MAKING FAIReR ASSESSMENTS POSSIBLE

FINAL REPORT OF
EOSC CO-CREATION PROJECTS

EUROPEAN OVERVIEW OF CAREER MERIT SYSTEMS

AND

VISION FOR RESEARCH DATA IN RESEARCH CAREERS

PROJECT DESCRIPTION

This report is a deliverable of EOSC Co-Creation projects (i) "European overview of career merit systems" and (ii) "Vision for research data in research careers", funded by the EOSC Co-Creation funding. Further information on these projects can be found here: **https://avointiede.fi/en/networks/eosc-co-creation**.

EOSCsecretariat.eu has received funding from the European Union's Horizon Programme call H2020-INFRAEOSC-2018-4, Grant Agreement number 831644 Europeana is an initiative of the European Union, financed by the European Union's Connecting Europe Facility and European Union Member States (**https://pro.europeana.eu/** and **https://www.europeana.eu**).

PROJECT TEAM

The project was led and coordinated by the Federation of Finnish Learned Societies, Finland:

Henriikka Mustajoki (ORCID 0000-0002-5625-833X)
Janne Pölönen (ORCID 0000-0003-1649-0879)
Elina Koivisto (ORCID 0000-0002-7637-5509)
Elina Pylvänäinen (ORCID 0000-0003-4829-0484).

Dragan Ivanović (University of Novi Sad, Serbia; ORCID 0000-0002-9942-5521) and **Kathleen Gregory** (DANS, The Netherlands; ORCID 0000-0001-5475-8632) have been contracted as experts to carry out in depth analyses, surveys, vision-building and drafting the report.

Valerie Brasse (IS4RI, France: ORCID 0000-0001-5395-4818) and **Joonas Kesäniemi** (CSC/Aalto University, Finland: ORCID 0000-0002-3770-0006) have been contracted as experts to carry out case-studies.

The steering group consisted of 5 persons representing key stakeholders and expertise: **Maijastiina Arvola** (EOSC-Nordic/ CSC, Finland), **Tim Engels** (ENRESSH/ECOOM Antwerp, Belgium), **Elisabeth Gadd** (INORMS/Loughborough University, UK), **Bregt Saenen** (European University Association, Belgium), **Kata-Riina Valosaari**, (Academy of Finland). The steering group contributed by advising and commenting work-in-progress, and disseminating survey, vision-draft and final results. We also acknowledge the contributions of all experts who participated in the survey, case-studies, bootcamps, and public consultation.

CITE AS:

Henriikka Mustajoki, Janne Pölönen, Kathleen Gregory, Dragan Ivanović, Valerie Brasse, Joonas Kesäniemi, Elina Koivisto, Elina Pylvänäinen (2021): Making FAIReR assessments possible. Final report of EOSC Co-Creation projects: "European overview of career merit systems" and "Vision for research data in research careers".

https://doi.org/10.5281/zenodo.4701375

Report design & infographics
KRUT Collective, krut.fi

Federation of Finnish Learned Societies

EUROPEAN OPEN SCIENCE CLOUD

EXECUTIVE SUMMARY

In this report, the co-creation team builds on research and co-creation processes to argue for creating FAIReR academic assessments. Such assessments are rooted in both the FAIR guidelines for data management and policies for the responsible assessment of research (FAIReR = FAIR + Responsible). Specifically, FAIReR assessments emphasise diversity, communities, and dialogue.

Recommendations for realising FAIReR assessments:

1. Communities co-create the meaning of diversity in assessments
MAKE IT MEANINGFUL We need to know what we want to value and evaluate. To do this, we start by considering the goals of open science and do not limit our evaluations to what is technically possible or easy to measure. We take into consideration the diversity of practices, outputs, missions and impacts of academic work, and differences between fields. In the case of research data, for example, such practices may include sharing (open) data, creating FAIR data, reusing data, or cultivating expertise in creating or curating FAIR data.

2. Communities build assessments on infrastructures capturing diversity
MAKE IT POSSIBLE We need to make it possible for researchers to report, make visible, and explain their diverse outputs, activities and impact of their work. Integration of relevant information from different sources is facilitated by open assessment infrastructure. In the case of research data, information on creating, publishing and sharing research data needs to be reliable, comprehensive, comparable and structured.

3. Communities reward diverse open science and FAIR practices
MAKE IT REWARDING We need to include a broad range of outputs, activities and impacts of academic work in criteria for hiring, promotion and funding. In the case of research data, this may include shared or open data, indications of data reuse, or acting as data steward.

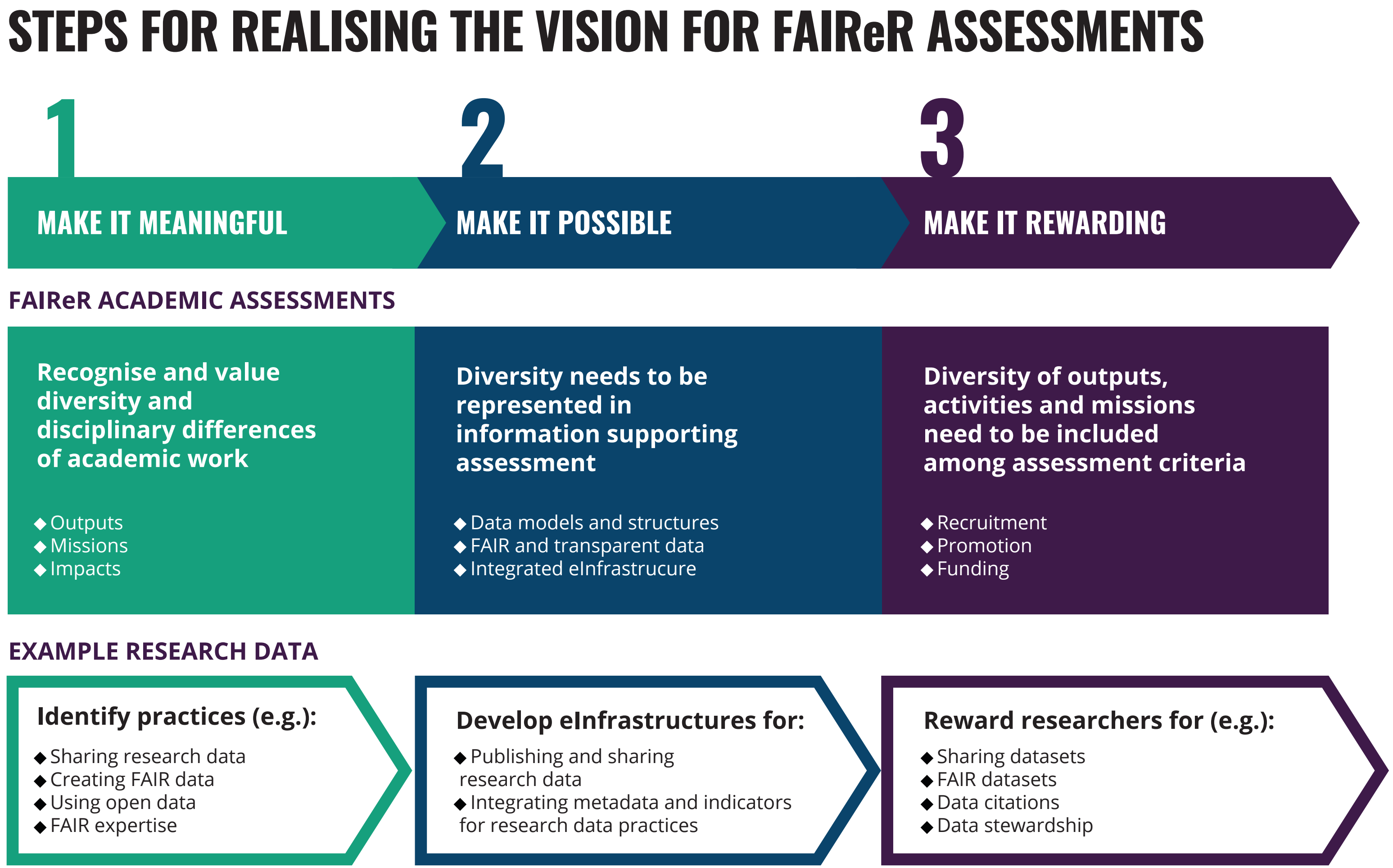


Figure 1. Steps for realising the vision for FAIReR assessments.

TABLE OF CONTENTS

INTRODUCTION

Open science, digitalisation, interdisciplinarity and internationalisation change the production, dissemination, impact and accountability of academic work. European institutions face increased global competition for positions and funding; growing numbers of academic personnel and students; as well as underfunding challenges. These changes must be reflected in the future academic assessment practices.

Researchers, funders and policy makers are also working to encourage open science, including the sharing and reuse of research data [1, 2]. Making data FAIR, or *findable, accessible, interoperable* and *reusable*, is seen as a key for enabling this shift to open science [3]. At the same time, proponents of *responsible academic assessments* are calling for a move away from entrenched, metrics-based systems to one in which research, evaluation processes and indicators are made transparent and held accountable [4,5].

We argue that open science can be encouraged and rewarded by developing FAIReR assessments recognizing open science outputs and activities. This requires a variety of stakeholders - research communities, policy makers, funders and publishers - to work together to address social and cultural barriers and challenges. It also requires creating a technical infrastructure, which makes responsible assessments of open science practices and outputs possible.

In this report, we focus particularly on two aspects needed to make assessing open science practices both rewarding and technically possible. These were addressed through the course of two intertwined EOSC Co-Creation projects. The first project, *European Overview of Career Merit Systems,* aimed to survey the existing landscape of the policies, technical systems and data models used in researcher evaluations and to evaluate how these systems support (or do not support) the responsible assessment of open science in academic careers. The second project *Vision for Research Data in Research Careers* aimed to understand the current state of assessing open science and data practices and to co-create a vision and roadmap for how these practices can be responsibly taken into account in academic careers.

We discuss the findings and outcomes from this work here, drawing on work conducted during an intense seven-month period of research and co-creation. During this time, we (i) performed an extensive review of policy documents, reports, manifestos for responsible assessments and the academic literature and (ii) conducted a survey and detailed case studies of five infrastructures. These studies were used to create the overview of the current state of practices and infrastructures presented in Section 1 (Overview) of this report. We also (iii) designed, moderated and engaged in four co-creation bootcamps with experts in open science, research data and academic assessments. We drew on the insights developed during the course of these bootcamps to propose the vision and roadmap, validated through a round of open public consultation, which compose the final two sections (Vision and Roadmap) of this report.[1]

1 A detailed description of our methodologies, as well as more detailed analyses, are included in Annex 1-7.

1 OVERVIEW

A fully-referenced report is available in Annex 2: Document review - Current state of responsible assessment of open science and research data.

The purpose of the overview of career-merit systems is to understand the current state of policies and information related to researcher assessment, and specifically how FAIR data figures in the picture. The overview is based on two kind of analyses:

1 Review of policy documents, surveys, reports and research related to incentives and rewards.

2 Analyses and comparisons of the data models and information systems for collecting publication and career merit information, consisting of:

> **A.** Online Survey on academic assessment systems distributed to European institutions.
>
> **B.** Five in-depth OS-CAM case-studies highlighting the key challenges for creating a European FAIR data environment for career assessment.

1.1. DOCUMENT REVIEW: CURRENT STATE OF RESPONSIBLE ASSESSMENT OF OPEN SCIENCE AND RESEARCH DATA

Key takeaways

- Guidelines for responsible assessments have common themes which need to be incorporated into FAIReR assessments: diversity, transparency, reflexivity and robustness of data and metrics. Academic assessments and practices of open science are situated within communities.

- European policy-makers and institutions are committed to encouraging and rewarding open science practices, including sharing research data. The move away from a narrow focus on research, publications and metrics towards a broader range of assessment criteria remains limited. Open science is also narrowly conceptualised, referring often only to open access publishing.

- The FAIR principles are often invoked. Measuring levels of FAIRness is challenging and perhaps not appropriate for judging individuals. Altmetrics for data and data citation are also in early stages; qualitative approaches for assessing open science offer promise.

- Information and data produced by researchers, institutions and infrastructures remains scattered and difficult to (re)use in assessments. There is momentum for developing an interlinked infrastructure based on persistent identifiers which integrates research entities and facilitates interoperability between research information systems.

CURRENT STATE OF PRACTICES

Common themes in responsible assessment

Responsible assessments are rooted both in anti-discrimination and equality legislation and in ethical guidelines for the responsible conduct of research and evaluation. Such guidelines converge on the themes of *diversity* (of research activities/outputs, disciplinary norms and professional roles), *transparency* (of systems and indicators) and *reflexivity* (about indicators and practices). They also highlight the need for robust, comparable data which can be used to calculate indicators to use in informing a variety of assessments.

Communities are also recognized as playing key roles in establishing understanding, trust and commitment in responsible assessment practices, as well as in determining (open) data and research practices.

Well-supported but not fully adopted; publications remain the norm
Numerous policies, articles and reports call for embedding open science criteria within processes for recruitment, career advancement and funding decisions. Alternative rewards and incentives, e.g., badges in journals for data sharing and management, have also been proposed by different stakeholders to further encourage open science activities. Despite this recognition, criteria and rewards for open science have not yet been fully adopted in practice.

Recognising a diversity of open science outputs and practices, as described in the Open Science Career Assessment Matrix (OS-CAM) [6], has also not materialised in practice. Publications, metrics based on journals and personal qualifications (e.g., skills and expertise specified on academic CVs) continue to be de rigour in assessing research and academic work. Existing CV portfolio templates, which can be used to simplify assessment processes, also only partially support a diversity of research outputs and practices.

Open science equates to open access
This lack of diversity is reflected in how open science is represented in many policies and practices, where *open science* is often equated with *open access* publishing. Guidelines and policies for encouraging open access publications are also more fully developed and adopted than policies encouraging sharing (open) data or publishing other research outputs. Many researchers' knowledge of open science skills is limited to making publications and data openly available.

Policies for research data also hone in on one aspect of data practices, namely data sharing and management, e.g., by requiring the creation of data management plans (DMPs). Such policies are not routinely enforced; nor are data sharing and management regularly considered in academic assessments. We find that other data practices, such as reusing data, are rarely incentivised.

Challenges remain for measuring FAIR data
The FAIR principles for data management are also taken up in many documents, although open science experts and researchers may not have the same view about the adoption of the FAIR principles. The FAIR principles apply not just to data but to other research objects, such as software and protocols. There is also some recognition that FAIRness is not a binary, either-or concept but rather, one that exists on a spectrum.

Multiple stakeholders see a need to develop metrics to indicate the level of "FAIR-ness" of data, or measuring how data align with the FAIR principles. Numerous efforts are underway to manually or automatically evaluate the FAIRness of data, although there are also clear messages that measures of FAIRness should not be used to make judgements about individuals. Assessing FAIRness is complex, as creating FAIR data involves the work of many entities, e.g., repositories, researchers, and standards developers. Assessments of FAIRness should go beyond the data in question and extend to examinations of other elements of the scholarly ecosystem, including frameworks and infrastructures.

Recurring theme: Alternative indicators and qualitative approaches in assessment
Although altmetrics (e.g., tweets and downloads) offer an alternative to publication-based metrics, they are rarely incorporated into researcher assessments, perhaps because of known challenges, such as accuracy and potential manipulation. Using customised metrics and citations for other types of research output, such as to data and software, is increasingly encouraged. Standards for data citation in particular are being created in multiple initiatives, although data citation infrastructures and practices are still in a state of development.

There is a danger that new metrics and quantitative indicators may be misapplied and misinterpreted. Qualitative approaches to assessment, e.g., narrative CVs or descriptions of research activities, are often called for as a way to counter this possibility. Despite a general interest and trend towards integrating qualitative methods into academic assessments, such approaches are not often used to assess open science or research data.

CURRENT STATE OF ASSESSMENT INFRASTRUCTURE

An open question: Do research information systems support assessments recognising open science?
Current research information systems (CRIS), which capture information about a wide range of research outputs and activities, and models and platforms for preserving digital research objects are commonplace. The information recorded in CRIS systems can be used to easily calculate many traditional, publication-based metrics; less is known about how these systems capture activities related specifically to open science and research data.

While CRIS-derived publication metrics can be easily integrated into other webpages and environments, they are also limited by the robustness and accuracy of the information initially recorded. Research information systems are often either locally created and managed or are outsourced to commercial providers. Local systems face challenges of interoperating with other research information systems due to differing local standards.

Trending: Interlinked infrastructures built on persistent identifiers
There is a trend toward integrating isolated systems and linking entities that exist in different locations, e.g., linking research data in repositories with CRIS systems; linking publications to researcher portfolios; or linking researchers to institutions. These linkages usually rely on persistent identifiers (PIDs) to uniquely identify entities of interest. While PIDs exist for many entity types, e.g., researchers, institutions and objects, there are also entities for which no PID exists.

There are many existing efforts which could be built on to create an interlinked scholarly infrastructure using PIDs and many research information models which describe research data and other digital research objects. The widely-adopted CERIF format [7], e.g., for representing research information has been applied to various research domain objects, including indicators and metrics used in academic assessments. A range of registries exist for data, publications, guidelines, policies and standards although there is a notable lack of a registry specifically for indicators and metrics used in assessment.

Recognition that technologies need people
Assessment and data infrastructures must be built, maintained, curated and adapted to meet the needs of various communities. Training and technical support, provided by data stewards, curators or research software engineers, are also necessary roles, which should be recognised and rewarded.

1.2. SURVEY AND CASE-STUDIES ON OPEN SCIENCE CRITERIA IN DATA MODELS AND INFORMATION SYSTEMS

While European research performing and funding organisations are strongly committed to open science policies, relatively little practical progress has been made in recognizing Open Science in career assessment. The Open Science Career Assessment Matrix (OS-CAM) [6] provides a well-established framework with examples of 42 possible assessment criteria across six main areas of Open Science activities: research output, research process, service and leadership, research impact, teaching and supervision, and professional experience. We conducted an online survey and five in-depth case-studies to investigate to what extent the assessment infrastructures at European institutions make possible capturing and supporting assessments based on diversity of open science outputs and practices.

Key takeaways

- Institutions support assessments with qualitative and quantitative data from various local einfrastructres (typically institutional or national CRIS - current research information systems) and global platforms (ORCID, Web of Science, Scopus, etc).

- Rewarding researchers for Open Science is currently lacking reliable, comprehensive, well-structured and comparable qualitative and quantitative data and metrics about most Open Science outputs and practices.

- The only Open Science activity, which is well covered in the current assessment infrastructures, are research publications, notably publishing in open access journals and self-archiving. There is a need to develop local and global information systems, data models, terms and vocabularies, as well as consistent use of PIDs, for capturing the diversity of Open Science activities.

- The OS-CAM criteria are often very difficult to apply in practice. A more detailed description of the OS-CAM criteria is also needed.

- Open Science infrastructure for integrating, interoperating and sharing (according to FAIR principles) the Open Science career assessment data from international, national and institutional research information systems, databases and data models is needed.

Survey for Academic Assessment Systems

Firstly, an online survey to research performing and funding organisations was carried out to understand how current academic assessment systems and infrastructures support reporting and evaluating open science practices using a selection of OS-CAM criteria (a full report of survey study is provided in Annex 3). We received 24 complete or mostly complete responses from institutions in 16 European countries, and one institution in Australia. The main findings of survey study are:

- Local eInfrastructures used for assessments offer only very limited support for recording information about the diverse Open Science outputs and activities. Open access publications and self-archiving of publications is a relatively well-covered area; however, other aspects of open science activities and outputs require much further development of information systems and platforms.

- Web of Science and Scopus remain the major information sources and platforms (in addition to local eInfrastructures) to support research assessment, despite the criticism related to their lack of coverage of diversity of research outputs.
- Harvesting data from global resources and infrastructures suffers from missing and/or inconsistent use of PIDs for some entities, and requires additional effort by librarians and other personnel to check, consolidate and enrich information.
- Data used in the academic assessment and results of assessment are rarely made open, and even if they are, that is usually related to the group assessment.

OS-CAM Case Studies

Secondly, case-studies were carried out to analyse in-depth to what extent the 42 OS-CAM criteria are covered in selected well-established international and national research information platforms (a full report of case-studies is provided in Annex 4). It was specifically analysed if the fulfilment of each of the OS-CAM criteria could be automatically checked and/or requires manual work from the evaluator in case of CERIF (International), ORCID (international), Research.fi (Finland), NARCIS (the Netherlands) and CRIStin (Norway). The main findings of the case-studies are:

- The OS-CAM criteria are often very difficult to apply in practice. The OS-CAM criteria should be described in much more detail, and they should be joined with a set of examples about the evidence of fulfilling those criteria.
- The OS-CAM criteria related to the well-established and popular ways of acquiring academic merit such as journal articles and other traditional research outputs are much better covered by analysed infrastructures than more holistic ones that include for example social impact, teaching and professional experience related information.
- Missing open science related terms and vocabularies are the main obstacle to supporting the OS-CAM criteria in the reviewed assessment infrastructures.

2 VISION

Relevant annexes: Annex 1: Vision for FAIReR assessments, Annex 5: Bootcamp methodology and overview and Annex 6: Public consultation for vision and roadmap.

2.1. CO-CREATING FAIRER ASSESSMENT CULTURE WITH COMMUNITIES

Our research - the policy and document review, survey and bootcamps - emphasises a need for engaging communities in researcher career assessments. We need to include a broad range of outputs, activities and impacts of academic work in criteria for hiring, promotion and funding. The FAIReR assessment of researcher careers takes into consideration this diversity and fosters community ownership and trust.

Communities
Communities are not defined alone by disciplinary domains. They are formed within disciplines, across organisational and national boundaries; and within organisations and nations, yet bridging disciplinary boundaries. Determining what is valued in researcher careers requires dialogue both within and between such communities.

