose of

evaluation (e.g., monitoring, learning and improvement, or resource allocation and career assessment) and the level of assessment (e.g. individual, unit, institution, country). This template will first be developed for the GraspOS pilots, then modified for general use.

### 5. Contextual factors template

TBD

### 6. Narrative template

The aim of the narrative template is to facilitate structured and evidence-based input of narrative information to support assessment. Building on the Résumé for Researchers template,[44] the narrative template will provide prompts and definitions, and possibly a module, for recognizing a broad range of qualities, impacts, contributions and Open Science practices, with instruction as to the documentation of evidence if required (e.g., using Openness profiles).

### 7. Strategy template

The aim of the strategy template is to facilitate the assessment of the entity in light of its own aims and strategy. Building on the Dutch Strategy Evaluation Protocol (SEP), the strategy template helps evaluands to outline their aims and ambitions (e.g., in research, education, outreach, Open Science), the plan of action to achieve these aims, as well as the documentation and indicators suited to monitor their achievement.

## Guidelines & Checklists

### 8. Open Science Assessment Guide

The aim of the Open Science guidelines as a starting point is to help focus on operationalizing the assessment specifics of Open Science (also in relation to RRA).

### 9. Guidance on the diversity of OS contributions, roles, and activities

The aim of the Guidance on the diversity of OS contributions is to ensure that a wide range of practices and activities are considered, and that all who contributed are recognized.

### 10. Assessment team guidelines

The aim of the assessment team guidelines is to help identify the kinds of roles needed in an evaluation team, for example to ensure that the basics of the evaluation (SCOPE 1 and 2) can be translated to the next stages. Here we address assessment actors. Above, in the Stakeholder mapping template, we address a wider range of participants who are needed for consensus, for

[44] See https://royalsociety.org/-/media/policy/projects/research-culture-images/2019-10-research-culture-resume-for-researchers-template.pdf for the template and https://royalsociety.org/topics-policy/projects/research-culture/tools-for-support/resume-for-researchers for more information on the template.

example on the values and purpose of the assessment. In practice, there would likely be an overlap between the assessment team and the stakeholders.

### 11. Guidelines for evaluators and evaluands

The aim of guidelines for evaluators and evaluands is to help understand the basic framework of rules and principles for individuals to conduct themselves in the role of evaluators and evaluands according to international/national legal regulations (e.g. laws and rights on gender equality and non-discrimination, European Charter for researchers), research integrity and ethics codes (e.g. All European Academies ALLEA), and key RRA and metrics recommendations (e.g. DORA, Leiden Manifesto, CoARA).

### 12. Guidance for overcoming common obstacles in implementing RRA

This guide aims to help address common obstacles identified in the GraspOS Surveys. The two surveys conducted for the landscape analysis show that the situation and challenges of the nine GraspOS pilots vis-à-vis CoARA Agreement and assessment practices are indeed very similar compared to the 54 landscape survey participants from 19 European countries. Based on these survey results it is now possible to identify guidelines to overcome common obstacles.

### 13. Guidance on equity, diversity, inclusion (EDI)

The aim of the Guidance on equity, diversity and inclusion (EDI) is to facilitate consideration in an entity's assessment to the aspects contributing to such issues as career stage, field or discipline, multi-, inter-, and trans-disciplinarity, basic vs. applied research, inter-sectorality, gender, sexual orientation, racial/ethnic origin, socio-economic status, disability and language.

### 14. Guidance on translating values, purpose and context into an assessment protocol

The aim of the Guidance is to support the selection of tools and indicators for an assessment and to ensure that the selected tools and indicators are aligned with what is valued about the entity under evaluation and also with the purpose and context of the assessment.

### 15. Checklist for responsible assessments

This checklist for responsible assessments will build on and further develop recommendations identified in the OS-aware RRA approaches landscape report and incorporate insights from a similar resource developed by TSV for use in Finland.

(see appendix 1, the Self-evaluation tool for culture of open scholarship services.

### 16. Guidance/template on what to include, how to document an assessment protocol

The aim of guidance and associated template is to provide a working definition for an OS aware RRA protocol and the kinds of information included for both the assessment event and the subsequent publication of the assessment protocol.