Engaging in dialogue with research communities enables the creation of a FAIReR assessment culture that takes into consideration the diversity of practices, outputs, missions and impacts of academic work, and differences between fields. This co-creative process is also a way to foster community ownership and trust. It allows communities to develop ownership in the policies and criteria used to assess their work; co-created assessment criteria are owned by the community for the community.

Research communities should not only have ownership of assessment processes and criteria, but they should also have a sense of ownership about the data used in assessments. Community-owned assessment infrastructures support the curation of data for FAIReR assessments, and make data as open as possible and reusable via open infrastructures. Community-owned infrastructure can integrate data on research activities as well as analytical services and tools provided by commercial parties. Ownership of assessment criteria, processes, data and infrastructures further enhances trust between evaluators and those being evaluated, ensuring that assessments benefit all parties involved. Trust is essential for the cultural change required for creating FAIReR assessment culture.

Elements of FAIReR assessment culture

1. Diversity of practices and communities

FAIReR research career assessments are field-specific and community-defined. Notions of openness, quality and data are grounded and differently defined within research cultures because:

A. Research is embedded within cultural, linguistic, social, economic, and political contexts.
B. Practices of finding, accessing, integrating and reusing data, as well as practices of data description and sharing, vary both between and within disciplinary fields.
C. Research performing and funding organisations have specific strategic priorities, diverse values and missions.
D. Assessments are carried out at a variety of levels (e.g., for institutions, research units and individual researchers) and for a variety of purposes (e.g., funding allocation, organisational rankings, promotion, hiring and awarding academic degrees).

2. Transparency of assessments
FAIReR research career assessments are transparent and open. In order to make the evaluation results verifiable, assessments need to be open interms of:

A. Policies
B. Assessment criteria
C. Assessment methods (including indicators and infrastructure)
D. Data used in assessments

3. Reflexivity of processes
FAIReR research career assessment processes are continually evaluated together with research communities. Co-creation is a mutual and reflexive learning experience that provides a way for assessment policies to ad just to changes in research environments and be adapted to local contexts.

2.2. ASSESSMENT INFRASTRUCTURE FOR FAIRER ACADEMIC ASSESSMENT

Assessment infrastructure is an essential non-profit open science service, which is governed and owned by the community, and funded collectively by governments, as well as research funding and performing institutions. An interlinked FAIReR assessment infrastructure is essential to capture the full diversity of research information. An infrastructure is needed for integrating qualitative and quantitative data from, and facilitating interoperability between, international, national and institutional research information systems and infrastructures.

Comprehensiveness
Open and FAIR assessment infrastructure should be developed in order to make FAIReR assessments possible. Rewarding researchers for diverse open science practices requires reliable, comprehensive, well-structured and comparable data and metrics to support the assessment process. In the case of research data, information on creating, publishing and sharing research data needs to be reliable, comprehensive, comparable and structured.

Interlinking of research information
Information produced by researchers, institutions and infrastructures remains scattered and difficult to use and reuse in assessments. An infrastructure for integrating qualitative and quantitative data from, and facilitating interoperability

between, international, national and institutional research information systems and infrastructures is needed. The FAIReR assessment eInfrastructure might represent a comprehensive global research information ecosystem built by using PIDs and semantic web technologies.

Openness

The infrastructure should be principally built on community-owned, community-curated and openly available data on research. The infrastructure built on top of openly available data through GUI and API facilitates reusing research information produced locally. Those openly available data besides input data for the assessment process should also include data describing the process of assessment and the result (output) of the assessment process. Therefore, the infrastructure should enable describing and publishing assessment criteria and indicators in machine-actionable format in accordance with FAIR principles. Moreover, publishing of academic assessment results should be a balance between needs for openness and transparency of assessment process on one side, and researchers' privacy on the other side.

The eInfrastructure architecture

Relevant annex: Annex 7: Vision of the FAIReR assessment eInfrastructure

Figure 2 (p. 19) presents a vision for developing FAIReR assessment infrastructure. Green rectangles are used for already existing platforms/services which can be used for building this research infrastructure ecosystem. Although those platforms/services already exist across the world, they should be maintained and further extended in accordance with the open science paradigm. Purple cloud in the middle of the diagram represents integration of all those services and platforms under one umbrella. It will improve visibility/discoverability of platforms/services, and on the other side it will enable collecting a comprehensive list of achievements of a researcher or group. Moreover, three more purple rectangles should be developed as a part of this vision with the goal of making a basis for building local Academic assessment platforms (dark blue rectangle) which will support responsible academic assessment. New elements of the envisioned architecture are briefly described below, while the more detailed description of all elements of the architecture might be found in the Annex 7.

TECHNICAL VISION OF THE FAIReR ASSESSMENT eINFRASTRUCTURE

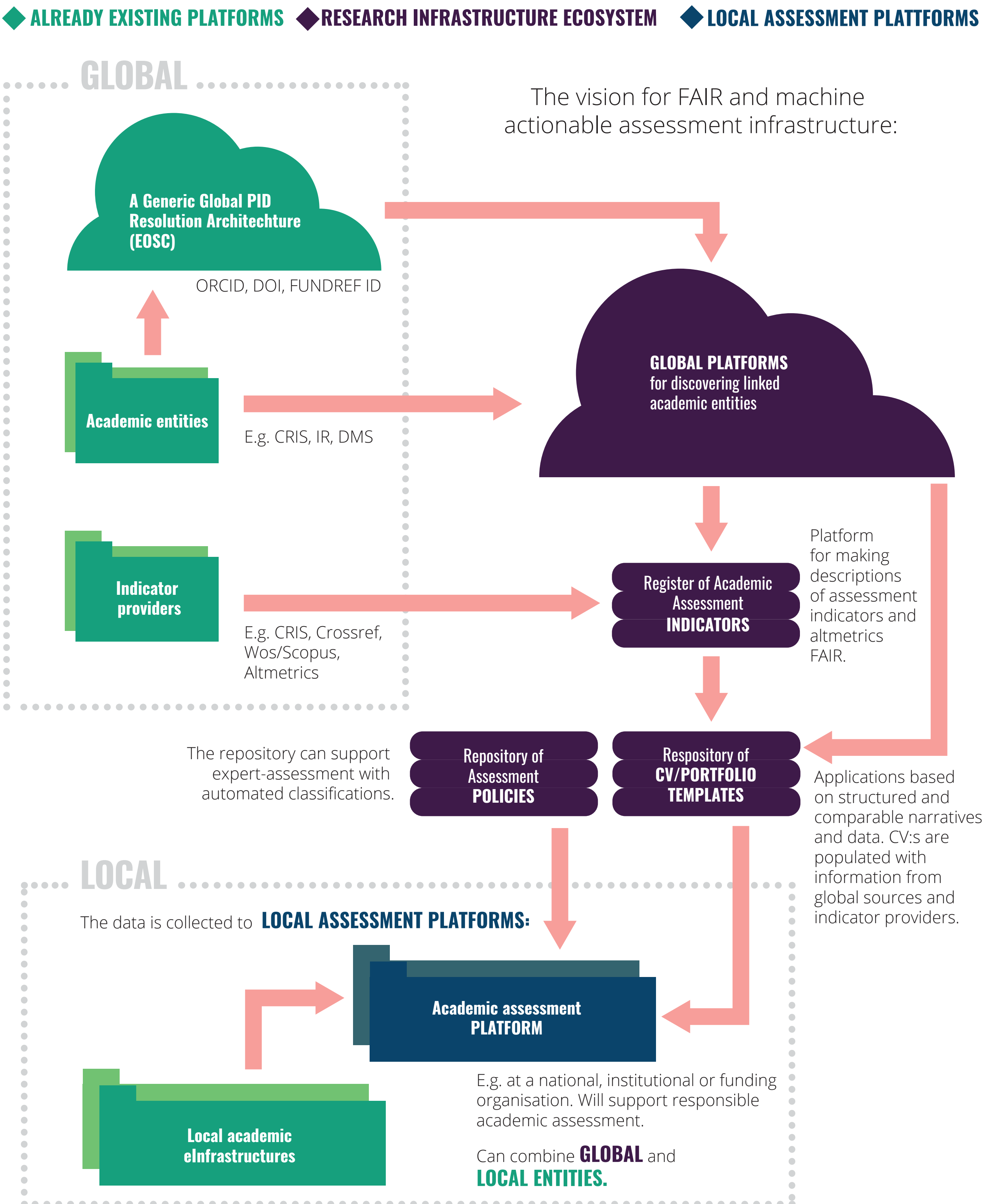


Figure 2. eInfrastructure architecture for responsible academic assessment

◆ **REGISTER OF ACADEMIC ASSESSMENT INDICATORS** is a platform for making descriptions of assessment indicators and (alt)metrics FAIR. The description should also include a list of indicator providers, e.g., platforms which might be used to calculate the value of an indicator (national and institutional CRIS, Crossref, Dimension, Microsoft Academic, WoS, Scopus, etc).

◆ **REPOSITORY OF CV/PORTFOLIO TEMPLATES** - CVs/Profiles Templates shape all applications in the same format. Machine-actionable CVs/Profiles Templates can be integrated with other infrastructure elements presented at Figure 2. Repository should store CVs/Profiles Templates represented by using machine executable instructions for collecting and formatting data for assessment. Templates represented in this way enable making researchers' CVs/Portfolios by using those templates and collecting data from linked academic entities' platforms. Besides CVs/Profiles Templates in machine readable format, the repository preserves templates' descriptions in a rich metadata format and makes templates FAIR.

◆ **REPOSITORY OF ACADEMIC ASSESSMENT POLICIES** - Some academic assessment criteria can be represented in machine executable format. Machine executable academic assessment criteria can automatically produce final classification based on complex rules built on top of input data provided by evaluators, applicants or indicator providers. Repository of academic assessment policies enables storage, discovering and execution of policies and its criteria. Besides academic assessment policies in machine executable format, the repository preserves pdf files and its descriptions in a rich metadata format making policies FAIR.

◆ **ACADEMIC ASSESSMENT PLATFORMS** should centralise academic assessment processes in one institution or region, and make those processes more efficient, fair and transparent. It might be used by applicants and evaluators. Moreover, it should be connected with "Repository of CVs/Profiles Templates" through its API for the purpose of collecting information available in global platforms and for uniform formatting of data, in machine and human readable format, which should be assessed. Applicants can extend those collected information from global sources with qualitative narratives and case reports of their research activities and impact. Furthermore, it should be connected with "Repository of academic assessment policies" for partial automatization of the process of writing assessment reports.

The eInfrastructure characteristics

The proposed eInfrastructure minimises assessment data re-entry and maximises assessment data re-use by making assessments' indicators, CV templates and policies FAIR. Therefore, the complete assessment process can be more transparent and enable the fair treatment for all applicants. Although it is optional due to privacy issues, it is also possible to store CVs/Portfolios in "Repository of CV/Portfolio templates" and results of assessment in the "Repository of academic assessment policies" and make the academic assessment process completely transparent. Moreover, the eInfrastructure might decrease the efforts of applicants and evaluators in the process of application and assessment. The proposed eInfrastructure supports a mix approach - quantitative and qualitative measures, usage of narrative evaluations and indicators. The architecture supports including plural characteristics and opening up the range of contributions in the assessment process for the purpose of objectivity and integrity. Moreover, the proposed eInfrastructure can be used for experimenting with novel approaches, evaluating assessment processes, or even for cyclical and iterative assessment.

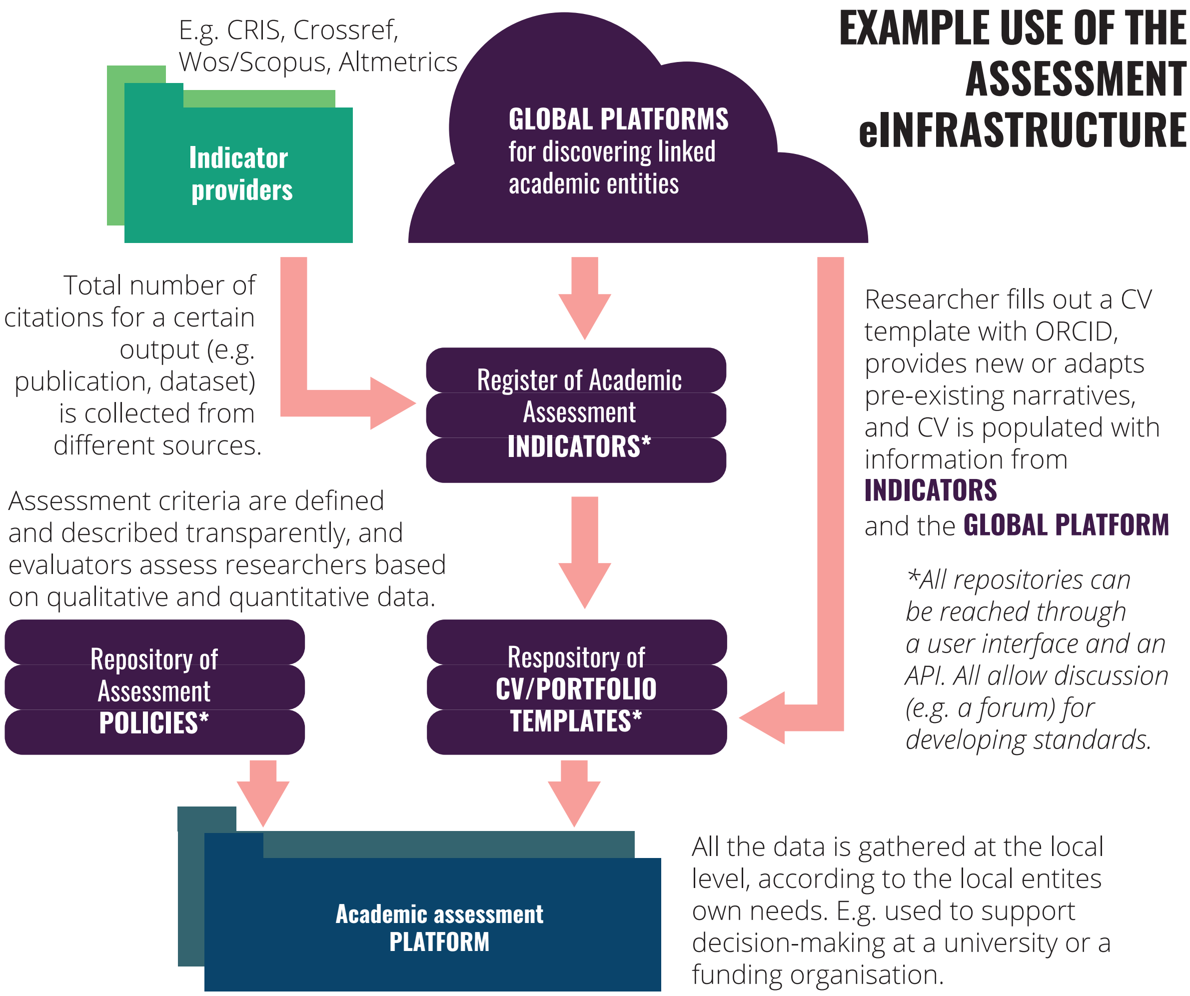


Figure 3. An example case of how the eInfrastructure could be used.

3 ROADMAP

Changes in academic assessment culture shake the research community at its core. Decisions around assessment define what is held important, valuable and where the research community wishes to go from here. Although “what” and “why” are important, nothing will happen without a “how”.

This change is going to take time and requires significant shared effort and investment in order to become reality. When moving from vision to reality, we draw on our own research - policy review, survey - as well as co-creation process of bootcamps with experts in the areas of open science, research data and research evaluations.

3.1. BARRIERS FOR FAIRER ASSESSMENT CULTURE

If creating a FAIReR assessment culture were easy, it would have been done already. The changes in assessment culture meet challenges, which must be overcome. Awareness of barriers supports developing a roadmap from vision to reality. Barriers to this change include:

A. Cost in developing infrastructure

B. Lack of resources for qualitative assessment

C. Limitations in expertise in responsible assessment

D. Inconsistencies in policies to regulate open science, data management and assessments

E. Cultural and practical investment in the current assessment and evaluation system

Given the scale of these challenges, a multi-stakeholder program - including government agencies, research funding and performing organisations, open science and research assessment and management communities, public and commercial service providers - is critical to achieve a shared global view.

3.2. PRIORITIES FOR ACTION

Policy collaboration

Because research and scholarship are international, there is a need for a global and shared vision for assessment. Shared vision will benefit all stakeholders. Countries and organisations make up the community, but none of them can alone change the international research assessment culture. The change requires simultaneous international and local policy development.

One of the challenges to implementation could be the range of guidelines and policies which have been developed to encourage and regulate open science, data management and assessment at different levels. Such policies rarely align with each other, making it difficult for researchers and institutions to know which policy to follow when.

International policy development creates a space for local innovations to flourish, which again feed and support international policies.

Priorities for policy development:

A. International agreements and/or MoUs on FAIReR assessment vision and policies

B. Co-creation of national and organisational FAIReR assessment policies

C. Support and populate platforms of best practices for FAIReR assessments and policies for mutual learning

D. Adopting, and where necessary creating, shared taxonomies of research contributions and contributors in policies regarding FAIReR assessments

Investment in assessment data infrastructure

Changes require resources - in time, energy and financial investment. Change in assessment is no different. Policy development above requires investment in time and energy, while infrastructure requires considerable financial investments. To make FAIReR assessments possible, the research communities require new infrastructures for gathering, storing and sharing assessment data at institutional, national and international level.

Key to success is finding a balance between harmonisation and diversity.

Balancing the use of metrics and qualitative assessments is also a question of wise use of resources. The key to decreasing the role metrics in assessments is to make better use of the time and effort experts invest in assessments. The experts' work can also be facilitated by producing good and reliable data to support qualitative evaluations, for example by developing infrastructure and services for the production, use, storing and sharing of structured and guided narrative descriptions and case reports. By making expert assessments more open (e.g., open peer review), the outcomes of an assessment can be used more than once. The experts' workload in peer review needs to be balanced with merit and rewards for peer review work. Ideally, performing assessments should be integrated in a natural way in the process of doing research.

Priorities for FAIReR assessment infrastructure investments:

A. International funding call by EOSC, or other similar organisation, to begin developing the technical solution for shared assessment infrastructure and required data models

B. Building a FAIReR assessment infrastructure with mutually shared architecture supporting quantitative and qualitative assessment information (Annex 7), which

- Has shared data models and PIDs
- Builds on existing infrastructures
- Provides different levels of access for different types of users
- Incorporates consent from researchers
- Minimises assessment data re-entry and maximises assessment data re-use

C. Establishing an international forum for dialogue between professionals, including open science and evaluation experts, researchers, data stewards and research software engineers to support development of FAIReR assessment infrastructure.

REFERENCES

[1] Saenen, B., Morais, R., Gaillard, V., & Borrell-Damián, L. (2019). *Research Assessment in the Transition to Open Science*(p. 48). European University Association. https://eua.eu/downloads/publications/research%20assessment%20in%20the%20transition%20to%20open%20science.pdf

[2] Sveinsdottir, T., Proudman, V., & Davidson, J. (2020). *An Analysis of Open Science Policies in Europe, v6*. SPARC Europe; Digital Curation Centre. https://zenodo.org/record/4005612#.YBk-6eB7mXG

[3] Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., & others. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3. https://doi.org/10.1038/sdata.2016.18

[4] Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature News*, *520*(7548), 429. https://doi.org/10.1038/520429a

[5] Wilsdon, J., Allen, L., & Belfiore, E. (2015). The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management. https://responsiblemetrics.org/wp-content/uploads/2019/02/2015_metrictide.pdf

[6] O'Carroll, C., Rentier, B., Cabello Valdès, C., Esposito, F., Kaunismaa, E., Maas, K., … & Wauters, P. (2017). *Evaluation of research careers fully acknowledging open science practices-rewards, incentives and/or recognition for researchers practicing Open Science*. Publication Office of the European Union. https://op.europa.eu/en/publication-detail/-/publication/47a3a330-c9cb-11e7-8e69-01aa75ed71a1

[7] euroCRIS. (2017). CERIF Data-model. https://github.com/EuroCRIS/CERIF-Data-Model

ANNEX 1

VISION FOR FAIReR ASSESSMENTS

INTRODUCTION

Open science, digitalisation, interdisciplinarity and internationalisation change the production, dissemination, impact and accountability of academic work. European institutions face increased global competition for positions and funding; growing numbers of academic personnel and students; as well as underfunding challenges.[1] These changes must be reflected in the future academic assessment practices.

European policy-makers and institutions are strongly committed to encouraging and rewarding open science practices, including the sharing and reuse of research data.[2] Researchers need to be recognised for contributions to teaching and learning, innovation, culture and societal change. Yet the move away from a narrow focus on research, publications and metrics towards a broader range of assessment criteria remains limited.[3]

1 European University Association (2020). *Universities without walls: A vision for 2030*. [link]; For statistics on academic personnel, see *EUROSTAT*: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=R_%26_D_personnel#Researchers.

2 European Commission Working Group on Rewards under Open Science. (2017). *Evaluation of research careers fully acknowledging Open Science practices: Rewards, incentives and/or recognition for researchers practicing Open Science.* Publications Office of the European Union. https://data.europa.eu/doi/10.2777/75255; Wouters, P., Ràfols, I., Oancea, A., Kamerlin, L., Holbrook, J. & Jacob, M. (2019). *Indicator Frameworks for Fostering Open Knowledge Practices in Science and Scholarship*. European Commission. https://op.europa.eu/en/publication-detail/-/publication/b69944d4-01f3-11ea-8c1f-01aa75ed71a1#; Sveinsdottir, T., Proudman, V., & Davidson, J. (2020). *An Analysis of Open Science Policies in Europe*, v6. SPARC Europe; Digital Curation Centre. https://zenodo.org/record/4005612#.YBk-6eB7mXG.

3 Saenen, B., Morais, R., Gaillard, V., & Borrell-Damián, L. (2019). *Research Assessment in the Transition to Open Science*. European University Association. https://eua.eu/downloads/publications/research%20assessment%20in%20the%20transition%20to%20open%20science.pdf; European Commission (2020) *Six Recommendations for implementation of FAIR practice by the FAIR in practice task force of the European open science cloud FAIR working group*. https://ec.europa.eu/info/sites/info/files/research_and_innovation/ki0120580enn.pdf

Europe needs a vision for FAIReR Assessments built on the *FAIR principles for data management* and policies guiding *Responsible assessment* (FAIReR = FAIR + Responsible). The FAIR principles, guidelines for making data *findable* (F), *accessible* (A), *interoperable* (I) and *reusable* (R), are key to enabling a shift to open science.[4] In FAIReR Assessments, research data, as well as the criteria, data and metrics informing assessments, are transparent and FAIR.[5]

Diversity is the guiding theme throughout this vision for FAIReR Assessments. Diversity in this context means recognising different outputs, roles and impacts of academic work, and respecting differences between fields. Starting with the DORA declaration (https://sfdora.org/), several international statements outline guiding principles for responsible research assessment methods, criteria and data.[6]

FAIReR Assessments build on principles of community governance, co-creation, co-curation and dialogue. Responsible assessments are also rooted in legislation regarding, for example, equality, anti-discrimination and data protection,[7] and in ethical guidelines for the responsible conduct of research and evaluation.[8] Re-

4 Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., & others. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3. https://doi.org/10.1038/sdata.2016.18

5 Science Europe. (2016). *Position Statement on Research Information Systems*. https://www.scienceeurope.org/our-resources/position-statement-on-research-information-systems/

6 Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature News*, 520(7548), 429. https://doi.org/10.1038/520429a; Wilsdon, J., Allen, L., & Belfiore, E. (2015). *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. https://responsiblemetrics.org/wp-content/uploads/2019/02/2015_metrictide.pdf; Curry, S., de Rijcke, S., Hatch, A., Pillay, D., van der Weijden, I. & Wilsdon, J. (2020). *The changing role of funders in responsible research assessment: progress, obstacles and the way ahead*. Research on Research Institute. https://doi.org/10.6084/m9.figshare.13227914.v1.

7 E.g. McCrudden, C., & Prechal, S. (2009). *The Concepts of Equality and Non-Discrimination in Europe: A Practical Approach*. European Commission. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjh1cyVkorwAhVE_SoKHfoTAWEQFjAAegQIBhAD&url=https%3A%2F%2Fec.europa.eu%2Fsocial%2FBlobServlet%3FdocId%3D4553&usg=AOvVaw2gKC7lo9sECSstYKKiZNKP; Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with regard to the Processing of Personal Data and on the Free Movement of Such Data, and repealing Directive 95/46/EC (General Data Protection Regulation), 88 (2012). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679.

8 Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Har Sham, M., Barbour, V., Coriat, A.-N., Foeger, N. & Dirnagl, U. (2020). The Hong Kong Principles for assessing researchers: Fostering re-

search communities have played a key role in establishing understanding, trust and commitment in responsible assessment practices at institutional and national levels.[9]

Vision for FAIReR Assessments includes development of an open and FAIR assessment infrastructure. Rewarding researchers for diverse open science practices requires reliable, comprehensive, well-structured and comparable data and metrics to inform assessments. Information produced by researchers, institutions and infrastructures remains scattered and difficult to use and reuse in assessments. An infrastructure for integrating qualitative and quantitative data from, and facilitating interoperability between, international, national and institutional research information systems and infrastructures is needed.[10]

The aim of our EOSC project is to co-create a common vision for FAIReR assess-

search integrity. *PLOS Biology*. **https://doi.org/10.1371/journal.pbio.3000737**; Biagetti, M. T., Gedutis, A., & Ma, L. (2020). Ethical Theories in Research Evaluation: An Exploratory Approach. *Scholarly Assessment Reports*, 2(1), 11. **https://doi.org/10.29024/sar.19**; Mustajoki, H., & Mustajoki, A. (2017). *A New Approach to Research Ethics: Using Guided Dialogue to Strengthen Research Communities*. Routledge. **https://doi.org/10.4324/9781315545318**.

9 Saenen, B., Hatch, A., Curry, S., Proudman, V. & Lakoduk, A. (2021). *Reimagining Academic Career Assessment: Stories of innovation and change* **https://eua.eu/downloads/publications/eua-dora-sparc_case%20study%20report.pdf**; VSNU, NFU, KNAW, NWO & ZonMw (2020). *Room for everyone's talent: towards a new balance in the recognition and rewards for academics*. **[link]**; Working group for the responsible evaluation of a researcher (2020). *Good practice in researcher evaluation. Recommendation for the responsible evaluation of a researcher in Finland*. The Committee for Public Information (TJNK) and Federation of Finnish Learned Societies (TSV). **https://doi.org/10.23847/isbn.9789525995282**; Himanen, L. & Gadd, L. (2019). Introducing SCOPE – a process for evaluating responsibly. *The Bibliomagician*. **https://thebibliomagician.wordpress.com/2019/12/11/introducing-scope-aprocess-for-evaluating-responsibly/**; COST Action ENRESSH (2017). *Challenges of the evaluation of social sciences and humanities research (SSH)*. **https://enressh.eu/wp-content/uploads/2017/05/Evaluation_of_SSH_final.pdf**.

10 SPARC Europe (2020). *Scoping the Open Science Infrastructure Landscape in Europe*. **https://zenodo.org/record/4159838#.YDYYcmgzZ3i**; Waltman, L. (2019). *Open Metadata of Scholarly Publications: Open Science Monitor Case Study*. European Commission. **https://doi.org/10.2777/132318**; Puuska, H.-M., Guns, R., Pölönen, J., Sivertsen, G., Mañana-Rodríguez, J., & Engels, T. (2018). *Proof of concept of a European database for social sciences and humanities publications: Description of the VIRTA-ENRESSH pilot*. CSC & ENRESSH. **https://doi.org/10.6084/m9.figshare.5993506**; Mahieu, B., Arnold, E., & Kolarz, P. (2014). *Measuring Scientific Performance for Improved Policy Making*. European Parliamentary Research Service. **http://www.europarl.europa.eu/RegData/etudes/etudes/join/2014/527383/IPOL-JOIN_ET(2014)527383(SUM01)_EN.pdf**; European Commission. (2010). *Assessing Europe's University-Based Research*. **https://web.archive.org/web/20190317171523/http://ec.europa.eu/research/science-society/document_library/pdf_06/assessing-europe-university-based-research_en.pdf**

ments to make rewarding open science practices possible.[11] The focus of this vision is on assessing individuals for purposes of hiring, promotion, funding, but the vision for FAIReR Assessments is relevant also for research groups and institutions. We recognise that assessments themselves cannot be standardised across Europe, due to diverse institutional needs, strategic goals, and disciplinary standards.

II VISION FOR FAIRER ASSESSMENTS

FAIReR assessments are rooted in both the FAIR guidelines for data management and policies for the responsible assessment of research. Specifically, FAIReR assessments foreground *diversity, communities*, and *dialogue*.

IN ORDER FOR FAIRER ASSESSMENTS TO BE REALISED:

1. Communities co-create the meaning of diversity in assessments
Make it meaningful. We need to know what we want to value and evaluate. To do this, we start by considering the goals of open science and do not limit our evaluations to what is technically possible or easy to measure. We take into consideration the diversity of practices, outputs, missions and impacts of academic work, and differences between fields. In the case of research data, such practices may include sharing (open) datasets, creating FAIR datasets, reusing data, or cultivating expertise in creating or curating FAIR data.

11 Nosek, B. (2019). Strategy for Culture Change. *Center for Open Science Blog*. https://www.cos.io/blog/strategy-for-culture-change

2. Communities build assessments on infrastructures capturing diversity
Make it possible. We need to make it possible for researchers to report, make visible, and explain their diverse outputs, activities and impact of their work. Integration of relevant information from different sources is facilitated by open assessment infrastructure. In the case of research data, information on creating, publishing and sharing research data needs to be reliable, comprehensive, comparable and structured.

3. Communities reward diverse open science and FAIR practices
Make it rewarding. We need to include a broad range of outputs, activities and impacts of academic works in criteria for hiring, promotion and funding. In the case of research data, this may include shared or open data, indications of data reuse, or acting as data steward.

STEPS FOR REALISING THE VISION FOR FAIReR ASSESSMENTS

1 MAKE IT MEANINGFUL	2 MAKE IT POSSIBLE	3 MAKE IT REWARDING
FAIReR ACADEMIC ASSESSMENTS		
Recognise and value diversity and disciplinary differences of academic work	**Diversity needs to be represented in information supporting assessment**	**Diversity of outputs, activities and missions need to be included among assessment criteria**
◆ Outputs ◆ Missions ◆ Impacts	◆ Data models and structures ◆ FAIR and transparent data ◆ Integrated eInfrastrucure	◆ Recruitment ◆ Promotion ◆ Funding
EXAMPLE RESEARCH DATA		
Identify practices (e.g.): ◆ Sharing research data ◆ Creating FAIR data ◆ Using open data ◆ FAIR expertise	**Develop eInfrastructures for:** ◆ Publishing and sharing research data ◆ Integrating metadata and indicators for research data practices	**Reward researchers for (e.g.):** ◆ Sharing datasets ◆ FAIR datasets ◆ Data citations ◆ Data stewardship

1. COMMUNITIES CO-CREATE MEANING OF DIVERSITY IN ASSESSMENTS

Creating a FAIReR assessment culture requires understanding and accounting for the diversity of both research practices and communities. It also requires co-creating assessment criteria, methods and practices in conjunction with research communities to foster ownership and trust.

Diversity within communities
Disciplinary domains alone do not define research communities. Communities form both within disciplines, across organisational and national boundaries; and within organisations and nations, yet bridging disciplinary boundaries. Co-creation requires engaging in an ongoing dialogue both within and between such communities.

Notions of openness, quality and data are grounded and differently defined within research cultures because:

A. Research is embedded within cultural, linguistic, social, economic, and political contexts.

B. Practices of using, finding, accessing, integrating and reusing data, as well as practices of data description and sharing, vary both between and within disciplinary fields

C. Research performing and funding organisations have specific strategic priorities, diverse values and missions.

D. Assessments are carried out at a variety of levels (e.g., for institutions, research units and individual researchers) and for a variety of purposes (e.g., funding allocation, organisational rankings, promotion, hiring and awarding academic degrees).

Co-creation as a way to identify diversity and foster community ownership and trust
Co-creation in general is a mutual and reflexive learning experience. Engaging in dialogue with research communities is a way to identify the diversity of practices and norms that need to be considered in FAIReR assessments. These dialogues also provide a way for assessment policies to adjust to changes in research environments and be adapted to local contexts. At the same time, co-creative pro-

cesses allow communities to develop ownership in the policies and criteria used to assess their work; co-created assessment criteria are owned by the community for the community.

Research communities should not only have ownership of assessment processes and criteria, but they should also have a sense of ownership about data used in assessments. Assessment infrastructure is an essential non-profit Open Science service, which is governed and owned by the community, and funded collectively by governments, as well as research funding and performing institutions. Community-owned assessment infrastructures support the curation of data for FAIReR assessments, and make data as open as possible and reusable via open infrastructures. Community-owned infrastructure can integrate data on research activities as well as analytical services and tools provided by commercial parties.

Ownership of assessment criteria, processes, data and infrastructures further enhances trust between evaluators and those being evaluated, ensuring that assessments benefit all parties involved. Trust is also essential for creating the cultural change required for creating FAIReR assessment culture.

2. COMMUNITIES BUILD ASSESSMENTS ON INFRASTRUCTURES CAPTURING DIVERSITY

An interlinked infrastructure supporting FAIReR assessments
An interlinked FAIReR assessment infrastructure is created to capture the full diversity of research information. The infrastructure provides ways to record quantitative and qualitative information about the diversity of outputs, activities and roles involved in academic work. Using PIDs and semantic web technologies, this infrastructure connects research information preserved in local and regional research information systems. Linking local systems with each other creates a comprehensive global research information ecosystem. The FAIReR assessment infrastructure builds principally on community-owned, community-curated and openly available data on research. Following the FAIR principles, the data on research - respecting privacy issues - is “as open as possible and as closed as necessary”.

Automated input and extraction of assessment data
APIs are integrated into existing information systems, technologies and workflows which compose the FAIReR assessment infrastructure. Machine-readable formats define the input of assessment data. This creates a standard, transparent process

for collecting and inputting assessment data and minimises manual data entry. The FAIReR assessment infrastructure also facilitates reusing research information produced locally.

Infrastructure supports FAIR criteria, FAIR indicators and community building
The infrastructure enables describing and publishing assessment criteria and indicators in accordance with FAIR principles.

The FAIReR assessment infrastructure also includes web-based communication channels, social media and networking tools (e.g., forums, open reviews, blogs and registries of good practices). This supports information exchange and community building between different professionals, including open science experts, researchers, data stewards and research software engineers. These communications provide one way for gathering information about community practices and uses of, for example, indicators and assessment criteria. Building the FAIReR assessment infrastructure is based on this ongoing dialogue.

3. COMMUNITIES REWARD DIVERSE OPEN SCIENCE AND FAIR PRACTICES

Every organisation has its own FAIReR assessment policy to use in evaluations undertaken in the course of, for example, recruitments, promotions and funding decisions. Such policies are created in collaboration with academic staff and diverse research communities.

FAIReR assessment policies take into account the diversity of outputs, activities and professional roles involved in research and academic work. Researchers and other actors are recognised and rewarded for practicing and encouraging open science, in accordance with the OS-CAM recommendations.

Examples of the diversity included in FAIReR assessments, specifically related to research data, include:

Outputs such as data management plans and shared/published metadata and datasets.

Activities such as teaching or mentoring data management skills, reusing existing data, participating in data management training or the peer-reviewing data.

Professional roles such as data librarians and stewards, research software engineers, evaluators and researchers.

Organisations commit to using FAIReR assessment infrastructures, with transparent assessment criteria, that support the use of researcher portfolios and qualitative descriptions of research.

FAIReR assessment policies include both quantitative and qualitative approaches to evaluation. Metrics must be transparent, both in terms of how they are calculated and in how they are applied in assessments. Appropriate quantitative indicators are accompanied with qualitative assessments. All assessment data is best evaluated using responsible expert review, which helps to counter possible biases or conflicts of interest among evaluators.

III FROM VISION TO REALITY

Changes in academic assessment culture shake the research community at its core. Decisions around assessment define what is held important, valuable and where the research community wishes to go from here. Although “what” and “why” are important, nothing will happen without a “how”.

This change is going to take time and requires significant shared effort and investment in order to become reality. As for the vision, in the move from vision to reality we draw on our own research - policy review, survey - as well as co-creation process of bootcamps with experts in the areas of open science, research data and research evaluations.

Barriers for FAIReR assessment culture

If creating a FAIReR assessment culture were easy, it would have been done already. The changes in assessment culture meet challenges, which must be overcome. Awareness of barriers supports developing a roadmap from vision to reality. Barriers to this change include:

A. Cost in developing infrastructure

B. Lack of resources for qualitative assessment

C. Limitations in integrity and expertise in responsible assessment

D. Inconsistencies of assessment policies and cultures between communities

Given the scale of these challenges, a multi-stakeholder program - including government agencies, research funding and performing organisations, open science and research assessment and management communities, public and commercial service providers - is critical to achieve a shared global view.

PRIORITIES FOR ACTION

1 Policy collaboration
Because research and scholarship are international, there is a need for a global, shared vision for assessment. Shared vision will benefit all stakeholders. Countries and organisations make up the community, but none of them can alone change the culture. The change requires simultaneous international and local policy development.

International policy development creates a space for local innovations to flourish, which again feed and support international policies.

Priorities for policy development:

A. International agreements and/or MoUs on FAIReR assessment vision and policies

B. Co-creation of national and organisational FAIReR assessment policies

C. Support and populate platforms of best practices for FAIReR assessments and policies for mutual learning

D. Adopting, and where necessary creating, shared taxonomies of research contributions and contributors in policies regarding FAIReR assessments

2 Investment in assessment data infrastructure
Changes require resources - in time, energy and financial investment. Change

in assessment is no different. Policy development above requires investment in time and energy, while infrastructure requires considerable financial investments. To make FAIReR assessments possible, the research communities require new infrastructures for gathering, storing and sharing assessment data at institutional, national and international level. Key to success is finding a balance between harmonisation and diversity.

Balancing the use of metrics and qualitative assessments is also a question of wise use of resources. The key to decreasing the role metrics in assessments is to make better use of the time and effort experts invest in assessments. The experts' work can also be facilitated by producing good and reliable data to support qualitative evaluations, for example by developing infrastructure and services for the production, use, storing and sharing of structured and guided narrative descriptions and case reports. By making expert assessments more open (e.g., open peer review), the outcomes of an assessment can be used more than once. The experts' workload in peer review needs to be balanced with merit and rewards for peer review work. Ideally, performing assessments should be integrated in a natural way in the process of doing research.

Priorities for FAIReR assessment investments:
A. International funding call by EOSC, or other similar organisation, to begin building the technical solution for shared assessment infrastructure and required data models

B. Building a FAIReR assessment infrastructure with mutually shared architecture supporting quantitative and qualitative assessment information, which

- Has shared data models and PIDs
- Builds on existing infrastructures
- Provides different levels of access for different types of users
- Incorporates consent from researchers
- Minimises assessment data re-entry and maximises assessment data re-use

Establishing an international forum for dialogue between professionals, including open science and evaluation experts, researchers, data stewards and research software engineers to support development of FAIReR assessment infrastructure.

EUROPEAN OPEN SCIENCE CLOUD

Federation of Finnish Learned Societies

ANNEX 2

DOCUMENT REVIEW - Current state of responsible assessment of open science and research data

ANNEX 2

DOCUMENT REVIEW – CURRENT STATE OF RESPONSIBLE ASSESSMENT OF OPEN SCIENCE AND RESEARCH DATA

Annex 2 dataset: *Characterizations of resources for Open Science and researcher merit information*. Dragan Ivanović, Kathleen Gregory, Elina Koivisto. doi: 10.5281/zenodo.4704351.

Key takeaways

1. Guidelines for responsible assessments have common themes which need to be incorporated into FAIReR assessments: diversity, transparency, reflexivity and robustness of data and metrics. Academic assessments and practices of open science are situated within communities.
2. European policy-makers and institutions are committed to encouraging and rewarding open science practices, including sharing research data. The move away from a narrow focus on research, publications and metrics towards a broader range of assessment criteria remains limited. Open science is also narrowly conceptualised, referring often only to open access publishing.
3. The FAIR principles are often invoked. Measuring levels of FAIRness is challenging and perhaps not appropriate for judging individuals. Altmetrics for data and data citation are also in early stages; qualitative approaches for assessing open science offer promise.
4. Information and data produced by researchers, institutions and infrastructures remains scattered and difficult to (re)use in assessments. There is momentum for developing an interlinked infrastructure based on persistent identifiers which integrates research entities and facilitates interoperability between research information systems.

1. INTRODUCTION

This annex provides a detailed description of the methodologies used and the findings from the document and policy review presented in Section 1 of this report. We begin with a description of our methods. We then provide a detailed characterisation of the reviewed resources in the Results section, where we discuss the current state of academic assessments, assessing open science, evaluating FAIR data and the technical landscape which could be used to support FAIReR academic assessments. We include a brief summary of the key points at the beginning of each section of the Results.

2. METHODOLOGY

We reviewed 281 resources including research papers, project reports, policies, data models and eInfrastructures descriptions and projects. We organised our work using a spreadsheet, which we used to characterise the resources (Available in Zenodo, 10.5281/zenodo.4704351). We began by collectively identifying a set of known resources using our own previous knowledge. We then expanded this list using two primary techniques.

- Link/citation chaining:
  - Analysis of inbound citation links using the Google Scholar citation database,
  - Analysis of outbound citation links in the reference list at the end of the papers, reports, policies,
  - Analysis of web sites of analyzed infrastructure projects to find relevant and similar projects.
- Information retrieval techniques:
  - Searching of citation and literature databases using keywords of the project,
  - Searching the web using search engines and keywords of the project,
  - Searching and reviewing feeds/posts at social networks accounts of participants in the project.

The analysed resources were then categorized into the following classifications for review. We also cross-referenced items to indicate their relevance to multiple categories.

- Research assessment
- CV/portfolios
- Open science
- FAIR
- Incentive structures
- Data models, formats, vocabularies
- eInfrastructures

We characterised resources using the following fields/columns. The spreadsheet presented at the end of this Annex contains a selection of these fields, which are indicated here with an asterisk.

Bibliographic information

- ***Title:** Title of the resource
- ***URL**: URL to the the resource
- ***Authors/Publisher:** Author(s), Organization(s) or Publisher responsible for producing the resource
- ***Year:** Publication year for documents or year of putting the service in the operation
- ***Full citation** (if available)
- ***Type:** Type of the resource, e.g. policy, recommendations, paper, MoU, information system

Relevance to the project

- **Relevance score**: Ranked internally as high, medium or low to assess the relevance of the resource for the needs of this project
- ***All relevant topics**: Indication of which aspect(s)/topic(s) of the final project report the resource could be of interest
- ***Relevance description**: Brief description of how the resource can be used in the final project report

Content of the resource

- **Stakeholders/consumers/audience**: For whom the resource has been created
- **Participants:** Who contributed/participated in the creation of the resource, e.g. participants in the survey
- **Short description:** Short description of the resource
- **Key messages/features**: Key messages of the document, results of projects/researches or key features of the service
- **Limitations:** Limitations of the resource, e.g. geolocation of participants, national service, which aspects have been taken into account into analysis

Provenance information

- **Provenance - cataloguer:** Team member who added the resource
- **Provenance - date**: Date of adding the resource into this spreadsheet
- ***Provenance - methodology of finding resource**: How the resource has been found

3. RESULTS – DETAILED CHARACTERISATION OF RESOURCES

In this section we provide a detailed overview of resources analysed in the overview. The following 14 gaps have been identified in assessment practices and infrastructures as a result of our analysis of the literature, projects and platforms review described in this section:

1. There is a lot of support for responsible assessment initiatives, but implementation is not at the same level.
2. There are also many stakeholders voicing support for open science, but implementation at the policy level and in researchers' own practices vary.
3. Guidelines and policies for encouraging open access publications are much better developed/adopted than for open data (and other research objects).
4. Numbers of open access publications and the production of FAIR data are common measures of the uptake of open science practices - less attention is given to other research outputs or the broad potential scope of open science activities.