### 17. Guidelines for evaluating the evaluation

The aim of Guidelines for evaluating the evaluation is to help establish the criteria for evaluating the evaluation and to identify which actors could be included (e.g., the evaluands).

## Open Infrastructure, Indicators & Data Sources

### 18. GraspOS tools & services catalog

GraspOS will deliver a set of catalogs of resources that can support Open-Science-aware Responsible Research Assessment events. The most important ones are the Tools Catalog, which will facilitate the discovery of enrichment and monitoring tools, and the Services Catalog, which will enable the discovery of enrichment, monitoring, and data services. Both catalogs will be available to the end-users through a Web-based front-end, which can be used as an inventory in which evaluators can find various options on tools and services that can address their needs in the context of specific assessment events.

### 19. Indicator toolboxes

The guidelines for 'indicator toolboxes' will be developed from relevant literature (e.g., the EC expert committee report on Indicators Framework) and observations of the specificities from each pilot.

### 20. Open research information sources

This will be a curated list of assessment-related open data sources. As these sources are often also provided by so-called open infrastructure, the CoARA working group Towards Open Infrastructure for Responsible Research Assessment (OI4RRA)[45] will be a key resource. The GraspOS Data Registry will serve as a resource that collects such sources and provides access information and other metadata regarding them.



[45] See https://coara.eu/app/uploads/2023/10/WG-Overview_Towards-Open-Infrastructures-for-Responsible-Research-Assessment_updated.pdf

# 5. Next steps

This deliverable brings the Open Science Assessment Framework (OSAF) concept into sharper focus. Looking ahead to the next development phase, the priorities are continued development of the guidelines and associated resources, advancing the Assessment Portfolio to proof of concept, and embedding the OSAF in the GraspOS federated infrastructure.

1. Develop the section 4 assessment guidelines and associated resources, drawing on the pilots' activity and insights from the D2.1 Landscape Analysis.

2. Mapping evaluable contributions to OS-aware RRA.
    1. Document OS practices and contributions from the pilot analyses
    2. Map these to RAiD metadata schema
    3. Identify RAiD metadata limitations
    4. Develop workarounds for OSAF
    5. (Ideally) develop assessment metadata scheme as RAiD extension

3. Integrating OSAF (WP2) in GraspOS services and infrastructure (WP3, WP4). Preferably, this would entail an integration meeting format, whereby assessment workflows are developed and used to identify and align end-to-end requirements.

# 6. Annexes

## Annex 1. Template for the pilot analysis

Introduction

This analysis is conducted to describe the current status of the pilot's research evaluation aims, context, and resources. The analysis:

a) establishes the pilot´s ambitions and level of maturity towards research assessment reform,
b) collects contextually relevant information to facilitate shaping of the pilot evaluation event focused on open science practices (supporting WP2), and
c) provides information on the initial requirements and the current best practices for indicators, data, tools and services (supporting WP3 and WP4).

The analysis has three parts. The pilots will describe:

**1) Pilot's state of affairs** in terms of
- 1.a) Open Science & Research Assessment practices
- 1.b) Tools/services and datasets used
- 1.c) Gaps

**2) Evaluation context**
**3) Pilot ambitions** in terms of developing new ways of evaluating/monitoring open science

Instructions on using the template

The template includes supporting questions under each section. However, as the pilots are diverse in their needs and vision, as well as their level of maturity in relation to research assessment reform and infrastructure, pilots can choose which of these questions are relevant to their setting. The aim of this analysis is not to rush the pilots into making quick decisions in regard to open science aware research assessment and how it should be conducted in the future, instead it is aimed at supporting the pilots in identifying their current status and future ambitions.

**Please use the headings provided in the template to structure your analysis. Please use the tables provided in the template. Other sections are free text questions.**

The analysis should be uploaded to the GraspOS Google Drive directory folder titled Pilot analyses in word format: https://drive.google.com/drive/folders/1zyCJq-2S6k7PnOHrSpp06ReChZUtc2FY?usp=share_link

**The deadline for the analysis is 14 July, 2023, EOD.**
More practical instructions will be provided in the upcoming monthly meetings for WP5 (to be scheduled later in March).

*Please check the goals and links to GraspOS as well as the KPIs listed in the proposal, and update them if necessary.*

| | | |
|---|---|---|
| Goals & links to GraspOS | from the proposal | updates if needed |
| KPIs | from the proposal | updates if needed |

1. Local state of affairs

a) Open science and/or research assessment (free text question)
SUPPORTING QUESTIONS to consider if relevant:

- In the context of your pilot, what does open science (OS) aware research assessment (RA) mean? How are OS practices considered and acknowledged in RA at the moment?
- What is the pilot's operational framework in terms of OS and/or RA: what kinds of policies, guidelines, principles, and/or best practices are there that refer to OS and/or RA (nationally, institutionally, within a discipline, whatever is the relevant level for the pilot)?
- What is the pilot's operational environment in terms of OS and/or RA: what kinds of developments are going on at the moment within the pilot's context, and how is the pilot connected to those developments? Does the pilot have access (influence) to these developments?
- The dominant practices within OS and/or RA: who or what are the dominant actors, i.e., who is in charge, what are the dominant practices, where and how is OS and/or RA defined and/or developed, and what is the level of awareness of OS and/or responsible research assessment (RRA) within the dominant practices and/or actors? Can you identify current RA practices that are not responsible?

b) Tools, services and data used
*Instructions: Please use the table to list the targets of evaluation or monitoring, and the tools, services or data sources used. Also list any challenges you may have encountered using different tools, services or data sources.*

| Target of assessment/monitoring | Tools, services used *) | Data source(s) **) | Challenges ***) |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |

*) indicators, metrics, monitoring tools (dashboards, PowerBI, aggregations, etc.), services supporting calculations, implementation, different templates (portfolio, cv, narratives, etc.)

**) include also if they are internal or external, national or international, if there are issues with quality of data, etc.
***) limitations, restrictions, openness to interpretations, etc. to do with tools, services and/or data sources used
SUPPORTING QUESTIONS to consider if relevant:
(free text question)

- Is there a current research information system (CRIS) in use, and if so, is it:
  - Pure, Converis, Symplectic Elements, or some other solution?
  - in-house solution?
- Is there an institutional repository (for publications), if not which repositories are you suggesting/using
- How is your infrastructure linked to the EU/EOSC?
- Is there a personnel management system?
- Is there a researcher profile system or some tools to create online CVs and/or academic profiles?
- Are there other local platforms in use?
- Is there support/information on OS available
- Are there research practices that touch upon FAIR or open data, or software?

c) What are you missing currently in terms of tools, services, data (indicators, metrics, quality, protocols, practices,...)? (free text question)

2. Evaluation Context (free text question)
SUPPORTING QUESTIONS to consider if relevant:

- What is the broader function of research evaluation for this pilot context? For example, in what ways do research evaluation events link to collective values (e.g in a mission statement or research agenda, or in normative practices)?
- Who are the intended stakeholders, audience(s), and beneficiaries for the outcome of the pilot evaluation?
- Which contextual factors are relevant for shaping the research evaluation criteria and content (e.g. level of analysis, institutional structures, disciplinary structures, etc.)
- Who are the key actors (by title) in preparing and conducting this pilot evaluation event?

Relevance of practices
Are the following relevant to monitoring OS in your setting? Are they relevant to research evaluation for researchers or your organization?

*Instructions: Please indicate which (if any) of the following open practices are considered or acknowledged in monitoring and/or evaluation within the pilot setting. Also, indicate on which types of OS activity (e.g., outputs, usage, practices) data is collected. Add new rows if necessary.*