5. There is no clear approach on how to measure the quality of data. The definition of “quality” and “excellence” in evaluation policies is also unclear.
6. There are examples of proposed incentives and rewards for data sharing or data management without evaluation of the FAIRness of data. There is not a standard approach for measuring the FAIRness of data.
7. Data sharing and data management are discussed in policies, although these policies are often not enforced. Data sharing and management are also not often considered in researcher assessments. Assessing and rewarding instances of data reuse are even less common.
8. The focus in openness/FAIRness of data used in research evaluations is usually on making the criteria and methods used in evaluations open, including transparency in indicator development. Fewer documents call for making the data used in evaluations open.
9. Tools for making CVs/portfolios do not fully support the diversity of research objects’ types and scholarly communication channels.
10. There is a need to develop a scholarly infrastructure of interoperable, linked, transparent systems to support open science and robust evaluations which can be based on persistent identifiers assigned to each research ecosystem entity to enable tracking indicators for open science results.
11. There are well-known persistent identifiers for some research entity types, but not for all.
12. There are initiatives for building registers for discovering available data, publications, infrastructures and services, guidelines and policies, but there is no initiative for building registers for discovering available indicators and metrics.
13. A standardised semantic model for the representation of new indicators and a rule-based engine for execution of rules for classification of research objects according to the values of indicators are missing.
14. Research software engineers (RSEs) could provide software support and implementations of helpful research eInfrastructures, thus RSEs are important stakeholders in development of open science, and they deserve a path for a career which is not clearly present in the academy at the moment. Data stewardship also needs to be supported as a career path, in part to address the gap in researchers’ skills and knowledge of open science practices.

3.1. RESPONSIBILITY IN ACADEMIC ASSESSMENTS

Summary:

- *There is a growing recognition that research assessment needs to be responsible, in line with current calls for responsible research, i.e. DORA, and the Leiden Manifesto, to take into account different practices across disciplines, and to avoid common pitfalls in previous (mostly quantitative) assessments of research which rely on journal publications.*
- *There is much support for responsible assessment initiatives, but implementation is not at the same level.*
- *Transparency has become increasingly important in evaluations in the recent years.*
- *There is no clear approach on how to measure the "quality" and "excellence", and definition of those terms in evaluation policies is unclear.*

Making a shift toward responsible academic assessments requires major changes in practice from institutions, policymakers, funders and research communities. In a survey conducted by Wellcome about research and assessment cultures, researchers state that conditions are made worse by the complex network of incentives from governments, funders and institutions that seem to focus on quantity of outputs, and narrow concepts of 'impact', rather than on quality (Wellcome, 2020).

Efforts to develop and spur responsible academic assessments are many. Numerous documents, policies, articles and initiatives offer recommendations and commitments to responsible assessments and improving the ways in which research output is evaluated by funding agencies, academic institutions, and other parties (e.g. Wildgaard, 2015; Beigel et al., 2020; Gadd, 2019; Curry et al., 2019; Hatch and Curry, 2020).

Among the most notable of these efforts is the San Francisco Declaration on Research Assessment, known as DORA[1]. DORA was initiated by a group of editors and publishers of scholarly journals who met during the Annual Meeting of The American Society for Cell Biology (ASCB) in San Francisco, CA, on December 16, 2012 to develop a set of recommendations. DORA is now a central

[1] https://sfdora.org/read/

voice in the move toward responsible assessments; the organization maintains an open collection of materials to facilitate the development of responsible research and researcher assessment policies and practices[2]. They also provide examples of implementation, including "Reimagining Academic Assessment: Stories of Innovation and Change"[3], which presents ten case studies of universities and national consortia highlighting key elements of institutional change to improve academic career assessment.

An overview of other efforts of note is provided briefly below:

- The well-known Leiden Manifesto is another important influence in the move toward responsibility in research assessments. The authors (Hicks et al., 2015) define best practices in metrics-based research assessment so that researchers can hold evaluators to account, and evaluators can hold their indicators to account.

- The Metric Tide report (Wilsdon et al., 2017) discusses and elaborates criteria of responsible metrics from the perspective of the UK's Research Excellence Framework (REF), and also provides recommendations for research information management systems in the context of academic assessments. The five main principles include robustness, humility, transparency, diversity and reflexivity.

- Twenty-two academic leaders, funders and scientists gathered together at a workshop in Washington in 2017 and defined 6 principles for assessing scientists for hiring, promotion, and tenure (Moher et al., 2018). Among the principles identified by the group is that "Publishing all research completely and transparently, regardless of the results, should be rewarded".

- The Hong Kong Principles (Moher et al., 2020) focus on the need to drive research improvement through ensuring that researchers are explicitly recognised and rewarded for behaviours that strengthen research integrity. There are five principles and for each of them a rationale for its inclusion. These principles include: responsible research practices; transparent reporting; open science (open research); valuing a diversity of

[2] https://sfdora.org/resource-library/
[3] https://sfdora.org/dora-case-studies/

types of research; and recognising all contributions to research and scholarly activity.

- The INORMS Research Evaluation Working Group[4] was established to consider how best to ensure that research evaluation is meaningful, responsible and effective through two main work-packages: 1) An approach for evaluation of rankers and its ranking/assessments criterias has been published[5]; 2) a description of processes for doing research evaluation responsibly (SCOPE) is a result of the second work-package (Himanen and Gadd, 2019).

- The International School on Research Impact Assessment (ISRIA) published ten-point guidelines for an effective process of research impact assessment which are based on experience of more than 450 experts from 34 countries (Adam et al., 2018).

- The Global Young Academy organized a workshop in 2016 on publishing models, assessment, and open science (Curtiss and Gramatté, 2018). After the workshop, the participants created a list of 15 recommendations for improving publishing models, assessment practices, and open science uptake. There mirrors another initiative of young researchers to adopt research assessment practices to open science at the University of Utrecht (Algra et al., 2018).

- National initiatives such as Recommendation for the Responsible Evaluation of a Researcher in Finland are also of note. (Working group for responsible evaluation of a researcher, 2020). Although this is a national scope document, some recommendations can be applied more widely. The Strategy Evaluation Protocol (SEP) 2021–2027 for the Netherlands (VSNU et al., 2020) is another national effort, which outlines a protocol for context-sensitive research assessment of research units. This protocol takes into account the aims of an institution in research assessments, the quality of research and its societal relevance, and the viability of a research unit's goals based on scientific results from the last 6 years. SEP

[4] https://inorms.net/activities/research-evaluation-working-group/
[5] https://arma.ac.uk/rethinking-the-rankings/

incorporates the principles of the new recognition and rewards framework outlined in the Room for everyone's talent statement[6].

- COST-Action ENRESSH - European Network for Research Evaluation in the Social Sciences and the Humanities, which started in 2016 and ended in 2020, is another example of a regional effort to propose clear best practices in the field of SSH research evaluation[7]. The Latin American Forum for Research Assessment (FOLEC) is an example of a regional space for debate and exchange on the meanings, policies and practices of the research evaluation[8].

- Helsinki-Initiative on Multilingualism in Scholarly Communication, launched in 2019, highlights that language is an important aspect of diversity in research, and aims specifically to promote the value and recognition of multilingualism in assessment, evaluation and funding procedures[9]. The Helsinki-Initiative emphasises that language biases are produced in both evaluations based on research metrics as well as evaluation based on expert assessment.

The documents described above recognize that academic assessment occurs in diverse contexts: grants, hiring, career progression, and self-evaluation. The documents also detail different principles which should be included in responsible evaluations, e.g. transparency, integrity, fairness, humility, reflexivity, robustness, competence of evaluators, and diversity of outputs. A summary of how these themes are present in DORA, The Leiden manifesto and the Metric-tide report is presented in Figure 1.

[6] http://vsnu.nl/recognitionandrewards/wp-content/uploads/2019/11/Position-paper-Room-for-everyone%E2%80%99s-talent.pdf
[7] https://enressh.eu/; https://www.cost.eu/actions/CA15137/#tabs|Name:overview
[8] https://www.clacso.org/en/folec/
[9] https://www.helsinki-initiative.org

	DORA	METRIC TIDE	LEIDEN MANIFESTO
METHOD	1. Do not use journal-based metrics, such as Journal Impact Factors, as a surrogate measure of the quality of individual research articles, to assess an individual scientist's contributions, or in hiring, promotion, or funding decisions."	2. Humility: recognising that quantitative evaluation should support – but not supplant – qualitative, expert assessment	7. Base assessment of individual researchers on a qualitative judgement of their portfolio 1. Quantitative evaluation should support qualitative, expert assessment
CRITERIA	2. Be explicit about the criteria used in evaluating [researchers] and clearly highlight, especially for early-stage investigators, that the scientific content of a paper is much more important than publication metrics or the identity of the journal in which it was published.	5. Reflexivity: recognising and anticipating the systemic and potential effects of indicators, and updating them in response.	8. Avoid misplaced concreteness and false precision 9. Recognize the systemic effects of assessment and indicators 10. Scrutinize indicators regularly and update them
DATA	11. Be open and transparent by providing data and methods used to calculate all metrics. 12. Provide the data under a licence that allows unrestricted reuse, and provide computational access to data, where possible.	1. Robustness: basing metrics on the best possible data in terms of accuracy and scope 3. Transparency: keeping data collection and analytical processes open and transparent, so that those being evaluated can test and verify the results	4. Keep data collection and analytical processes open, transparent and simple 5. Allow those evaluated to verify data and analysis
DIVERSITY	3. For the purposes of research assessment, consider the value and impact of all research outputs (including datasets and software) in addition to research publications, and consider a broad range of impact measures including qualitative indicators of research impact, such as influence on policy and practice.	4. Diversity: accounting for variation by field, and using a range of indicators to reflect and support a plurality of research and researcher career paths across the system	2. Measure performance against the research missions of the institution, group or researcher 3. Protect excellence in locally relevant research 6. Account for variation by field in publication and citation practices

Figure 1. Comparison of three responsible assessment guidelines: DORA, the Leiden manifesto and the Metric-tide report.

There is agreement that responsible evaluations should consist of a mix of qualitative and quantitative assessments (Schmidt, 2020). Quantitative indicators should not be used for qualitative aspects; quantitative assessment should support, but not supplant qualitative approaches and expert assessment (Helmer et al., 2020). Publication metrics and metrics derived from publications, e.g. those calculated using bibliometric or citation databases (e.g. Web of Science, Scopus, and Google Scholar) are imperfect, criticised, and at times discouraged for use in research assessments.

Science Europe and the European University Association signed a joint statement demonstrating a commitment to building a strong dialogue between members who share the responsibility of developing and implementing more accurate, open, transparent, and responsible approaches that better reflects the evolution of research activity in the digital era[10].

Despite various such commitments to responsible assessment, we found indications that recommendations are not implemented, or are just beginning to be implemented in actual practice, at least for research performing organisations (RPOs).

---

[10] https://www.scienceeurope.org/our-resources/joint-statement-on-research-assessment/

For example, the report *"UK Progress Towards the Use of Metrics Responsibly - Three years on from The Metric Tide"* discusses problems and presents the current state of adoption of the Metric Tide recommendations in the UK (Saenen et al., 2019). The authors found that signing statements such as The Metric Tide, and/or developing institutional policies on the use of research metrics does not ensure that metrics will be used responsibly. Implementing these principles is a challenge, involving change at many levels, e.g., cultural changes and changes to systems, data collection, and analysis tools which use robust and credible data. Funding organisations and governmental policies may have more teeth in encouraging change (Saenen et al., 2019).

Transparency and quality

Transparency and trustworthiness are important for responsible researcher evaluation. Transparency can contribute to fairness of the assessment process (Working group for responsible evaluation of a researcher, 2020), as well as to quality of assessment process and research results.

To enable transparent metrics and informed evidence-based decision-making, e.g., eight UK universities began to develop Snowball Metrics[11], an initiative which provides a standardised approach to indicator development using freely available sources and recipes. Transparency is also becoming more important in evaluations for the needs of selection of research proposals in funding according to the results of a survey conducted by Science Europe (del Carmen Calatrava Moreno et al., 2019).

The European University Association's survey of RPOs (Saenen et al., 2019) further found that 63% participating organisations reported that information about research assessment approaches is publicly available, mostly on websites, while 34% of organizations make this information internally available. However, information about processes for assessing researchers is limited, and the transparency of assessment data itself is not addressed.

Although determining quality is a central premise of assessments, formal definitions of quality remain elusive (del Carmen Calatrava Moreno et al., 2019). There are, however, some initiatives to define criteria for proper research behaviour, and to maximise the quality and robustness of research, e.g. The

---

[11] https://snowballmetrics.com/

European Code of Conduct for Research Integrity[12]. The Global Young Academy "Optimizing assessment - Promoting Excellence" group[13] argues that "scientific excellence" should refer not only to excellence in research outputs, but also to excellence in connecting science to society, in teaching and mentoring scientists, in science management, and in providing scientific advice to policy makers.

Some organisations identify non-formalised concepts of 'quality' and 'excellence' in criteria as important mechanisms for differentiating the top tier applications in any applicant pool (Science Europe, 2020). Nevertheless, requesting 'excellence' from researchers and proposals, as a catch-word without qualification, may lead to a variety of unintended consequences. At an individual level, demanding 'excellence' can exacerbate systemic biases, promote individualism rather than 'team science' (Vogel et al., 2019), lead to a reduction in research integrity, and may not account well for the methodological nature of research.

3.2. ASSESSMENT OF RESEARCHERS

Summary:

- *Researcher assessment occurs in diverse ways, for example in the context of applying for funding (at the RFO level) or within RPOs when decisions are made about hiring, promotion and tenure.*
- *Two key features emerge that are important across assessment processes: (i) publication information and (ii) information about the researcher (i.e. his/her CV, academic background, collaborations, experience).*

There are diverse forms of assessment undertaken by diverse actors. Assessment occurs when decisions are made about tenure, promotion, hiring and awarding funding. Decisions are made by "peers" (other researchers within a discipline who are experts), universities - with staff members at different organisational levels - and funders.

According to a survey conducted by Science Europe of 31 research funding organizations (RFOs) and five research performing organizations (RPOs), the assessment practices of European funders for grant proposals are well-established, and rely on external single-blind peer review and external panel

---

[12] https://allea.org/code-of-conduct/

[13] https://globalyoungacademy.net/activities/optimising-assessment-promoting-excellence/

reviews. The RPOs participating in the study relied on external open review and internal panel review for the assessment of full research proposals (del Carmen Calatrava Moreno et al., 2019).

Assessment processes normally depend heavily on policies at national levels, particularly those of funders (European University Association, 2018). However, the different actors conducting assessments don't always work together or have the same policies and assessment criteria. A survey of European universities and RPOs (n=260) in 2019 found that institutions tend to be mostly autonomous in their assessment of the careers of individual researchers as well as the work of research groups and the allocation of internal funding (Saenen et al., 2019). This autonomy is exemplified by cases of individual universities discussed in the reviewed literature, such as that of the Utrecht University (Benedictus and Miedema, 2016), who moved ahead with the development of holistic policies for evaluation, despite a prominence of policies at national and institutional levels emphasising metrics-based evaluation.

Two key features emerge that are important across assessment processes: (i) publication information and (ii) information about the researcher (i.e. their CV, academic background, collaborations and experience).

3.2.1. PUBLICATION INFORMATION

Citations & bibliometric indicators
Summary:

- *Publication metrics based solely on journal publications, often calculated using bibliometric databases (e.g. Web of Science, Scopus, Google Scholar) are imperfect, criticised, and therefore even discouraged to be used for research assessment (e.g. DORA and critics of the Journal Impact Factor).*
- *Despite this, publication metrics continue to be used in, and perhaps dominate, academic assessment.*
- *There are some attempts/initiatives to move away from relying solely on publication metrics in research assessment at both the RFO and RPO level, notably by introducing more qualitative forms of assessment.*
- *The importance of counting other types of research output, primarily shared research data, in researchers' assessment has been recognised and partially adopted in some policies for evaluation.*

- *Metrics to non-traditional (not textual publications) research objects including (primarily) research data are still in a state of development, and can not be directly adopted from textual publication metrics.*
- *Although still developing in many disciplines, data citation is increasingly encouraged and supported through the development of new infrastructures and guidelines.*
- *Citations to data and other non-traditional outputs are not usually considered in researchers' assessments.*

Citations & bibliometric indicators for traditional outputs/journal articles

Publication metrics, i.e. number of publication, citation counts and derived bibliometric indicators, characterise much of the discussion on academic assessment. The report "Next Generation Metrics" (Wilsdon et al., 2017) provides an overview of these metrics. Traditional/conventional indicators measure output at the end of the research process, usually based on number of publications and number of citations calculated from popular global bibliometric databases such as the Web of Science, Scopus, and increasingly Google Scholar.

Publication metrics are imperfect. There are various problems with metrics, which are typically based on journal/article output, including problems with accuracy and comprehensiveness of data sources. Early career researchers (ECRs) may be disadvantaged and need to be evaluated according to other criteria, beyond bibliometric indicators (McNut, 2014). ECRs may also face unique challenges in adopting open science practices in a system that rewards more traditional ways of doing science (Allen and Mehler, 2019).

Especially the use of indicators based on aggregated citation counts of publications published in the same journal, notably the Journal Impact Factor (JIF), are increasingly criticized when used in researcher's evaluation. DORA, e.g., which currently[14] has 16,740 individual and 2,113 institutional signatories, discourages the use of the impact factor in researcher assessments. Aroeira and Castanho (2020) add to the many claims that the number of citations and JIF are not reliable measures of a paper's impact. Others claim that journal impact factors and other publication-source citation indicators should not be used as a standard of comparison between disciplines, because citation practices depend

[14] As of January 2021

very much on the subject area, with the result that a high impact factor for one discipline may look extremely low in comparison with another (Nisonger, 2004).

Despite their well-known limitations, journal metrics continue to be used in academic assessments. A study on academic assessment practices conducted by European University Association found that the majority of universities reported relying on a limited set of evaluation criteria and practices, most of which were based on research publications. Institutions reported using both quantitative and qualitative approaches in assessments, 75% of surveyed institutions (n=260) used the JIF to assess the research output of individual researchers, 70% the h-index, and a minority of respondents used other metrics, e.g., CiteScore, Source Normalized Impact per Paper (Saenen et al., 2019).

McKiernan and colleagues (2019) found that JIF-based evaluation is also present in North America (USA and Canada). Kraus (2014) interviewed 16 individuals with international perspectives on scholarly communication issues including open access and research impact issues. The article (Kraus, 2014) provides an overview of the discussions with those 16 interviewees. One conclusion of the discussion is that there is no “one best way” in the path toward open access. Although the majority of scholars still believe that an article published in an elite journal will receive higher recognition because of the impact and status of the journal, interviewees mentioned altmetrics as a way to measure the impact of research manuscripts promoted through social media channels. However, a survey organised by Research on Research institute (RoRi) shows that RFOs are phasing out JIF- based assessments of project proposals (Curry et al, 2020). Fifteen respondents in this study (43%) have eliminated the use of journal metrics in the evaluation practices, and a further five (14%) are planning to do so.

This observed tendency to evaluate researchers based on journal metrics has also been confirmed by numerous research articles. Jappe (2020) recently conducted a meta-evaluation of European assessment practices from 2005-2019. This study found that the JIF is sometimes used as a substitute for missing citation data (i.e. for recent publications). There is also a lack of access to citation databases across EU countries’ institutions which limits the type of bibliometric evaluations that organisations can conduct.

The aforementioned study conducted by Science Europe (del Carmen Calatrava Moreno et al., 2019) finds a trend amongst RFOs to move away from relying solely on bibliometric indicators in evaluations. The majority of participating organisations reported favouring qualitative assessments for decision-making. Some of the surveyed funders make use of bibliometric indicators, in addition to qualitative approaches. Of those who do, most consider the number of publications in high-ranking journals, the number of highly cited publications, the h-index and the total number of citations the researcher has accumulated.

This is in line with recommendations for responsibility, i.e. using a mix of quantitative and qualitative assessment.

Citations & bibliometric indicators for other outputs, particularly data

The importance of counting other types of research output, primarily shared research data, in researchers' assessment has been recognised and partially adopted in some policies for evaluation.

Numerous policies consider a variety of works besides publications. For example, about half (48%) of respondents in the survey conducted by EUA (Saenen et al., 2019) responded that other types of research output, besides publications, are important in academic assessment. Another study conducted by Science Europe (Science Europe, 2020) also confirmed this trend, finding that "36% of surveyed organisations had recently broadened the spectrum of non-publication research outputs (datasets, software, hardware, and so on) considered for assessment, with a further 13% planning to make this change."

The Research Excellence Framework[15] (REF) is a national framework for research assessment in the UK. Assessment prescribed by this framework is based on expert peer review, although usage of bibliometrics indicators is suggested to the reviewers. Four research outputs, some of which can be outputs other than publications, are required. Most of these policies focus only on the number of these "alternative" outputs (e.g. data, software, etc.) rather than on other quantitative indicators.

According to a SPARC Europe survey of funders, 26 out of the 60 funders who participated in the survey consider datasets in grant evaluations (Fosci et al., 2019).