| | Monitoring | Research evaluation |
|---|---|---|
| ***Open practices considered/acknowledged in monitoring and/or evaluation*** | | |
| ***FAIRness*** | | |
| ***Reproducibility*** | | |
| ***Interdisciplinarity*** | | |
| ***Research Outputs usages*** | | |
| ***Collaboration*** | | |
| ***Other practices*** *[please add rows if needed]* | | |
| ***Data collected or used for monitoring and/or evaluating OS*** | | |
| ***Open research outputs*** *[E.g., which research results are relevant? Publications, data, software, services, tools, other? Is FAIR data important in monitoring or assessment in your setting? If yes, how is it captured? Are you using assessment tools? Does institutional/national repository offer FAIR by design /default?]* | | |
| ***Links among research outcomes*** *[E.g., for monitoring: does it make any sense on how objects are linked together. E.g., provenance]* | | |
| ***Usage of research outcomes*** *[E.g., is usage (clicks/downloads, citations) taken into account in monitoring/assessment? If yes, how? If yes, how do you have access to these?]* | | |

| | | |
|---|---|---|
| ***Usage of services***<br>*[E.g., EOSC is keen on publishing and sharing services. How important is this for monitoring research or open science in your environment?]* | | |
| ***Data on practising OS***<br>*[E.g., researchers practice OS in many ways that are not captured by OA/FAIR infrastructure. E.g., data stewardship, citizen engagement, open-source community software. What other practices do you, or would you like to measure? How do you record these?]* | | |
| ***Other***<br>*[please add rows if necessary]* | | |

3. **Pilot ambitions** in terms of developing new ways of evaluating/monitoring open science (free text question)

SUPPORTING QUESTIONS to consider if relevant:

- Why evaluate OS: what does the pilot use the evaluation/monitoring of OS for at the moment, and what the pilot hopes to use it for in the future?
- What does the pilot want to develop and why: what is valued, i.e., what would be evaluated/monitored, and how could/should it be evaluated/monitored?
- Gap analysis between what the pilot wants to develop and what are the dominant practices: what do you think you CAN change, and how? How can GraspOS support here, i.e., what is the value proposition to dominant actors and practices to cooperate with GraspOS. What is the added value of being a part of the project?
- Do you anticipate some challenges or barriers involved in developing new ways of evaluating/monitoring open science (to do with, for example, lack of data or information, lack of indicators, issues to do with local privacy regulations, etc.)?

Considering how GraspOS tools/services (WP3) and datasets (WP4) may be helpful

*<u>Instructions:</u> Please indicate in the second column of the next table which GraspOS tools/services and datasets could be interesting for your pilot by briefly describing relevant use case scenarios (no obligation to use them is implied - just expression of relevance/interest). In addition, please include in the third column some special requirements for these tools/services or datasets that may be important for your pilot. Detailed descriptions for all GraspOS tools/services and datasets can be found in*

*Appendix A. If a tool/service or dataset seems to be irrelevant to your use cases just leave blank the respective cell.*

| **GraspOS service: Enrichment of research outputs with missing attributes** | | |
|---|---|---|
| **Component** | **Usage scenarios** | **Requirements** |
| OpenAIRE Broker<br>*License: AGPL*<br>*TRL (start/end): 9/9*<br>*Partners: OpenAIRE, CNR* | | |
| OpenAIRE IIS text mining modules<br>*License: AGPL*<br>*TRL (start/end): 9/9*<br>*Partners: OpenAIRE, ARC* | | |
| OpenAIRE Metadata Validator<br>*License: AGPL*<br>*TRL (start/end): 7/9*<br>*Partners: OpenAIRE* | | |
| SCRE Pipeline<br>*License: AGPL*<br>*TRL (start/end): 9/9*<br>*Partners: OPERAS* | | |
| **GraspOS service: Enrichment of research output links** | | |
| **Component** | **Usage scenario** | **Requirements** |
| Semantic Citation Classifier<br>*License: MIT*<br>*TRL (start/end): 4/7*<br>*Partner: UNIBO* | | |
| BIP! citation classifier<br>*License: GNU/GPL*<br>*TRL (start/end): 3/7*<br>*Partner: ARC* | | |
| **GraspOS service: Enrichment of research outputs with novel metrics** | | |
| **Component** | **Usage scenarios** | **Requirements** |
| **EC KIP OS indicators**<br>*License: GNU/GPL*<br>*TRL: (start/end): 6/9*<br>*Partner: ARC* | | |