[15] https://www.ref.ac.uk

Metrics for other types of outputs are still in a state of development. Metrics used for publications may not be suitable or have the same meaning when applied to other forms of scholarly output, i.e. data. Bibliometric citation, e.g., is not equivalent to data citation. Data citation involves a host of unique socio-technical challenges - the need for links to other objects, context for data to make sense, allowing for a variety of contribution roles, and more (Borgman, 2016). There is also the potential for various forms of citation aside from an article citing data. These could include, e.g. instances where data cite data - as a form of provenance, where a dataset derived from another needs to reference the original data (Lowenberg et al., 2019).

Citation practices for research data are also not standardized, and are taken up in various ways in different disciplines and communities. Not all data used in the research (i.e. for purposes of instrument calibration or results verification) would be necessarily cited in a paper (Federer, 2019). The meaning of citations also needs to be interrogated in regards to issues of credit and attribution. For example, who should be credited when data are created collectively, or when decisions about data collection are made years in advance (Borgman, 2016)? There are calls to go slowly and carefully with the development of data metrics, especially with the development of data citation and to take into account co-developing norms for citing alternative outputs (Stuart, 2017).

Data citation is, however, increasingly encouraged, although not necessarily in academic assessment contexts. There are projects and initiatives for infrastructural developments which aim to develop the social and technical infrastructure necessary for data citation (e.g. DataCite[16], Scholix[17], OpenAIRE Research Graph[18], Make Data Count[19]). Recommendations for data citation formats and guidelines are also prescribed by the FORCE 11 Joint Declaration of Data Citation Principles (Data Citation Synthesis Group, 2014). Principle Two calls for preserving metadata to include provenance information and to recognise all contributors to data, recognising that a single style may not fit all data types. Although citation is seen as being well-suited to "showcasing research impact" of data, efforts to standardise and promote data citation have had limited success so far (Konkiel, 2020).

---

[16] https://datacite.org/
[17] http://www.scholix.org/
[18] https://graph.openaire.eu/
[19] https://makedatacount.org/

Citations to various scholarly outputs, in particular data but also to other types of outputs, are not yet taken into account in the majority of academic assessment policies, although this has been discussed in the literature and at scientific conferences and workshops dealing with academic assessment. The European University Association's study (Saenen et al., 2019) asked respondents which altmetrics are used at institutions to measure the societal outreach of research, 35% of respondents report considering data citation in academic assessment, and 10% considered using DataCite as a source.

There are signs of support for the idea of data citation among researchers. In various surveys, data citation ranked highly as a motivating factor for sharing data (Fane et al., 2019). Also from another angle, outside of researcher practice, there are thoughts that citation would encourage data sharing (Konkiel, 2020).

Altmetrics

Summary:

- *Alternative metrics (altmetrics) can be used in assessments of attention and impact of contemporary research objects (research data, source code, news, etc.).*
- *There are issues with the application of altmetrics in researchers' assessment (reliability, accuracy, gaming by bots).*
- *There are also issues in applying altmetrics to non-traditional research outputs, i.e. data, in terms of versioning and granularity. These issues are not unique to altmetrics but also arise when applying other metrics to data.*
- *Altmetrics are rarely used in assessments and there is no strong evidence/indication it will be changed in the near future.*

Altmetrics in general

In recognition of the fact that data (and other outputs) are not literature, and that citation may not adequately capture the range of uses for these outputs, alternative metrics, or 'altmetrics', are being investigated as a way to demonstrate attention received by a range of outputs (Fecher and Friesike, 2014).

Altmetrics can be assigned for a variety of outputs such as books, blogs, reports, data or software. Sugimoto (2017) defined altmetrics as research indicators based on social media activity measuring attention such as shares, likes, tweets, comments and downloads. There are altmetrics aggregators helping in collecting diverse alternative metrics from various sources. Examples of such aggregators include PlumX[20] and Altmetrics[21].

There are many potential benefits of alternative metrics. Altmetrics could be used to assess new scholarly formats/outputs, and new types of usage could be recorded by alternative metrics. Altmetrics also offer potential in the assessment of interdisciplinary research and the impact of scientific results on the society as a whole, not only on the scholars, because alternative metrics can include the views of all citizens and not only other scholars (Wilsdon et al., 2017; Ravenscroft et al., 2017). Altmetrics can also be used for different purposes such as self-assessment and career development.

However, there are also challenges in the adoption and usage of alternative metrics in academic assessments (Haustein, 2016). Challenges include the heterogeneity of altmetrics, as most are tied to provider platforms and their regulations - i.e. Twitter and their rules for sharing data; the quality of the underlying data; their potential to be gamed by bots, and the lack of, among other things, consistency and accuracy. Data collected by altmetrics aggregators (e.g. Altmetric, PlumX) are also skewed toward the parent companies' data sources (Digital Science / Dimensions, Elsevier); and aggregators do not have meaningful disciplinary benchmarks (Konkiel, 2020).

Fenner (2014) identifies questions that need to be addressed before altmetrics can be used in assessment. In addition to raising concerns mirroring those above, Fenner questions if altmetrics can be applied across disciplines and over time to measure research, whether they measure impact or something else, and whether application of altmetrics in assessment can produce undesired incentives.

Altmetrics in general are rarely used in academic assessments[22], although there is some support for the idea (Konkiel, 2020). However, on the other side, more researchers disagree than agree that altmetrics in general should "count toward scholarly reputation" (Jamali et al, 2016)

According to the results of the Research Assessment in the Transition to Open Science survey organised by European University Association in 2019, altmetrics

[20] https://plumanalytics.com/
[21] https://www.altmetric.com/
[22] https://www.ref.ac.uk/ and https://www.arc.gov.au/excellence-research-australia

(views, downloads, etc.) are ranked as being less important than other indicators, including traditional publications (Saenen et al., 2019). Some funders in the Science Europe interview study report from 2019 are considering using altmetrics but the majority do not and do not indicate that they are considering it (del Carmen Calatrava Moreno et al., 2019).

Altmetrics for data

There are some additional key challenges specific to data, in addition to the general limitations discussed above. Data are dynamic, meaning that data can be updated and versioned. It is challenging to bring these versions together or to make comparisons across them. Similar to citation databases, there are different data sources from individual repositories to aggregators. Those sources can also define and use numerous metrics.

The authors of "Open Data Metrics: Lighting the Fire" (Lowenberg et al., 2019) suggest differentiating data usage at different levels, e.g., different levels of views (of metadata, images, etc.) and downloads (file level or dataset level). They warn that without standardisation of counting views and downloads, any definition of views and downloads used by various stakeholders can be arbitrary. Counts are not metrics per se, but those counts can be standardised in order to create a metric. The COUNTER Code for Research Data[23] is a promising way to standardise data usage counts, but at the moment there are many different ways of counting which are implemented. There are calls for Individual repositories to report their usage metrics to centralised data brokers like DataCite, but this rarely occurs (Konkiel, 2020). Transforming counts to a measure of data reuse is a thorny issue and raises a lot of questions: What type of use? Scholarly? Societal? By whom and for which purposes?

3.2.2. RESEARCHER INFORMATION

Summary:

- *Information about a researcher such as academic background, collaboration, skills, expertise and qualifications are taken into account in the researcher assessment process.*
- *Researcher evaluation is not divorced from research evaluation. Research evaluation often occurs at the publication level (i.e. through peer review).*

[23] https://www.projectcounter.org/code-practice-research-data/

Although there are different evaluation systems at the publication level and individual level, those systems are tightly linked.

According to the results of the Research Assessment in the Transition to Open Science survey organised by European University Association, also many other aspects of scholarly work - besides publications - are important in evaluating researchers (Saenen et al., 2019). These include attracting funding (75% most important/important); collaborations, supervision, and teaching (over 60% most important/important). Bibliometric indicators and quantitative approaches can be also used for measuring collaborations and co-authorship. Moreover, the potential spectrum of criteria for the needs of self-evaluation reporting is broader for evaluating researchers' groups than those used for individual research careers, including, for example, the number of doctoral candidates active at or graduating from a research unit.

There are also some calls for a broader recognition of roles enabling the diversification and vitalisation of career paths (VSNU et al., 2019; Working group for responsible evaluation of a researcher, 2020).

As the EU report "Indicator Frameworks for Fostering Open Knowledge Practices in Science and Scholarship" (Wouters et al., 2019) emphasises, researcher evaluation is not divorced from research evaluation. There are different reputation/reward systems operating in science at the community level (mostly peer review of research - i.e. in publications, conferences) and at the individual level (tenure promotion, grants). Although the latter is our focus in this document, the two forms of evaluation are tightly linked in the current system, as evidenced by the coupling between publications and evaluative systems.

The authors of the EU report suggest that the community level of research evaluation, particularly peer review, could also be made more open. Some work is already underway to address this issue, for instance through recommendations for diversifying peer review, preprints, open pre- and post-publication review, and services offering support for these activities, e.g. PubMedCommons, Qeios, Winover, F1000 (Munafò et al., 2017).

CV / portfolio templates

Summary:

- CV/portfolio templates simplify assessment based on standardised CVs/portfolios.
- CV/portfolio templates only partially support diversity of research objects' types and scholarly communication channels at the moment.

CV/portfolio templates support resaercher's assessment with standardised information structure. The Finnish National Board on Research Integrity TENK, Universities Finland UNIFI, Rectors' Conference of Finnish Universities of Applied Sciences Arene and the Academy of Finland prepared the first template for the researcher's curriculum vitae (CV) in accordance with the responsible conduct of research for Finnish research organisations in 2012. This curriculum vitae template[24] was updated in 2020. This template takes into account contemporary trends in evaluation of open science and social networks activities, as well as other aspects of research assessments. The scope of the template is limited to one country, although the good practices from this template have the potential to be adopted at the European level.

In Finland, a vocabulary for researcher merit activities[25] is human and machine readable, which is in accordance with FAIR principles (I2), and enables classifying researcher's activities to groups making assessment of researcher's results easier and more transparent. Unfortunately, some open-science-related terms are not present in the vocabulary at the moment.

The ACUMEN Portfolio is a tool for evaluators and scholars for evaluation of scientific works and careers (ACUMEN Consortium, 2014). The ACUMEN Portfolio enables researchers to highlight their achievements, including not only publication information but also other tasks, such as digital methods of collaboration, researching, teaching and contributing to society. ACUMEN supplements the traditional CV because it enables researchers to present themselves in the most positive way by highlighting key achievements rather than giving an exhaustive list of references. However, it does not recognise sharing open data as an achievement.

---

[24] https://tenk.fi/en/advice-and-materials/template-researchers-curriculum-vitae

[25] https://koodistot.suomi.fi/codescheme;registryCode=research;schemeCode=aktiviteetitjaroolit

The NIH Central Resource for Grants and Funding Information published a template, instructions, and samples for a researcher's curriculum vitae (CV) in the application process for NIH Grants and Fundings for the purpose of evaluation of a project team capacity to complete the project's tasks[26]. A project applicant can cite up to four publications or research products that highlight their experience and qualifications for the proposed project. Research products can include a wide range of research digital objects such as audio or video products; conference proceedings such as meeting abstracts, posters, or other presentations; patents; data and research materials; databases; educational aids or curricula; instruments or equipment; models; protocols; and software or netware. However, the NIH template is not widely adopted, it is developed only for the specific purpose of application for NIH grants.

Résumé for Researchers[27] is a standardised template for a narrative description of a researcher's achievement. It provides a standardised narrative CV format to facilitate the recognition of a range of research contributions divided into the modules publications, funding and awards, as well as activities such as public engagement, training and knowledge exchange.

CV/portfolio templates only partially support the diversity of research objects' types and scholarly communication channels at the moment.

Piwowar and Priem (2013) discussed risks and benefits of including altmetrics in building CVs. They recommended that altmetrics should not be considered as a replacement for careful expert evaluation of CVs and other traditional metrics, but should be considered as a supplement if they are accurate, clear and meaningful. The addition of altmetrics in a CV can provide additional information about any type of research objects and uncover the impact of just-published objects. Including altmetrics in CVs will reward researchers who opened and provided additional description of their research objects through numerous channels and encourage others to do so in the future.

Best practices for the evaluation of CVs are needed. Those best practices need to be published and promoted to encourage a broader recognition of the range of verifiable contributions individuals can make to the knowledge system, including teaching and peer review, and the production of a broad range of open science

[26] https://grants.nih.gov/grants/forms/biosketch.htm
[27] https://royalsociety.org/topics-policy/projects/research-culture/tools-for-support/resume-for-researchers/

output types including datasets (European Commission, 2018b). The career narrative should be the basis for a researchers' evaluation as it provides the crucial context in which indicators can be interpreted (European Commission, 2018b).

3.3. OPEN SCIENCE EVALUATION

Summary:

- *Metrics for open science play 2 roles: (i) monitoring development of open science within the scientific system as a whole; (ii) measuring performance in order to reward OS practices at group and individual level*
- *With a few notable exceptions, evaluating open science practices at the group/individual level is in the recommendation stage, rather than the implementation stage.*

Metrics for OS play 2 roles: (i) monitoring the development of open science within the scientific system as a whole; and (ii) measuring performance in order to reward OS practices at group and individual levels (Wilsdon et al., 2017; European Commission, 2018a).

The measurement and evaluation of open science is still in the development phase. Although there has been much interest in this topic over the past three years, metrics for open science remain relatively un-used in academic assessments (Morais and Borrell-Damián, 2018; Saenen et al., 2019).

Making recommendations for assessing and evaluating open science outputs and activities is much more common in the reviewed resources than reports of implementation or practical developments. Recognising this problem, a report from the EU Open Science Policy Platform presents practical commitments to implementation and measures progress in adopting current recommendations with the aim of stimulating open science to move beyond the recommendation phase to an implementation phase (Lawrence et al., 2020).

In this section, we discuss the developments and recommendations toward evaluating open science along two themes which we identified in the reviewed resources: (i) the need to **foster open science** practices through guidelines, technical developments, training, and changes to reward and incentive infrastructures and (ii) recommendations and ideas for **assessing open science**.

3.3.1. FOSTERING OPEN SCIENCE

Summary:

- *There is a need to foster open science practices through guidelines and policies, development of eInfrastructures, trainings and support, and rewards/incentives*

Although some researchers see collective benefits of sharing research data, they are also at times reluctant to share data, and do so only selectively (Fecher et al., 2017). Some see research datasets as a raw material for article publications; others believe that other researchers could publish with their datasets first. On the other hand, Baccigotti[28] for example claims that open science and intellectual property rights (IPRs) do not have to be on opposite sides, and that both open science and IPRs are necessary, beneficial and consistent with each other.

The Workshop Report on Digital Transformation in Scholarly Communication argues, from the researcher perspective, that scholars (will) need access to an increased number of academic outputs to perform research in the future (Science Europe, 2019). This topic has also been discussed in a report of the Expert Group on the Future of Scholarly Publishing and Scholarly Communication to the European Commission (Guédon et al., 2019).

PLAN-E[29] is the Platform of National eScience Centers in Europe which unites the efforts of more than 30 eScience organizations across Europe to strengthen the European position in the eScience domain. PLAN-E stresses the importance of general openness in scientific achievements, results and reviewed publications, wherever possible enriched by providing access to underlying data, descriptions of procedures followed and/or details of equipment used in getting the published results.

[28] https://www.the-guild.eu/blog/confessions-of-a-knowledge-transfer-manager.html)
[29] https://plan-europe.eu/

Guidelines and policies for fostering open science

Summary:

- *Guidelines and policies are a way to encourage open science/open data practices at both the RFO and RPO level.*
- *Guidelines and policies for encouraging open access publications are much better developed and adopted than policies for open or shared data and other research objects.*
- *Although policy enforcement is not the norm, there is a feeling that funders should enforce policies requiring researchers to share data.*
- *Sharing data is not always possible, therefore policies and incentives should be flexible to some level.*
- *Guidelines and policies should be aligned and good practices should be recognized.*

Establishing clear guidelines and policies, including relevant legal information, has been identified as a way to encourage open science/open data practices (Morais and Borrell-Damián, 2018). Robust funder requirements are seen as key to a widespread uptake of open science (Leonelli et al., 2015). The Budapest Open Access Initiative[30], e.g., issued an early guideline for the development of open access policies in institutions of higher education and in funding agencies. Differently defined policies for open science are being implemented at the funder, institutional and national levels (European Commission, 2018b; The Serbian Ministry of Education, Science and Technological Development, 2018).

Despite recognition that open science involves a broad host of practices and considerations such as open/shared data, open methods and open education[31] (Guy and Ploeger, 2015), many policies focus on regulating and encouraging open access to publications alone (The PASTEUR4OA consortium, 2016; Leonelli et al., 2015).

Similarly, a recent survey conducted by SPARC Europe of 60 European funders found that 61% of respondents reported having a policy for open access to publications; nearly the same number (69%) reported not having a data policy (Fosci et al., 2019). A study from 2017/2018 found a similar trend among RPOs; the majority of surveyed institutions (n=321) had open access policies, while 40% lacked or were not planning on developing data policies (Morais and Borrell-

---

[30] https://www.budapestopenaccessinitiative.org/

[31] http://fosteropenscience.eu/

Damián, 2018). There are, however, calls from organisations such as LERU (the League of European Research Universities) to develop data policies, particularly policies that are in line with FAIR data principles (Ayris et al., 2018).

A similar trend is also present at the national level, where open science usually equates to open access publications and, at times, open data (Sveinsdottir et al., 2020). SPARC Europe conducted analysis of open science / open data policies in 2020 and identified 14 national policies, 11 of which were from EU member states. The majority included open access to publications and research data in one policy. This study also highlighted two more holistic policies from France and the Slovak Republic, which include other "open" projects - i.e. open government partnerships to address open education, governments open to dialogue, and open justice.

Although it only pertains to open access publications, Plan S is a notable initiative supported by cOAlition S[32], an international consortium of research funders. Launched in 2018, this initiative of more than 20 organisations requires that scientific publications resulting from publicly-funded research must be published in compliant Open Access journals or platforms, beginning in 2021. There is also a guideline on how to follow the principles and implement the plan.

The FAIRsFAIR's analysis of the data policy landscape (Davidson et al., 2019) and data practices (Whyte et al., 2019) in 2019 and found that there is still much that needs to be done to foster and harmonise policies in order to support the aims of the European Open Science Cloud and realise the vision of Turning FAIR into Reality (TFiR). Additionally, FAIRsFAIR has prepared a series of practical recommendations for policy enhancement to support the realisation of a FAIR ecosystem (Davidson et al., 2020). The European Open Science Cloud FAIR Working Group has also recognised the need for developing and monitoring policies for FAIR data and research objects as being an important step (The EOSC Executive Board, 2020) .

Although numerous policies require or recommend data sharing and the creation of data management plans, these activities are not tied to researcher evaluation. There is also a general leniency in compliance, among funders, when policies are not enforced (Fosci et al., 2019). This observed leniency mirrors calls for stricter enforcement of policies at the RPO level (Ayris et al., 2018). Moreover,

---

[32] https://www.coalition-s.org/

some evidence suggests that scholars also believe that funders should both require data sharing and take a harder line in enforcing policies, i.e. by withholding funding when data are not shared (Fane et al., 2019).

The EOSC-Pillar 'National Initiatives' Survey aims to gain insight into the stage of open science development in the five EOSC-Pillar countries: Austria, Belgium, France, Germany and Italy (Boldos et al., 2020). The results of this survey show that funding bodies frequently impose mandatory rules for data management plans (DMPs, 40%), open access publications (36%), open research data (32%), compliance of data to the FAIR principles (32%), publication of data in a repository (28%) and the long-term availability of data (28%). Bloemers and Montensanti (2020) proposed a framework for research funders to drive the transition toward FAIR Data Management and Stewardship Practices.

Other top-down approaches seek to further define a broad standard for open access to research output, while allowing for possible gradations, at the policy level. Dutch Ministry of Education, Culture and Science suggests that funders set the default in data sharing to open access, but allow a choice of access options to account for disciplinary and data differences: from open and free downloads to application and registration-based access (Dutch Ministry of Education, Culture and Science, 2016).

Challenges remain, as some policies may be complicated or not applicable to certain research areas or for certain data[33]. The issue of appropriately sharing data containing personal information, e.g. health information in biomedical settings, is well known within the research data community (Corti et al., 2019). Challenges also arise in research resulting from collaborations with industrial partners, who may not be willing to participate in or fund collaborative work if data sharing is mandated (Ali-Khan et al., 2017). Strict policies may be unlikely to encourage open practices in such situations.

The diversity of policies can be problematic. Guidelines and policies should be aligned and good practices should be recognised. Earlier work has been undertaken to bring together open access policies at the European level as a part of the Pasteur4OA project[34]. The PASTEUR4OA's Knowledge Net activities to support the development and alignment of open access policies in Europe

[33] See for example https://www.gida-global.org/care
[34] http://pasteur4oa.eu/

continued as the Knowledge Net integrated into the OpenAIRE infrastructure from September 2016.

ROARMAP[35] is a searchable international registry of open access mandates and policies adopted by universities, research institutions and research funders that request depositing research article outputs in open access repositories. There are more than 1,000 catalogued policies in the repository at the moment of writing this report. Again, the majority of catalogued policies are related to open access publications, although some policies could also cover openness of other research objects. ROARMAP exists in order to keep an accurate record of policies, mandates and details of open access repositories across the world.

Need for technical infrastructures for fostering open science

Summary:

- *Technologies do not bring change on their own, but eInfrastructures should support open science and evaluation practices for open science.*
- *Persistent identifiers should be assigned to each research ecosystem entity to enable tracking indicators for open science results.*
- *There is a need to develop a scholarly infrastructure of interoperable, linked, transparent systems to support open science and robust evaluations.*

Science Europe organised a two-day workshop in November 2019 entitled Digital Transformation in scholarly communication. The event report was published after the workshop (Science Europe, 2019). One of the conclusions was that technologies do not bring change on their own, and a crucial point in defining new scholarly publication systems is the interaction between technology readiness and the scholars' culture, for example on research evaluation methodologies. Similarly, representatives of organisations signing the DORA declaration reported difficulties in changing scholars' ways of thinking and existing habits.