| | | |
|---|---|---|
| BIP! Services (Toolbox)<br>*License: GNU/GPL*<br>*TRL (start/end): 9/9*<br>*Partner: ARC* | | |
| **GraspOS service: Enrichment of research outputs with usage data** | | |
| **Component** | **Usage scenarios** | **Requirements** |
| **EOSC accounting for service**s<br>*License: Apache*<br>*TRL (start/end): 5/7*<br>*Partner: GRNET* | | |
| UsageCounts<br>*License: AGPL*<br>*TRL (start/end): 9/9*<br>*Partner: OpenAIRE* | | |
| OPERAS Metrics<br>*License: MIT*<br>*TRL (start/end): 7/9*<br>*Partner: OPERAS* | | |
| **GraspOS service: OS Institutional Dashboard** | | |
| **Component** | **Usage scenarios** | **Requirements** |
| OpenAIRE Institutional Monitor Dashboard<br>*License: AGPL*<br>*TRL (start/end): 9/9*<br>*Partner: OpenAIRE* | | |
| **GraspOS service: EOSC OS Researcher Dashboard** | | |
| **Component** | **Usage scenarios** | **Requirements** |
| BIP! Scholar<br>*License: GNU/GPL*<br>*TRL (start/end): 7/9*<br>*Partner: ARC* | | |
| **GraspOS resource: Scholarly resources** | | |
| **Resource** | **Usage scenarios** | **Requirements** |
| OpenAIRE Graph<br>*Partner: OpenAIRE* | | |
| OpenCitations | | |

| | | |
|---|---|---|
| *Partner: UNIBO* | | |
| Scholexplorer<br>*Partner: OpenAIRE* | | |
| BIP! DB<br>*Partner: ARC* | | |
| OpenAIRE Usage Counts<br>*Partner: OpenAIRE* | | |
| OPERAS Metrics<br>*Partners: OPERAS* | | |
| **FAIRCORE4EOSC RAiD**<br>*Partners: CWTS, OpenAIRE, CSC* | | |

# Annex 2. GraspOS Tools/Services descriptions

| **GraspOS service: Enrichment of research outputs with missing attributes** | | |
|---|---|---|
| **Component** | **Description** | **GraspOS extensions** |
| OpenAIRE Broker<br>*License: AGPL*<br>*TRL (start/end): 9/9*<br>*Partners: OpenAIRE, CNR* | A subscription-notification service that enables content providers (repositories, CRIS systems, aggregators, knowledge graphs, publishers) to enrich content with additional metadata. It utilizes the OpenAIRE Graph, which has a key role in the de-centralization and local re-use of all metadata available. | Extend current data model to cover additional research results, relationships between products, classifications, and other metrics included in metadata records. |
| OpenAIRE IIS text mining modules<br>*License: AGPL*<br>*TRL (start/end): 9/9*<br>*Partners: OpenAIRE, ARC* | An information inference service that enriches scholarly data with automatically inferred metadata, based on a flexible big data processing pipeline supporting full-text and metadata mining. Current modules include citation extraction (article-data, article-software, product-grant, product-organizations); subject inference (Frascati and SDG); community context. | Extend with methods for the completeness of metadata (e.g., resource types, subjects, licensing typology) and, in particular for the pilots, deeper classifications (>Frascati Level 4) to be used in capturing discipline specific characteristics. |
| OpenAIRE Metadata Validator<br>*License: AGPL*<br>*TRL (start/end): 7/9*<br>*Partners: OpenAIRE* | A rule-based service that provides metrics on FAIR assessment of content providers (data, software, publication repositories/journals, CRIS systems) based on metadata of the research objects. | Extend with configurations for custom and domain-specific FAIRness metrics at the level of the individual records and average-based metrics for data sources, funders, institutions. Publish as a stand-alone service (now embedded in PROVIDE Dashboard). |
| SCRE Pipeline<br>*License: AGPL*<br>*TRL (start/end): 9/9*<br>*Partners: OPERAS* | A general and configurable AI-assisted content acquisition and processing pipeline for documents and projects that aggregates and performs data cleaning and semantic processing. | Extend with methods for the completeness of metadata for SSH. Upgrade so outputs match the specifications from RDA IG Scientific Knowledge Graphs. |
| **GraspOS service: Enrichment of research output links** | | |
| **Component** | **Description** | **GraspOS extensions** |

| | | |
|---|---|---|
| Semantic Citation Classifier<br>*License: MIT*<br>*TRL (start/end): 4/7*<br>*Partner: UNIBO* | A tool that performs automatic annotation of citations in academic papers, by processing bibliographic data and the text in which the citation appears. It enriches each individual citation with structural and semantic features combining deep learning and problem reduction techniques. | Extended to detect a larger set of citation intents. We will extend the training dataset and refine the classification techniques used so far in order to increase accuracy and coverage of domains. New data fusion techniques will also be implemented - for instance, the reduction to Semantic Retrieval (SR). Multi-faceted classification will also be explored, and we will build a more flexible and robust module to export citation data and link them to existing datasets. |
| BIP! citation classifier<br>*License: GNU/GPL*<br>*TRL (start/end): 3/7*<br>*Partner: ARC* | A tool that automatically annotates citations based on their intent. | Currently the tool is focused on CS publications, the plan is to be extended for other fields. |
| **GraspOS service: Enrichment of research outputs with novel metrics** | | |
| **Component** | **Description** | **GraspOS extensions** |
| **EC KIP OS indicators**<br>*License: GNU/GPL*<br>*TRL: (start/end): 6/9*<br>*Partner: ARC* | A suite of advanced software components for indicators on impact and collaboration of OA based research outputs on citation and network analysis (with emphasis where possible on timeliness): Field-Weighted Citation Impact - FWCI scores, Collaborative Index (CI), Degree of collaboration (DC), Collaborative coefficient (CC). | Extend to offer aggregate metrics for funders, research groups and institutions and with models that cover additional research products, not just publications. Use pilots for benchmarking and further extensions. |
| BIP! Services (Toolbox)<br>*License: GNU/GPL*<br>*TRL (start/end): 9/9*<br>*Partner: ARC* | A suite of advanced software components that leverages scholarly knowledge graphs (Crossref, OpenCitations, OpenAIRE Graph) and produces advanced citation-based indicators for publications and researchers that capture popularity (= short-term impact), influence (= long-term impact), impulse (= initial momentum). | Extend to offer aggregate metrics to research groups and institutions and to provide additional metrics that will cover (a) EOSC service usage and (b) the level, impact, collaboration, and timeliness in OS practice for researchers, groups, and institutions. |
| **GraspOS service: Enrichment of research outputs with usage data** | | |
| **Component** | **Description** | **GraspOS extensions** |

| | | |
|---|---|---|
| **EOSC accounting for service**s<br>*License: Apache*<br>*TRL (start/end): 5/7*<br>*Partner: GRNET* | A tool that collects service metrics such as Virtual Access Metrics and aggregates them according to EOSC/community rules. It would be used as a data source for impact metrics for the services as part of the federated metrics infrastructure. | Extend with an add-on and APIs to provide aggregate level metrics and reports. |
| UsageCounts<br>*License: AGPL*<br>*TRL (start/end): 9/9*<br>*Partner: OpenAIRE* | A service that collects usage data from OS content providers repositories, journals, and other scientific data sources. Using COUNTER code aggregates, deduplicates and delivers standardized activity reports (SUSHI) about research usage and uptake. Offers data collection from MakeDataCount. | Extend with an add-on and APIs to provide metrics and reports for funders, organizations and individual researchers. Be the key technology in the OA Trust eBook and the IDS prototype. |
| OPERAS Metrics<br>*License: MIT*<br>*TRL (start/end): 7/9*<br>*Partner: OPERAS* | The service collects usage and alternative metrics for OA monographs and books: downloads, web visits, tweets, Wikipedia mentions, etc. It is based on a shared data model. | Extend with an add-on and APIs to provide aggregate level metrics and reports, accommodating also OA journal articles. |
| **GraspOS service: OS Institutional Dashboard** | | |
| **Component** | **Description** | **GraspOS extensions** |
| OpenAIRE Institutional Monitor Dashboard<br>*License: AGPL*<br>*TRL (start/end): 9/9*<br>*Partner: OpenAIRE* | A service built on the OpenAIRE Research Graph providing monitoring services for research outputs of research actors. The service offers OS statistics focusing on products, funders, and sub-units of a given institution. It is highly configurable to accommodate mix & match indicators and metrics. | Extend with additional indicators. Enhance the existing administration backend with pre-set templates and support others for building additional ones tailored to their needs. |
| **GraspOS service: EOSC OS Researcher Dashboard** | | |
| **Component** | **Description** | **GraspOS extensions** |
| BIP! Scholar<br>*License: GNU/GPL*<br>*TRL (start/end): 7/9*<br>*Partner: ARC* | A dashboard that leverages data from scholarly knowledge graphs (OpenAIRE Graph Crossref, OpenCitations) and ORCID to display custom reports containing a variety of | Extend to provide detailed reports on researchers' OS practice (level and impact). The reports will include intuitive visualizations and will also cover groups of researchers. |