Brian Nosek, in his vision for cultural change (Nosek, 2018), emphasises that a necessary element to making open science possible is the development of supporting eInfrastructures. Other groups, e.g., FORCE 11[36], a community of scholars, librarians, archivists, publishers and research funders with a goal of promotion of open science and the FAIR principles, similarly see the effective use

[35] http://roarmap.eprints.org/
[36] https://www.force11.org/

of information technology as being a key component in encouraging open science.

Metadata linking every piece of scientific knowledge to a unique and persistent identifier is the cornerstone of such infrastructures (Wilsdon et al., 2017). Authors of the Next Generation Metrics report also recommended assigning PIDs at all levels (for all objects, individuals, institutions) to support citation, as well as linking work to individuals (Wilsdon et al., 2017).

The ORCID identifier is seen as being key, with some calling for ORCIDs to be mandatory for all applicants and participants in European funding programs (European Commission, 2018b; Wilsdon et al., 2017). An example of a platform recognising and promoting open research based on PIDs (ORCID and DOI) is Rescognito[37]. Well-known persistent identifiers exist for some research entity types, but not for all. This may change, however, given existing initiatives and projects working to develop PIDs, e.g. Freya[38], the PID Forum[39], and the PID architecture for the EOSC (Schwardmann et al., 2020).

Ficarra and colleagues present the current state of the development of digital infrastructures for fostering open science, based on the results of a survey (Ficarra et al., 2020). This study is a step towards gaining a better understanding of the current open science ecosystem, and it demonstrates the need for still further development. The study also highlights issues with the sustainability of research infrastructures due to unstable funding streams. It calls for funding agencies, governments, and institutions to make a strategy for effectively and structurally funding infrastructures for open science.

Besides this report, there are many general calls for the need to develop interoperable, linked, and transparent scholarly infrastructures and research information models to support open science and robust evaluation (Mirowski, 2018; European Commission, 2018b; Dutch Ministry of Education, Culture and Science, 2016; Materska, 2019).

Similarly, in data-related documents, the discussion of necessary infrastructures remains at a high level. Laine (2018) called for data infrastructures supporting multiple types and uses of data. Perrier and colleagues (2020) suggested making

---

37 https://rescognito.com/
38 https://www.project-freya.eu/en
39 https://www.pidforum.org/

data sharing easier through improving infrastructure support which should allow making data available quickly and seamlessly. However, a survey conducted by SPARC Europe (Fosci et al., 2019) found that the majority of funders do not currently provide support for data infrastructures.

Training and support for fostering open science

Summary:

- *Training is also seen as playing a critical role in facilitating the uptake of open science practices through the development of researchers' skills.*
- *There are indications that many researchers' knowledge of open science skills is limited primarily to basic skills for making publications and data open access.*
- *The professionalization of data stewardship is recognised as being pivotal to the transition to open science.*
- *Research software engineers (RSEs) are also important stakeholders in development of open science.*

Training is seen as playing a critical role in facilitating the uptake of open science practices through development of researchers' skills. Numerous documents point toward the need to enhance open science literacy (e.g. European Commission, 2018b; Smederevac et al., 2020) and to offer differentiated professional development and training opportunities for researchers in different disciplines and career stages (Ayris et al., 2018; Forsström et al., 2020).

Proposals for open science training opportunities seem to center around skills for sharing and working with data. Leonelli and colleagues (2015) proposed that training should include support for skills such as self-archiving, different data sharing options as well as the ethical, social and regulatory aspects of big data. The Open Science Skills Working Group of the European Commission conducted a survey (1,277 researchers participated) with the aim of proposing suggestions for open science training (O'Carroll et al., 2017). They found that most researchers' conceptions of open science skills were limited to making publications and data open access and that most did not have knowledge about other aspects of open science, e.g. citizen science, open education, open notebooks and open peer review.

The FOSTER training program[40], developed in a project running from 2017-2019, was developed to support a diversity of open science practices; it contains a significant number of resources for data sharing, data management, and data mining. These resources include a blend of face-to-face training events, e-learning courses and online training resources. The project educated researchers and trainers about reproducible research (e.g. through using lab notebooks), open evaluation, and embedding open science and legal aspects pertaining to open science in regular workflows. These materials are also grouped together in the Open Science Training Handbook[41]. Other training for researchers interested in open science is organised regularly within the NI4OS project[42], and in the data science summer schools organised by the RDA alliance[43].

The authors of the "Amsterdam Call for Action for Open Science" called for national authorities and research funders to formally recognise the profession of data stewardship and provide data stewards with career opportunities (Dutch Ministry of Education, Culture and Science, 2016).

One of four thematic sections in the FAIRsFAIR report about identifying recommended actions to enable FAIR data is developing professional support for FAIR data (Theme C, Molloy et al., 2020). There are recommendations to further define and manage the costs and resources involved in the support of FAIR and to implement models for coordinating and supporting data stewards and research software engineers. Similarly, the European Commission (European Commission, 2019) analysed costs of FAIR implementations and recommended securing coverage of data stewardship and management, as well as data infrastructure operational costs needed for making the transition to open science possible. Other FAIRsFAIR deliverables also identified roles of data stewardship in a competence center (Herterich, 2019).

Research software engineers (RSEs) could provide both software support and implementations of research eInfrastructures. RSEs are therefore seen as being important stakeholders in the development of open science, deserving career paths which are not currently present in the academy (Seibold, 2019). Cohen and colleagues (2021) also highlight the importance of research infrastructure and

[40] http://fosteropenscience.eu/
[41] https://www.fosteropenscience.eu/content/open-science-training-handbook
[42] https://ni4os.eu/
[43] https://www.rd-alliance.org/

research software engineers. The importance of RSEs has been recognized in some disciplines and countries. For instance, some technical fields have started embracing research software engineering roles[44], but others highlight limited opportunities in fields such as the biomedical sciences (Kucharski et al., 2020).

Rewards & incentives for fostering open science

Summary:

- *Although there is a recognised need for incentives to foster open science, there is a lack of effective ones for encouraging practices of open science.*
- *To appropriately recognise the work involved, and in order to provide an incentive for data sharing, individuals who initially gather data should receive appropriate and standardized credit that can be used for academic advancement, for grant applications, and in broader situations.*
- *Rewards based on traditional metrics of academic success are often used to motivate researchers to share research data.*
- *Data reuse is a goal on its own, not just sharing data. While there are many proposed benefits for data reuse, there are minimal incentives at the moment for reusing dat*a.
- *There is conflicting evidence about whether or not the move from recommendation to implementation of incentives for open science practices is in progress.*

There are many broad calls for funders and institutions to develop incentives and rewards for individuals, groups and projects which are applicable to all disciplines, career stages, a range of outputs and a range of career directions (see for example European Commission, 2018b). According to the results of a survey on open access organised by European University Association in 2018 (Morais and Borrell-Damián, 2018), however, there is a lack of effective incentives to encourage practices of open science.

The majority of proposed incentives focus on using traditional reward structures (i.e. citations) and other measures of academic success important in scientific work, such as collaborations, job opportunities, opportunities for career advancement and funding opportunities (summarised in McKiernan et al., 2016, also see Wilsdon et al., 2017).

[44] See for example https://www.ukri.org/opportunity/research-software-engineer-fellowships-2020/

Citations to a range of outputs including research data are commonly seen as a way to incentivise sharing one's own data or outputs. For example, the European Commission's Expert Group on Altmetrics stated in their report (Wilsdon et al., 2017) that open science practices should be recognised and rewarded "via the currency of citations, by linking every piece of scientific knowledge to a unique and persistent identifier."

This matches researchers' own motivations for sharing their data. The 2019 version of the annual State of Open Data report (Fane et al., 2019), which presents the results of a survey of over 8500 researchers, found that 61% of respondents report citations as the favored way of receiving credit for sharing data. Other factors reported as being valuable mechanisms for receiving credit for data sharing in this study include consideration in job interviews (45%) and receiving co-authorship on papers (42%).

An open question is whether data citations are rewards/incentives for sharing data, or for making them useful and used by others. Pierce and colleagues (2019), e.g. encouraged rewarding/recognising "generating data that has become useful to other researchers". Data citation could be a way to encourage reuse of data by others (although not necessarily to reward it in this instance). By linking to rich metadata, citation could provide contextual information that would facilitate understanding/evaluation of data by potential reusers (Groth et al., 2020).

Data citations are seen as incentives for data sharing; they are also viewed as a means of providing credit which could translate into use in academic assessments. Foster and colleagues (2019), e.g., recognized the role of data citation in assigning credit and attribution and provided an example of data citation format.

The Data Citation Principles[45], published by the FORCE 11 Data Citation Working Group[46] promote citation practices that are both human understandable and machine-actionable. These principles do not include recommendations for specific implementations, but encourage communities to develop practices and tools that embody these principles. Eight principles have been defined including the second principle, Credit and Attribution, which recommends giving credits to

---

[45] https://www.force11.org/datacitationprinciples
[46] https://www.force11.org/datacitation/workinggroup

all contributors to research data. The principles are defined for research data, not for other research objects, e.g., software code, multimedia, etc.

Tanhua and colleagues (2019) identified an issue with giving credit to datasets authors; datasets are often aggregated into data products, therefore it is not obvious how to give appropriate credit to all of the scientists who generated the data. Bierer and colleagues (2017) proposed a system for recognition of data authorship. The authors of this paper proposed a set of responsibilities for a researcher who should be cited as a data author. These include substantial contributions to the original acquisition, quality control, and curation of the data.

Including requirements for open science practices within the hiring and promotion process (for instance in job descriptions), performance appraisals and promotion criteria for all or most research and teaching staff, is another commonly mentioned theme (Ayris et al., 2018), which we discuss further in section 3.3.2.

Other possible incentives for open practices are also mentioned or proposed, although not as frequently. Many researchers report ethical motivations for adopting open science practices, i.e. data sharing (Fane et al., 2019; Ali-Khan et al., 2017). Other proposals for encouraging open practices include the use of badges (in journals) for sharing and describing data in a way that enhances reusability (Kidwell et al., 2016). Other initiatives also encourage assigning open science badges[47]. Implementing badges is associated with an increased rate of data sharing (Kidwell et al., 2016); as of 2016, 68 journals offered open science badges.

In another initiative, Shrestha and Vassileva (2018; 2019) built a blockchain-based framework to allow data owners to protect their data and to get rewards (digital tokens) from sharing their data. The proposed framework supports building a verifiable record of the provenance, accountability of access and incentives for data owners.

It is also notable that many incentives (in the data realm) focus on rewarding and incentivising data sharing rather than data reuse. While many benefits for data reuse have been discussed in the broader literature (e.g. reproducibility, cost savings), there are minimal incentives at the moment for reusing data. Many different types of data reuse could be incentivised in different ways, as we

[47] https://www.cos.io/our-services/badges

further discuss in the below section regarding FAIR data. One proposed incentive for encouraging data reuse is to offer research grants and funding opportunities specifically for instances of reuse of data generated from earlier grants (as summarised in Perrier et al., 2020).

3.3.2. ASSESSING OPEN SCIENCE

What to measure when assessing open science

Summary:

- *Numbers of open access publications and the production of FAIR data are common measures of the uptake of open science practices; less attention is given to other research outputs or to the broad potential scope of open science activities.*
- *There is a convergence in research practices and policies on measuring aspects related to data sharing, data management and data reuse.*

The starting point for measuring and assessing open science practices should be defining the objectives and outcomes of open science (Wilsdon et al., 2017). Many considerations about what to measure fall along the lines of the outputs highlighted in previous sections, namely open access to publications and data. Less attention is given to other research outputs or the broad potential scope of open science activities.

For instance, open access monitoring efforts in Finland are based on counting open access publications (Ilva, 2017). Furthermore, Robinson-Garcia and colleagues (2020) measured the uptake of open access by universities worldwide using the number of open access articles published in journals indexed in the Web of Science. Therefore, measuring the number of articles in open access journals is one common factor, although another could also be to recognise article preprints, as suggested in (Ayris et al., 2018). In a study of funders in 2019, however, SPARC Europe found that open access to publications is still not a factor taken into consideration in grant evaluations (Fosci et al., 2019).

There is a broad recognition that FAIR data is critical to achieving the goals of open science (e.g. Burgelman et al., 2019) and that RFOs and RPOs should give credit/reward to researchers for making data and other outputs - methods, codes, publications - FAIR (European Commission, 2018b). There are numerous calls to recognise and incentivise FAIR data practices. For example, Stall and colleagues (2019) described the experience of the Geosciences, describing how societies and academies should specifically include and mandate data sharing and FAIR treatment of data.

This emphasis on FAIR, and the link between FAIR and research integrity within Europe, is also apparent in a recent analysis of major codes of conduct for research. Laine (2018) found by analysing codes of conduct across the world, that the European code is the only one which pays an almost equal amount of attention to data management (framed in terms of FAIR in the analysis) as to publications.

In terms of researcher practices, we see a convergence in the reviewed literature on measuring aspects related to data sharing, data management and data reuse. This is indicated, for example, in a recent review of 14 national open science/open data policies conducted by SPARC Europe (Sveinsdottir et al., 2020), which evaluated policies according to criteria including their mentions of FAIR, data sharing, data availability statements, data management plans, data citations and open licensing information. The analysis found that under half of policies mention FAIR data explicitly, while most do so implicitly by addressing individual aspects discussed in FAIR such as data accessibility or licensing.

Reflecting the fact that data citation is an emergent and evolving practice in many disciplines, only four of the policies reviewed in the SPARC Europe study mention expectations regarding data citation (Sveinsdottir et al., 2020). As it is already stated before in this report, while data citation is one way to measure data reuse, current framings are such that any potential reward for citing data belongs to the data sharer, rather than the person using the data.

The existence of data management plans offer another possible indication for FAIR data practices, although DMPs/Output management plans (OMPs) are not regularly taken into account in researcher evaluations. There are, however, recommendations for OMPs to be machine readable and regularly updated (European Commission, 2018b).

Sharing and managing data, protocols and other results in a FAIR way takes much *time and effort* on the part of the researcher. This type of *meaningful* sharing should be rewarded (Leonelli et al., 2015), perhaps more than data not described or managed in accordance with disciplinary/data management standards. Moreover, Leonelli and colleagues (2015) also noted that open science, by definition, is a *community* effort, rather than a purely individual one. Much research outputs are either unpublished or unusable, because researchers have not been given the capacity, time or incentives to make their research outputs open and accessible (Kucharski et al., 2020).

How to measure open science practices

Summary:

- *In order to measure open science practices, responsible metrics and indicators should be developed.*
- *The starting point for designing responsible metrics for open science should be defining the desired objectives and outcomes for open science.*
- *There is no one perfect solution due to different purposes of evaluation and practices in different fields.*
- *Ideally, next generation metrics should reflect the elements of the FAIR principles. They should be open, freely used, shared, interoperable, available in human and machine readable format, inclusive, and multidimensional.*
- *The benefits and consequences of using metrics for open science in researcher assessments should be evaluated regularly.*
- *Academic assessment changes have the potential to give new insights into open science results, practices and expertises.*
- *Indicators for open science practices should incorporate and build on existing indicators, and likely move away from some traditional metrics such as journal-level indicators.*

This section focuses primarily on recent developments for measuring open science, and briefly touches on open data metrics. Much of the discussion of how to measure is focused on (i) the development of indicators and increasingly (ii) how to develop and implement those metrics and indicators responsibly.

Broad approaches/recommendations

One of the most extensive and relevant sources reviewed here results from the work of an expert group commissioned by the EC, which resulted in the publication of a report on “Next Generation Metrics” (NGM - Wilsdon et al., 2017). This work builds on Herb’s recommendations for *open metrics* (Herb, 2016) and other calls for responsible metrics to propose recommendations for creating indicators for open science. NGM makes broad recommendations which include that an open science system should be grounded in a mix of expert judgement, quantitative and qualitative measures. It also emphasises that transparency and accuracy are crucial and advocates for making *better use of existing metrics for open science (i.e. usage, altmetrics, collaboration measurements).*

The call to apply, build on and modify existing metrics to measure open science is also mirrored in the Amsterdam Call for Action on Open Science (Dutch Ministry of Education, Culture and Science, 2016). However, both of these documents, as well as numerous others we reviewed, encourage moving away from existing journal-level metrics, as a proxy for individual quality, in particular the JIF (see for example European Commission, 2018b; Morais and Borrell-Damián, 2018; Ayris et al., 2018).

Building on the existing scientific practices of academic citation, the Next Generations Metrics report recommends using citations for every scientific output, highlighting that being able to do this depends on developing corresponding infrastructure and assigning PIDs, which is also mentioned in the previously mentioned Amsterdam call for Action on Open Science (Dutch Ministry of Education, Culture and Science, 2016). Other work (Pampel and Dallmeier-Tiessen, 2014) proposes not just usage of citations, but also including some type of “sharing factor” to indicate how much researchers share information for the good of society.

The starting point for designing responsible metrics for open science should be defining the desired objectives and outcomes for open science. As metrics are developed, there should also be a program of meta-research about the indicators to assess their likely benefits and consequences as they are applied in evaluation and also to identify unintended biases and consequences (Wilsdon et al., 2017; European Commission, 2018b).

It is recognised in the literature that a one-size fits all approach for measuring open science will not be adequate. The Amsterdam Call for Action (Dutch Ministry of Education, Culture and Science, 2016) acknowledges that there are different purposes of evaluation and correspondingly different definitions for what the ‘right’ criteria are. For instance, the HuMetricsHSS initiative[48] considers the specific positions of the humanities and social sciences (HSS) and their

[48] https://humetricshss.org/

scientific outputs in the existing research assessment policies. Moreover, Agate and colleagues (2020) claim there are discrepancies between what should be valued in HSS disciplines and what is currently assessed and rewarded in practices. Some HSS scholars try to adapt to the current academic assessment process by engaging in citation gaming practices, e.g. by insisting that their Ph.D. students cite them in every work.

There is also ra ecognition that openness is a continuum of practices, rather than an all-or-nothing situation, which could be thought of as being composed of different levels, where the most basic level could include self-archiving post-prints and the highest level openly sharing grant proposals, research protocols, and data in real time (McKiernan et al., 2019).

Metrics for open data

Lowenberg and colleagues (2019) proposed a vision for developing open data metrics (ODM). ODM are open, freely used, shared and built on by anyone. This vision also includes recommendations for metrics to be inclusive (applicable to all disciplines, communities all data types); to be expressed in machine-readable format; to be transparent; interoperable; and multi-dimensional, e.g. by not conflating multiple dimensions into a single number. Although not specific to data, the Snowball Metrics[49] initiative similarly calls for more open metrics, where calculations and recipes are made publicly available.

The Make Data Count[50] project (MDC) makes further recommendations for creating metrics for data. The initiative brings together representatives from DataCite, DataOne, and Crossref to address the significant social as well as technical barriers to the widespread development of open data metrics. The members of the initiative believe the scientific community values data citation, data usage, and data download statistics more than they value the metrics focused on social media. They also see the need for bibliometric studies about data citation and reuse to understand current practices, and work to help repositories develop services to normalize data usage and citation counts.

The need for developing open metrics in general is also highlighted in the Next Generation Metrics report. The report's authors stated whenever closed/proprietary metrics are used, they should be accompanied by open

---

49 https://snowballmetrics.com/
50 https://makedatacount.org/)

metrics (Wilsdon et al., 2017). Others also highlight the dangers of being locked in proprietary systems, which could potentially exacerbate platform capitalism, where various big players position themselves to package together all the functions including the calculations of metrics into proprietary platforms (Mirowski, 2018).

Any evaluation models that are implemented, as LERU (Ayris et al., 2018) recommends, need to be monitored, reflected upon and updated as needed.

Konkiel (2020) also argued for the responsible development of indicators for data in research, using the Leiden Manifesto as a guideline for proposing how indicators should be developed and applied. Among other points, she suggested that quantitative and qualitative evaluations should be used; expert opinion is especially needed to evaluate the quality of research data. She also highlighted the need to account for variation by field in publication and citation practices; to document those differences in data practices; to scrutinise indicators regularly and update them, and to keep data collection and analytical processes open. In addition to the Leiden Manifesto, Konkiel argued that understanding the strengths and limitations and biases of the data sources used to create indicators is needed.

Developers of indicators/metrics need to be cognizant that raw counts do not have much meaning out of context. In an analysis of biomedical data reuse, Federer provides an example - data about common health issues receive more requests compared to more rare diseases. Federer states that it could be argued that rare disease data is in fact more *valuable* than common disease data (Federer, 2019).

Frameworks, toolkits for measuring open science

There are some more recent developments to develop frameworks and indicators for open science activities as a whole (not just for data), which take context into account, as well as the idea of different purposes and levels for evaluation (Duarte et al., 2016)

The Open Science Career Advancement Matrix (OS-CAM) could serve as one such toolkit, or stand on its own as an evaluative framework for researchers (O'Carroll et al., 2017). This framework provides concrete examples of assessments for different types of open science activities across phases of research. This framework could be used in various contexts (e.g. grant evaluations, job applications, funding models) and at different levels of expertise (e.g., “learning about open science” for first stage researchers or “doing open science” for recognised researchers). The authors note that the weighting of different possible criterion (see Annex 4, Appendix 1) should reflect the background of the researcher being evaluated. The OS-CAM proposes rewarding data reuse.

Wouters and colleagues (2019) proposed a system of “indicator frameworks” and “toolkits” to guide indicator development for different evaluative purposes: *monitoring*, *learning,* and *resource allocation,* all in *descriptive* as well as *comparative* ways. The process for the development of indicators depends on the level of assessment, e.g., for researchers, research groups, institutions or systems. For instance, assessments of individual researchers, which could be used in career evaluations, should include career-oriented qualitative and quantitative indicators, based on the principles of responsible metrics. Academic assessments could adopt some existing indicators, but should be wary of using other indicators, e.g., journal-level indicators, which may not be as applicable to the range of open science practices.