| | | |
|---|---|---|
| | researcher-level RRA indicators capturing different aspects of their performance (e.g., productivity, impact, career stage) while considering the different roles of researchers in the respective works (according to the CRediT taxonomy). | |
| **GraspOS resource: Scholarly resources** | | |
| **Resource** | | |
| OpenAIRE Graph<br>*Partner: OpenAIRE* | A service that populates and provides access to a scientific knowledge graph that includes metadata and links between scientific products, organizations, funders, funding streams, projects, communities, and (provenance) data sources. | |
| OpenCitations<br>*Partner: UNIBO* | An infrastructure provided by a not-for-profit organization that publishes open bibliographic and citation data by the use of Linked Data technologies. Its main Index (COCI) currently contains more than 1.29M citation links. | |
| Scholexplorer<br>*Partner: OpenAIRE* | A service that populates and provides access to a graph of links between dataset and literature objects and dataset and dataset objects. | |
| BIP! DB<br>*Partner: ARC* | A dataset that contains citation-based impact indicators for more than 138M articles considering a citation network that consists of more than 1.672B citations. | |
| OpenAIRE Usage Counts<br>*Partner: OpenAIRE* | A dataset containing usage data for various types of research products (e.g., publications, datasets, etc.). | |
| OPERAS Metrics<br>*Partners: OPERAS* | The service collects usage and alternative metrics for OA monographs and books: downloads, web visits, tweets, Wikipedia mentions, etc. It is based on a shared data model. | |
| **FAIRCORE4EOSC RAiD**<br>*Partners: CWTS, OpenAIRE, CSC* | A resource collecting research activity identifiers (i.e., research projects), implemented as an EOSC-integrated service in the context of the FAIRCORE4EOSC project. | |

# Annex 3. Stakeholder mapping template

**GraspOS Pilot Assessment Stakeholder Mapping**

With this template, we aim to facilitate stakeholder mapping for the workshop on 2nd October 2023. Below are some questions to help you start thinking about this mapping exercise:

Start with what you value *about the thing you are evaluating/monitoring*:

- Who determined what is valued?
- Was it a collaborative effort?
- Who wasn't talked to (but might have been relevant)? Why?

Determining the purpose of the evaluation:

- Who defines the purpose?
- What does that imply for the assessment?

Step 1. Table for collecting stakeholder information

| **stakeholder name/title** | **affiliations(s)** | **role(s) in the assessment** | **relationship to the outcome** | **stage of evaluation: planning/conducting/utilizing** |
|---|---|---|---|---|
| | | | E.g. decision-making, consulting, following, evaluand(s), evaluator, reporting… | |
| *(example) Name /* Professor of… | CWTS, Leiden U QSS editor | member of evaluation organizing committee conducted self-assessment evaluand | decision-making evaluand reporting | |

Step 2. Illustrating stakeholder relationships

In this step, we would like to invite you to present your stakeholder mapping findings to us. This could be in the format of one slide with simple text, but this could also be a diagram, a drawing or another type of visual material that fits the data you will be presenting.