How to embed open science in researcher assessments

Summary:

- *Dimensions of open science should be embedded in recruitment processes, career advancement, funding criterias, etc.*

Many documents, policies and resources recommend embedding open science dimensions in processes for recruitment, career advancement and evaluation (e.g. European Commission, 2018b; Ayris et al., 2018).

The Sharing Rewards and Credit Interest Group[51], e.g., was established within the Research Data Alliance to work towards improving research crediting and rewarding mechanisms for researchers who want to organise their data and other research objects for community sharing. This group suggested changing policies and evaluations to give credits to data owners in order to motivate them to make the effort to share their data.

Recommendations made by LERU for RPOs specifically (Ayris et al., 2018) call for including open science factors in job descriptions, performance appraisals and

[51] https://www.rd-alliance.org/groups/sharing-rewards-and-credit-sharc-ig

promotion criteria which account for the multiple roles and responsibilities that researchers have. These recommendations suggest including open science research outputs, research processes, teaching and supervision activities, leadership, service to various groups (university, community, public), etc. in the process of hiring and promotion. Many of these different responsibilities are aligned with the OS-CAM framework discussed previously. There are also calls for RFOs to include open science aspects in grant evaluations and to reward open science activities with the "highest degree of visibility" (O'Carroll et al., 2017).

The Amsterdam Call for Action on Open Science (Dutch Ministry of Education, Culture and Science, 2016) also calls for addressing open science aspects in academic assessment at an international and national levels, encouraging national and international actors to acknowledge that national initiatives are reaching their limits, and that this is an area which needs a harmonised approach at the EU level.

Novel ideas for embedding open science activities into the evaluation of researchers have been proposed at different levels. The LERU recommendations, for instance, suggest that RPOs should create institutional catalogues listing *where* researchers have published data and work to develop ways to embed this information into assessments (Ayris et al., 2018).

Moreover, there are some novel ideas for promotion of open science at the journal level also. The Transparency and Openness Promotion Guidelines, e.g., is a set of eight standards with different levels of achievement to help journals promote and eventually reward open practices (Nosek et al., 2015). These standards promote, among other things, citation standards, data and code transparency. Although these standards are geared toward journal policy, they could potentially be applied to different situations and at different levels (RPO, RFO).

There are also new proposals for embedding open science evaluations at funding levels. Eisfeld-Reschke and colleagues (2014), for example, suggest that while prominent funding agencies have already embraced single elements of evaluating open science, new approaches could also be considered. They propose "Research Funding 2.0" and wonder whether existing "bureaucratic models" of delegating responsibility in funding could be replaced or augmented by involving researchers and/or citizens in the process.

Embedding open science in academic assessment requires changes at all levels, and by all stakeholders (Munafò et al., 2017). Researchers should also be engaged in creating "assessment criteria and practices, enabling researchers to exactly understand how they will be assessed and that open practices will be rewarded" (Nosek et al., 2015).

3.4. FAIR DATA EVALUATION

3.4.1. FAIRNESS OF RESEARCH DATA

FAIR data / digital objects and the accompanying FAIR ecosystem

Summary:

- *The FAIR principles apply not just to data, but also to other research objects, e.g., software, protocols, and other research resources and outputs.*
- *FAIR assessments must "go beyond the object itself" and include evaluations of other elements of the FAIR ecosystem including frameworks and infrastructures.*
- *The understanding of FAIR and importance of FAIR data is not at the same level for researchers and open science experts.*
- *There are some signs that FAIR data is moving beyond implementation to the adoption phase, and is becoming common practice within certain disciplines.*

According to the FAIR principles (Wilkinson et al., 2016), research data should be Findable, Accessible, Interoperable and Reusable. Although this set of principles is primarily defined for research data, there is some discussion and proposals for applying these principles to other outputs and infrastructures for research objects[52,53] (Goble et al., 2020).

The European Commission Report, Turning FAIR into Reality (TFiR) (Collins et al., 2018) undergirds many of the policies and the development of tools which has been undertaken in the context of EOSC, including work within the FAIRsFAIR project. This report emphasises that the FAIR principles apply not just to data, but also to other digital objects (DOs) such as software, protocols, other research resources and outputs. All FAIR DOs should follow FAIR guidelines, and "be

[52] https://eurocris.org/activities/conferences/cris-2018
[53] https://leidenmadtrics.nl/articles/publications-should-be-fair

accompanied by PIDs and metadata rich enough to enable them to be reliably found, used and cited" (Collins et al., 2018).

The TFiR report also recognizes that FAIR DOs can only exist in an accompanying ecosystem of FAIR services, metadata schemas, data stewardship, and policies. This recognition is also reflected in work by the FAIRsFAIR project to develop metrics to assess FAIR. This work suggests that *FAIR assessment must "go beyond the object itself"* and include evaluations of other elements of the FAIR ecosystem, such as enabling services and data repositories, which are vital to ensuring the longevity and sustainability of objects. This need to assess the FAIRness of the entire network (beyond an object, including repositories and services) is echoed in the document FAIR Ecosystem Components: Vision (L'Hours and von Stein, 2019).

Some take the concept of FAIR DOs and apply it also to computational workflows. Goble and colleagues (2020), for example, see these workflows as being themselves "research objects" which could help to further enhance open science if they were made findable, accessible, interoperable and reusable. The authors also state that FAIR computational workflows rely heavily on supporting infrastructures, which are not yet in place.

Interoperable frameworks and infrastructures supporting and defining community "practices for data sharing, data formats, and metadata standards" are key to making a FAIR ecosystem a reality (Collins et al., 2018).

Although there is much buzz about the FAIR principles throughout the reviewed literature, there appears to be some discrepancy between how the uptake of FAIR is understood by researchers and open science experts. According to a survey of researchers in the State of Open Data report (Fane et al., 2019), the FAIR principles are relatively unknown amongst researchers, who have more familiarity with the concept of open data. A study by JISC (Allen and Hartland, 2018) also reported inconsistent views among researchers on "what adherence to FAIR means in practice and how to evidence it" and a lack of awareness/practice in making data machine-readable.

Among the expert group working to implement open science principles across Europe (4 year project), however, there is a feeling that FAIR data is moving beyond implementation to the adoption phase, and is becoming common practice (Jacobsen et al., 2020). They note, however, that this awareness varies among disciplines.

Assessing FAIRness

Summary:

- *Manually or automatically evaluating the FAIRness of objects is called for by multiple stakeholders for various purposes.*
- *There are examples of proposed incentives and rewards for data sharing or data management without evaluating the FAIRness of data.*
- *There are many challenges in the evaluation of the FAIRness of data and there is no standard approach. FAIRness is not a binary measure (either/or).*
- *Not all elements of the FAIR principles can be globally assessed. Some need to be locally defined within disciplinary communities.*
- *Some FAIR metrics are related to the object itself, and some for the infrastructure and other objects in the collection.*

Why is the assessment of the FAIRness of digital objects or data needed? Often the answer lies in the need to assess/monitor the components of the ecosystem that will fulfill promises of data science (and possibly of open science).

Authors of the document FAIR Ecosystem Components: Vision (L'Hours and von Stein, 2019) state that multiple audiences need to assess the FAIRness of objects, and that multiple forms of assessment from self-assessment on the part of researcher to certification (e.g., on part of repositories) are necessary. They highlight that *funders*, as well as researchers, data stewards, and repository managers, need the results of both *automated and manual evaluation* of the FAIRness of objects. The authors also emphasise that **not** all elements of FAIR can be globally assessed, but rather need to be locally defined. Certain FAIR guidelines (particularly regarding reusability - i.e. the richness of metadata descriptions, accuracy and relevance), can only be understood against the background of local or disciplinary norms and expectations.

Using assessments of FAIR data in researchers' evaluations is called for, but not always done explicitly. More common wording is to reward researchers for practices that make FAIR data - e.g., data sharing and producing data management plans. FAIR is, however, mentioned explicitly at times in regards to researcher evaluations, but often vaguely. Koers and colleagues (2020) recommend that funders and institutions should "consider the *level of FAIRness* and data sharing as part of research assessment, among other criteria". The

Open Science Policy Platform Recommendations (European Commission, 2018b) include a recommendation that RFOs and RPOs should "*give credit for FAIR data* resulting from research work" and that publicly funded research *must make data FAIR*, citable and as open as possible.

This implies a need to measure/assess whether or not research outputs have been made FAIR or are in compliance with the FAIR principles. This is tricky and raises questions and challenges. When should such evaluation occur - at the end of the process, when an object is certified as "FAIR" or throughout the process, at different levels of gradation, as FAIRness is also a matter of degree[54,55], rather than a binary either/or category? How does/would the entire ecosystem necessary to achieve FAIRness affect the evaluation of individual researchers? How should disciplinary differences and differences in data be evaluated? What is evaluated - the *practice of the researcher*, or the *FAIRness of the object* produced?

Bishop and Hank (2018) presented a method for qualitatively measuring the FAIR Data Principles through operationalising findability, accessibility, interoperability, and reusability from a re-user's perspective, meaning researchers who reused data.

On a GitHub repository[56] accompanying the article "Evaluating FAIR maturity through a scalable, automated, community-governed framework" (Wilkinson et al., 2019), the authors have posted a *"philosophy of FAIR testing"*. This philosophy explicitly states that FAIR will have different requirements for different communities, that certain indicators/tests *may not be applicable to certain resources from a given community and that proposed indicators may need to be more specific for particular communities.* It also states that *"FAIR evaluations are not intended to be used as "judgement"*, but rather as a means to objectively (AND TRANSPARENTLY!) test if a resource has successfully fulfilled the FAIRness requirements that that community has established."

Other challenges also exist when considering developing FAIR metrics. FAIR has become so ubiquitous, that it is being interpreted and implemented in many ways, often outside of the context of the original focus on machine readability

---

[54] See https://github.com/FAIR-Data-EG/consultation
[55] Also see https://content.iospress.com/articles/information-services-and-use/isu824
[56] https://github.com/FAIRMetrics/Metrics/tree/master/MaturityIndicators; Current as of April, 2021

(see Jacobsen et al., 2020). However, measuring reusability in particular is difficult, as some aspects of the principles detailed under “R” “require human mediation and interpretation” (Devaraju et al., 2020). Assessing FAIR (as mentioned before) also involves assessments of an entire ecosystem, not just an object itself (Devaraju et al., 2020).

Examples of FAIR metrics

Summary:

- *Various attempts to define metrics for assessment of FAIRness of data are under development, but it seems there is no one widely adopted set of metrics at the moment.*
- *Measuring FAIR is still in early days, further development is needed and expected.*

In the last 2-3 years, various groups (FAIRsFAIR, some of original authors of the FAIR principles, and RDA) have begun work to develop metrics to assess the FAIRness of objects. As the RDA group stresses, regarding their own efforts: The question now is really on *what* aspects to measure rather than *how* to measure them (RDA FAIR Data Maturity Model Working Group, 2020). This being said, they (and others) propose various tests - automated and manual - to assess the FAIRness of objects (see Devaraju et al., 2020; Wilkinson et al., 2019).

The European Commission Expert Group on FAIR Data published a report Turning Fair into Reality in 2018 (Collins et al., 2018). The authors recognise the challenges involved in assessing FAIRness and calls for an expert group to develop next generation metrics that could support open science, and the development of a new set of metrics to develop the FAIRness of objects. In their vision (Collins et al., 2018), such metrics “will need to be defined by research communities based on their disciplinary interoperability frameworks for FAIR sharing.” Basic minimum standards for measurement are, however proposed: objects should have “discovery metadata, persistent identifiers and access to the data or metadata.” The authors also include a plea for responsibility in developing these metrics - “The development of FAIR metrics will need to be extremely mindful of the usually unintended – but all too often negative – consequences and behavioural shifts that result from the introduction of metrics.”

FAIRsFAIR proposes 15 metrics to evaluate the FAIRness of research data and other DOs in Trustworthy Digital Repositories (Devaraju et al., 2020). They include a description of each proposed metric; map each metric to the relevant FAIR principle; and align it with Core Trust Seal certification requirements. Although the metrics are designed to assess generic, cross-disciplinary metadata standards, e.g. DataCite, Dublin Core, schema.org, multiple possible example methods for assessing each metric have been provided as well. They provide recommendations for metrics both at the data and metadata level.

General themes in the recommendations include checking for globally unique, persistent identifiers - for both data and metadata; machine readability; persistence (i.e. of metadata if data is no longer available); standardisation of metadata (ideally following standard of “target research community), links between metadata, data, and related entities. Measuring reuse focuses on checks for assigning the appropriate license for the data, but also that metadata includes provenance information about data creation or generation (Devaraju et al., 2020).

The FAIRsFAIR project has also experimented with developing automated ways of assessing FAIRness. For instance, the F-UJI tool[57] is a beta web service developed to programmatically assess the FAIRness of research data objects.

Wilkinson, the lead author on the original FAIR principles, also led other efforts to develop frameworks for possible metrics (Wilkinson et al., 2018; Wilkinson et al, 2019). One of those frameworks defines “maturity indicators”. “Maturity indicators describe facets of FAIRness that can be objectively evaluated by a machine” and thus used to establish a “contract of expectations and capabilities between a data resource and an automated agent.” (Wilkinson et al, 2019)

Another interpretation of the term maturity indicator, although not the one stated by Wilkinson in his work, could be a recognition that FAIRness is a process and exists on a continuum of behaviours[58]; also mentioned by FAIR authors (Mons et al., 2017).

The design framework (Wilkinson et al., 2018) represents the first iteration by FAIR authors to develop a core set of FAIRness indicators which can be measured by semi-automated processes. This framework provides a template to

---

[57] https://github.com/FAIRsFAIR/fuji
[58] https://github.com/FAIR-Data-EG/consultation

be used within individual communities to derive *community-specific FAIR metrics. Responsibility* - recognises that any derived FAIR metrics are not an indication *impact*; and that any metrics that assess “popularity” of an object are not measuring its FAIRness.

Other aspects of this framework of interest in regards to developing responsible metrics, include: “1) Metrics should address the multi-dimensionality of the FAIR principles, and encompass all types of digital objects; 2) Universal metrics may be complemented by additional resource-specific metrics (at community level); 3) Metrics themselves must be FAIR; 4) Various approaches to FAIR assessment should be enabled (e.g., self-assessment, task forces, crowd-sourcing, automated); but these approaches need to be scalable; 5) Governance over the metrics, and the mechanisms for assessing them, will be required to enable their careful evolution.” (Wilkinson et al., 2018)

Wilkinson and colleagues (2019) expand this work, and details a series of (still developing) maturity indicators. These indicators aim for providing fully automated ways of assessing FAIRness and cover most of FAIR principles/guidelines, *with the exception of R1.2 ((meta)data include detailed provenance) and R1.3 ((meta)data meet community standards).*

A working group within the RDA has also developed a “FAIR data maturity model,” a list of indicators, ranked by priority, and possible evaluation methods. Although this list of indicators has much overlap with Wilkinson lists, there are also differences. This one includes separate indicators for data and metadata. It also includes more indicators regarding reusability, including R1.2 and R1.3. The “human” is also more present in this list of indicators. It includes some indicators to evaluate human access (i.e. RDA-A1-02M: Metadata can be accessed manually.) No indicators for interoperability are assigned the essential priority level.

Examples of FAIR self-assessment tools

Summary:

- *FAIR self-assessment tools can be used by researchers to better understand the FAIR principles and to improve FAIRness of their own research data.*

Other stakeholders have been working to develop self-assessment tools, which researchers can use as aids to assess/improve the FAIRness of their own data. The Data Stewardship Wizard (Pergl et al., 2019) is composed of a series of primarily multiple-choice questions to aid in the creation of data management plans. The tool clearly indicates alignment with the FAIR principles as the researcher proceeds to answer the questions.

Another recent tool, developed in the FAIRsFAIR project, is FAIR-Aware[59]. This also consists of a series of questions, not specific to creating DMPs, to assist researchers in assessing the FAIRness of their data before uploading it to a repository. The stated goal of the tool is to help researchers better understand FAIR, and it emphasises that the tool "is **not** meant to give you a score for the FAIRness of a specific dataset."

Examples of FAIR ecosystem tools and certifications

Summary:

- *FAIR digital objects, and their assessment, do not exist in isolation. There are developments assessing the trustworthiness or FAIRness of digital repositories and standards.*

FAIRsharing[60] is a registry of FAIR standards, databases and policies and the links between them. These resources are themselves FAIR. It is a carefully (manually) curated resource and aims to promote use of *existing resources* and standards in the FAIR ecosystem.

The Core Trust Seal[61] offers a core level certification for data repositories based on the DSA–WDS Core Trustworthy Data Repositories Requirements catalogue and procedures. Building on work to develop the Core Trust Seal, the TRUST principles (Lin et al., 2020) for digital principles have been proposed as a way for repositories to "establish transparency, responsibility, user focus, sustainability and adequate technology." The arguments for assessing the trustworthiness for repositories in large part follow the arguments for the responsible development of research metrics - i.e. the transparency of policies and mission. The principles also call for the inclusion of data metrics in repositories, interestingly as a way to demonstrate user focus.

[59] https://fairaware.dans.knaw.nl
[60] https://fairsharing.org
[61] https://www.coretrustseal.org/

3.4.2. OPENNESS/FAIRNESS OF DATA USED IN RESEARCH EVALUATIONS

Summary:

- *Criteria and data used in research evaluations should be transparent and open.*
- *The focus in transparency of data used in research evaluations is usually on making the criteria and methods used in evaluations open, including transparency in indicator development. Fewer documents call for making the data used in evaluations open.*

There are many calls, both within the open science and researcher assessment literature which we reviewed, to make the *criteria* used in evaluating researchers transparent and open (Working group for responsible evaluation of a researcher, 2020; Hicks et al., 2015; Wilsdon et al., 2015; Global Research Council, 2018). Often, when policy documents mention data - the focus is on open criteria behind academic assessments and including transparency in indicator development.

The Recommendation for the Responsible Evaluation of a Researcher in Finland (Working group for Responsible Evaluation of a Researcher, 2020), e.g., recommends providing public, accessible information about how researchers are evaluated, stating that *"the data, metadata and methods that are relevant to research evaluation, including but not limited to citations, downloads and other potential indicators of academic re-use, should be publicly available for independent scrutiny and analysis by researchers, institutions, funders and other stakeholders."*

In recommendations for RPOs, LERU (Ayris et al., 2018) also argues for transparent communication of policy and proper training of evaluators, calling for the development of "*institutional policies for recognising and rewarding open science practice* anchored in broad-based support; *communicate them clearly and transparently, make them easy to find and access, and provide proper guidance* or training to those who are involved in staff recruitment, appraisal and promotion in the university."

The calls for open metrics (e.g. Lowenberg et al, 2019; Wilsdon et al., 2015; Herb, 2016) highlight that open metrics go beyond altmetrics in that open metrics and adhere to the following:

1) research products and data sources for metric development need to be logically selected, *openly documented,* and chosen in line with the disciplinary norms;
2) *the data that underlies metrics, indicators, and measurements need to be open and accessible (preferably via automatic processes, e.g., API);*
3) open software - provision of software that was used for calculations;
4) logical, scientific, and documented *explanation of how data were derived and metrics were calculated*.

Laine (2018) defines open science evaluation as both "the notion that the evaluation processes and the data used to inform the evaluations should be transparent, as well as to broadening the range of things that are considered to be of merit, such as data outputs." In her analysis of multiple international research codes of conduct, she found that *"None of the codes mentions openness in the context of research evaluation methods and processes, be this for transparency of evaluation criteria, open metrics, or open peer review. Only the European code encourages a plurality of recognised research outputs in the evaluation*. What the codes do recognise as an openness-related issue is the need for transparency concerning possible conflicts of interest in evaluation."

3.5. TECHNICAL GROUND FOR FOSTERING OPEN SCIENCE AND RESPONSIBLE RESEARCH ASSESSMENT

Summary:

- *The development of open science and the move to responsible research assessment requires research eInfrastructures - software tools, platforms, databases, information systems, etc.*

In 2010, the European Commission's expert-group concluded that "assessment of university-based research is being hampered by a lack of reliable, comparable, and comprehensive data", and recommended that the European Union should "invest in developing a shared information infrastructure for relevant data to be collected, maintained, analysed, and disseminated across the European Union" (European Commission, 2010). In the same vein, a report to the European Parliamentary Research Service (Mahieu et al., 2014) recommends "the development of a European integrated research information system", "having features of a distributed infrastructure, inter-connecting the existing national research information systems”, facilitated by CERIF (the European standards for research information systems).

Also responsible assessment of researchers fully-recognising open science outputs and practices requires reliable, comparable, and comprehensive data, as well as integrated research eInfrastructures (eRIs) - software tools, platforms, databases, information systems, which provide assessment data on the whole range of research activities. The aim of this section is to identify the main limitations and gaps of available eInfrastructures (including data models, formats, and vocabularies), especially from the point of view of recognising open science in research assessment.

There are some overall resources for listing and comparing research eInfrastructures:

- Research networking tools (RN tools).[62] “RN tools serve as knowledge management systems for the research enterprise. RN tools connect institution-level/enterprise systems, national research networks, publicly available research data (e.g., grants and publications), and restricted/proprietary data by harvesting information from disparate sources into compiled profiles for faculty, investigators, scholars, clinicians, community partners and facilities”.
- Directory of Research Information System (DRIS)[63] is a euroCRIS initiative to map the available research information management infrastructure in Europe (and beyond).
- Scholarly Communication Technology Catalogue (SComCAT)[64]. SComCAT is a catalogue of scholarly communication technologies with the potential to foster the development of open science.
- SPARC Europe report Scoping the Open Science Infrastructure Landscape in Europe (Ficarra et al., 2020) provides an overview of 120 relevant Open Science Infrastructures from 28 European countries.
- The Framework for Open Science and Research (The Open Science and Research Initiative, 2016) also maps application services needed to support various stages of the (open) research lifecycle

62 https://en.wikipedia.org/wiki/Comparison_of_research_networking_tools_and_research_profiling_systems
63 https://eurocris.org/services/dris
64 https://www.scomcat.net/

3.5.1. RESEARCH INFORMATION

Summary:

- *International and national, general and domain specific data models and research information platforms built on top of them might be used for preservation of research information including research assessment indicators and recording credits and rewards given to researchers due to a contribution to open science.*

In this section we analyse international and national, general and domain specific data models and information systems which might be used for preservation of research information including preservation of research assessment indicators and recording credits and rewards given to researchers due to a contribution to open science. Moreover, we also analysed formats of data models including human readable and machine readable formats. Those formats could improve discoverability of information, and enable interoperability of systems which might lead to integration of the research landscape at European level. Besides machine readability of format of a data model, one important aspect of interoperability capacity of a data model is standardization of vocabularies used in the model.

Research information model

Summary:

- *A standardised research information model, interchangeable format and vocabulary could enable exchange of research metadata across Europe, which might lead to European central database of research metadata and much more transparent and aligned evaluation of researchers' achievement.*

The Common European Research Information Format (CERIF)[65] is a flexible data model for representing information about research ecosystem entities (Figure 2). It was developed and is being maintained by euroCRIS and is recommended by the European Commission to the Member States bodies responsible for collecting and preserving information on Research and Technological Development activities at national levels. This model could enable exchange of research metadata across Europe, which could support integrated European eInfrastructure of research metadata and a more transparent and comprehensive assessment of researchers' achievement.

[65] https://github.com/EuroCRIS/CERIF-DataModel

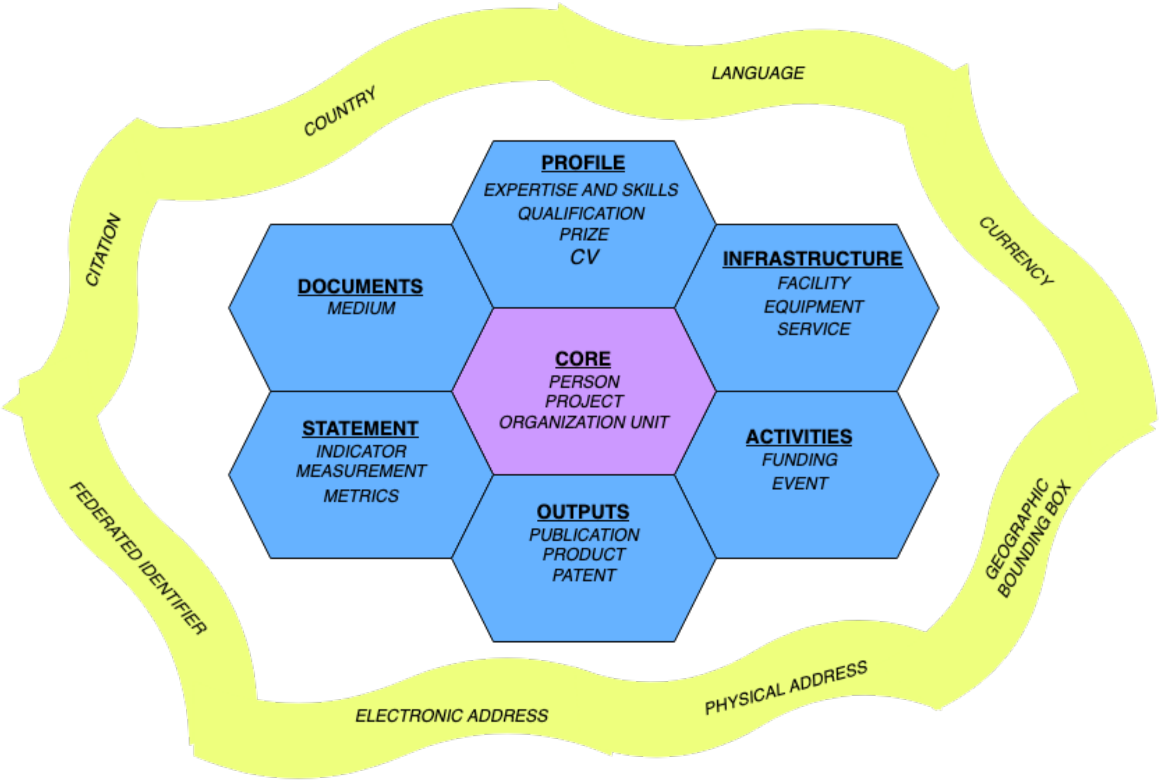


Figure 2. Representation of the CERIF model[66]

Some of the key features of this model are support for multilingualism, expressing time-framed (the temporal dimension) relationships, storing semantics of entities and relationships, usage of global persistent identifiers. However, modularity and extensibility at its core should be enabled to make decentralisation of the work on the standard and tools easier, and it is planned for the ongoing CERIF refactoring project[67].

Common format for exchanging CERIF data is XML, although there are some examples of usage of CERIF in a semantic web format such as RDF[68]. There is a semantic vocabulary associated with the CERIF[69]. This standardised thesaurus could enable full interoperability at metadata level across the European research eInfrastructures. Moreover, standardization of the vocabulary is a necessary step in the creation of a machine readable format. The vocabulary is implemented using the CERIF semantic layer in accordance with the Simple Knowledge Organization System (SKOS) W3C recommendation. SKOS is designed for any type of structured controlled vocabulary enabling easy publication and

[66] https://www.eurocris.org/services/main-features-cerif
[67] https://www.eurocris.org/cerif-refactoring-project-introduction
[68] https://vre4eic.ercim.eu/
[69] https://github.com/EuroCRIS/CERIF-Vocabularies

use of such vocabularies as linked data. The existence of this vocabulary is in accordance with the FAIR I2 principle: "I2. (Meta)data use vocabularies that follow FAIR principles". However, as for all vocabularies, and especially taking into account recent open science development, there is a question whether this vocabulary is comprehensive and includes all necessary terms at the moment, and how often it should be extended with new terms which might appear in the near future.

Platforms for implementation of research information systems

Summary:

- *There are examples of local, institutional, national and regional research information platforms built in-house as a home-made solution or built by adoption and customisation of an existing software solution.*

Research Information platforms can be developed at local, institutional, national, regional or European level. The solution can be built in-house as a home-made solution or some existing solution such as DSpace-CRIS, Pure, Converis, Sympletic Elements can be adopted and customised.

- DSpace-CRIS[70] presents an open-source CRIS (Current Research Information System) solution built by extension of DSpace, a well-known digital repository . Additionally, it can also be integrated with the CKAN platform for preserving research data. Moreover, this platform is integrated with some altmetrics and services (Altmetrics / Digital Science, PlumX / Elsevier), meaning it is possible to assess the impact of some research object.
- Pure[71] is a well known Elsevier CRIS solution used by more than 250 institutions across the world. It is possible to automatically feed data from Pure to the organization's digital repository. Pure supports integration with DSpace, ePrints, FEDORA and Equella digital repositories. Those digital repositories can preserve publications, data, and other research digital objects.
- Converis[72] is a Clarivate CRIS solution which integrates the management of research information and internal workflows, bringing together data from external and internal sources into a single platform. Converis can be

[70] https://www.4science.it/en/dspace-cris/
[71] https://www.elsevier.com/solutions/pure
[72] https://clarivate.com/webofsciencegroup/solutions/converis/

used to validate processes for open access digital repository and for bibliometric analysis, i.e. the platform could be used as a gateway to open research objects repository and to measure research objects impact using bibliometric indicators.

- Symplectic Elements[73] is a Digital Science CRIS solution which is a comprehensive and powerful scholarly information software solution which empowers organisations to collect, analyse, showcase and report on all of their scholarly activities. Some of the key features of this solution are automated data harvesting, rich researcher profiles, assessment and reporting functionality and digital repository integrations.

An example of a national CRIS is the Finnish national research hub[74] which provides all Finnish resources for evaluation and monitoring of science in one place. The platform improves the finding of information and experts on research and increases the visibility and impact of Finnish research. The platform does not require entering information manually and therefore does not increase the time researchers spend on reporting or other administrative work. The service contains descriptive information that makes it easy to find, for example, openly available publications and materials. Each reporting organisation is responsible for the accuracy, quality, completeness and timeliness of its own data. At the moment information about Publications, Projects, Infrastructures, Organisations, Science and Research News are collected, while integration with ORCID and Universities' CRIS systems is under implementation, as well as integration with Fairdata[75], the Finish national research data platform. A national Finish strategy for supporting open science, recognition of open research data, responsible research assessment and so on could be applied starting from the Finnish national research hub.

The Integrated Semantic Framework ontology modules for VIVO (the VIVO-ISF ontology)[76] provide a set of types (classes) and relationships (properties) to represent researchers and linked entities of the research ecosystem. The VIVO-ISF ontology incorporates elements of several well-known ontologies, including Dublin Core, Basic Formal Ontology, Bibliographic Ontology, FOAF, and SKOS. The ontology can be used to describe several roles of faculty members, including

[73] https://www.symplectic.co.uk/
[74] https://research.fi/en/
[75] https://www.fairdata.fi/en/
[76] https://wiki.lyrasis.org/display/VIVODOC110x/Ontology+Reference

research, teaching, and service. Semantic web application (VIVO)[77] is built on top of this ontology. When a VIVO instance is populated with researcher interests, activities and accomplishments, VIVO enables discovery of research entities. VIVO encourages research discovery, expert finding, network analysis and assessment of research impact. However, some types of bibliometrics which can be used for research assessment are not provided in the VIVO tool (e.g., citation counts). Moreover, there is no concept/entity representing research data in VIVO-ISF, although there are some efforts that have been extending the VIVO-ISF ontology to represent research impacts and objects including research datasets. VIVO highlights teaching expertise/responsibilities which is often a missing information in systems and platforms for monitoring researchers activities.

The "Core Data Set Research" for the German science system[78] is a specification of standards that are intended to enable the harmonized recording and storage of research information across research institutions in Germany. KDSF is the official Germany acronym, but it is also referenced as RCD in some literature. KDSF contributes to interoperability of Research landscape at national (Germany) level and contains metadata about publication and researchers. Unfortunately, there is no support for metadata about some research objects such as research data.

Research information systems and open science

Summary:

- *There is a trend of integrating research information systems with publication and data repositories which could, besides improving capacity for monitoring and evaluating the research landscape, promote open science.*

Current Research Information Systems (CRIS) can help in monitoring of science and evaluation of research objects representing results of research. If a CRIS is integrated with publication repository and data repository, then the CRIS could also be used for exploratory purposes, could be a part of a marketing of an institution, could promote open science and boost impact of researchers and institutions. The survey “Practices and Patterns in Research Information Management” (Bryant et al., 2018) shows a trend of integration of CRIS and institutional repositories (IR) which supports a more transparent reporting of results of funded projects and leads institutions toward open access goals. The

[77] https://github.com/vivo-project/VIVO/releases/tag/vivo-1.10.0
[78] https://www.kerndatensatz-forschung.de/

logical consequence is that we can expect a similar direction in the next period for integration of CRIS-IR with data repositories which will enhance wider uptake of the open science paradigm. For instance, there is an initiative for integration of data management plans with institutional systems at TU Delft[79]. Moreover, Tanhua and colleagues (2019) suggest an integrated system of data management rather than inventing new systems to address challenges. Integration should be based on PIDs. Schöpfel and Azeroual (2021) discussed the benefits and risks of merging CRIS and IR into one system. They found that IRs often fulfill the requirements of monitoring and assessment of institutional research performance while CRISs, beyond the processing of metadata, begin to store, preserve, and disseminate research papers.

The Science Europe Position Statement on Research Information Systems (Science Europe, 2016) defines recommendations how CRIS systems should be developed in the open science and FAIR data era. The document promotes the development of resilient research information systems by adopting the core principles of flexibility, openness, FAIRness and data entry minimisation:

- Flexibility: Research information systems should be flexible enough to allow for extensions in terms of the data objects covered, their definitions, metadata, and use of external data sources.
- Openness: Research information systems' data should be available for external use – in line with the principle 'as open as possible, as closed as necessary' and EU Directive 2013/37/EU1 – and their processing should never require the loss of ownership in underlying raw data by the originating institution.
- FAIRness: Research information systems should foster the findability, accessibility, interoperability, and reusability of the data that they store by implementing the FAIR Guiding Principles for research activity data.
- Data entry minimisation: Research information systems should minimise the need for entering data and facilitate the reuse of data entered manually in line with the motto 'enter once, reuse multiple times'.

Schöpfel and colleagues (2017) investigated the impact of open science development on further evolution of CRIS systems as one type of eInfrastructure for preserving research information and evaluation of research results and

[79] https://openworking.wordpress.com/2021/02/22/towards-better-efficiency-integrating-data-management-plans-with-institutional-systems/

researchers. The paper contributes to the debate on the evaluation of research data, especially in the environment of open science and open data, and is helpful in implementing CRIS and research data policies.

Sīle and colleagues (2017) conducted a survey which identified and described 23 national databases for research output that are currently in use in Europe. The survey shows a great diversity in bibliographic data collection practices in the different countries. The insights reported here can be used as a starting point for more detailed exploration of designs and contents of databases for publication information across Europe. However, this study was limited to bibliographic information for just one type of research objects - textual publication. Moreover, it was also limited to SSH fields.

A proof of concept of a European publication infrastructure integrating data from six institutions across four different countries has already been carried out in the framework of EU COST-Action ENRESSH (Puuska et al., 2018). Nevertheless, there is still a lot of work to be done to improve the standardization and interoperability of CRIS data to build large-scale international solutions that can compete with commercial bibliometric databases. Puuska et al. (2020) proposed integration of national publication databases as a solution for developing a high-quality and comprehensive information base on scholarly publications in Europe. Pölönen et al. (2020) argue that a large-scale data infrastructure for monitoring Open Science at the European level should reflect Europe's geographic, cultural, and linguistic diversity, and only institutional publication data, integrated from the local CRIS at the national and international levels, can provide the needed comprehensiveness.

3.5.2. RESEARCH OBJECTS

In this section we analyse metadata models for describing research digital objects representing research results such as digital textual publication, research data, software, workflows, etc. Standardised formats for research object metadata are important for further development and uptake of open science, because those standards facilitate the consumption and aggregation of metadata from multiple catalogs. Machine and human readable formats for the description of research objects improve the FAIRness of digital repositories based on this format, and thus improve discoverability of open research objects and foster uptake of open science paradigm. Moreover, we have analysed platforms enabling the preservation and publishing of research digital objects and good practices for making digital repositories, as well as certifications of digital repositories.

Research objects description

Summary:

- *There are numerous metadata formats for cataloguing research objects. These vary in popularity.*

In CERIF, research objects representing research results can be expressed by using Result entities: cfResultPublication, cfResultProduct, cfResultPatent. This means that metadata about research data, software, workflow should be stored using cfResultProduct entity. Examples of usage of the CERIF cfResultProduct for dataset storage can be found in the C4D[80]. Replacement of the ResultPublication, ResultProduct and ResultPatent triple of entities by a hierarchy with a single root (ResearchObject) where the descendants would include the Dataset class, among others is planned for the ongoing CERIF refactoring project[81].

The DataCite Metadata Schema[82] is a list of core metadata properties made in accordance with community standards, such as ORCID, the Open Funder Registry, DOI and Dublin Core. The purpose of this well-known schema is an accurate and consistent identification of a digital resource for citation and retrieval purposes. The schema includes a resourceTypeGeneral set of values including DataPaper, Dataset, InteractiveResource, Software, Workflow, meaning it can fully support the diversity of open science research objects. However, many metadata fields prescribed by this schema are optional (not mandatory) which might lead to metadata incompleteness in DataCite metadata records. Moreover, some fields are ambiguously defined (e.g., type and subtype).

OpenCitations Data Model[83] (OCDM) has been developed to model bibliographic and citation data in the OpenCitations Corpus[84]. It permits the publication of bibliographic and citation data as Linked Open Data in a JSON-LD format,

---

80 https://cerif4datasets.wordpress.com/?s=4+
81 https://www.eurocris.org/cerif-refactoring-project-introduction
82 https://schema.datacite.org/)
83 http://opencitations.net/model
84 http://opencitations.net/

thereby conferring machine readability and interoperability of the research data on the Web. Although OCDM has been developed for the need of OpenCitations Corpus, the OCDM has also been adopted by external projects in recent years.

DCAT[85] is a W3C recommended RDF vocabulary designed to facilitate interoperability between data catalogs published on the Web. The fact that DCAT does not cover semantic relations to organisations, persons, software, projects, or funding could introduce some challenges for research assessment and reporting. Moreover, the Dublin Core Type vocabulary used by DCAT includes classes for dataset, software, and service, but does not include classes for model and workflow.

The DCAT Application profile for data portals in Europe (DCAT-AP) is a DCAT based specification for the description of public sector data sets in Europe. The purpose of this profile is to enable a cross-data portal search for data sets and make public sector data better searchable and discoverable across borders and disciplines.

Research objects preservation

Summary:

- *There are examples of local, institutional, national and regional data repositories built in-house as a home-made solution or built by adoption and customisation of existing software solutions for building digital assets repositories (textual publications, data, software, etc) or software solutions specialised for building digital data repositories.*
- *There are tools helping in making data from data repositories FAIR and reusing data.*

An open-access digital repository can be built using repository platforms such as DSpace[86] and ePrints[87]. Although, DSpace was initially created to be a repository of digital textual publications, it is possible to create a data repository or even mix of all research objects repository[88] using the DSpace platform. The strengths that DSpace has in this area are:

[85] https://www.w3.org/TR/vocab-dcat-2/
[86] https://duraspace.org/dspace/
[87] https://www.eprints.org/uk/index.php/eprints-software/
[88] For example https://www.research-collection.ethz.ch/

- file type agnostic
- no theoretical file size limit, even though there might be limits in other places (OS, underlying software), DSpace itself has no known limit of data size
- flexible metadata schemas, allowing you to align with DataCite and other metadata schema
- DOI integration with DataCite
- different workflows and rules are possible on a per collection basis, giving an excellent starting point for a mixed Publication/Data set repository

ePrints shares many of the features commonly seen in document management systems, but is primarily used for institutional repositories and scientific journals. ePrints can be used for building repositories of publications or repositories of research data[89]. One example of ePrints publication repository is Zurich Open Repository and Archive[90] which is the primary directory of publications by researchers at the University of Zurich and provides access to the full texts.

Both platforms, DSpace and ePrints, implement the OAI-PMH protocol which enables exporting of metadata to repository aggregators such as OpenAIRE.

Previously listed platforms can be used for setting up digital resource repositories, but there are open source data management solutions which can be used by data publishers - national and regional governments, companies and organisations - for making their data open and available. Two popular and easily customized data management solutions are CKAN and Dataverse:

- CKAN[91] is an open source data portal and data management solution providing a streamlined way to make data discoverable and presentable. There is a page for each dataset record listing digital resources and a rich collection of metadata. Moreover, besides a rich CKAN core set of features, there are more than 200 community extensions which can fill almost any feature gap. However, there are still some issues and gaps. Authors of a dataset are not linked as entities assigned to data records, for instance, which might make researchers' assessment and reporting quite complex and error-prone. There is a specific internal metadata

---

[89] For example https://www.eprints.org/uk/index.php/flavours/eprints-for-research-data-1
[90] https://www.zora.uzh.ch/
[91] https://ckan.org/

format used by the CKAN platform, but metadata can be exported in the DCAT format.

- Dataverse[92] is an open source research data repository software which supports the FAIR Data Principles and data citation for data sets and files. Moreover, APIs for interoperability and custom integrations have been implemented as part of this software. Furthermore, the schema.org JSON-LD used by Google Dataset Search is supported, as well as DataCite integration. Dataverse supports the Make Data Count project by collecting and displaying usage metrics including counts of dataset views, file downloads, and dataset citations.

Besides adoption of the previous two solutions, a data management system can be built for the specific purpose of institutional or national data platforms such as the Fairdata platform[93]. The Fairdata platform enables verifiable and reproducible science and secure preservation of Finish researchers' digital research outputs. This platform enables publishing of research data in a FAIR repository. It is going to be integrated with Finnish national research hub, meaning links with other research information (researchers, organisations, equipements) will be established. The platform could be used for building metrics for evaluation of Finish researchers and research data published by them.

There are also tools helping to make data from data repositories FAIR and for reusing data. For instance, the FAIR Data Point[94] (FDP) has been developed to achieve data publication in a FAIR manner and foster their findability, accessibility, interoperability and reusability. The FDP is a software layer on top of datasets to expose them as FAIR and inter-linkable data. Moreover, it provides access to the data and metadata using REST-APIs in accordance with the W3C Linked Data Platform specification. The FDP contributes to the development of FAIR Data infrastructures across Europe which is aligned with the EOSC mission and could enable giving recognition to the research data authors through increased number of downloads and usage.

---

92 https://dataverse.org/)
93 https://www.fairdata.fi/en/
94 https://eudat.eu/communities/an-eudat-based-fair-data-approach-for-data-interoperability

The GO FAIR Initiative[95] coordinates and contributes to the coherent development of the Internet of FAIR Data & Services (IFDS)[96]. The scalable and transparent routing of data, tools, and compute (to run the tools on) is a key central feature of the envisioned IFDS. The development approach is largely based on the EOSC communication and the recommendations of an expert group on the EOSC.

The rOpenSci company[97] was founded to make scientific data reproducible. rOpenSci has developed an ecosystem of R open source tools for data science and sharing of research data, and run a community for discussion about issues related to research data eInfrastructures.

Platforms for publishing research objects

Summary:

- *Some existing web platforms can help in giving credits to all data contributors, besides data creators, including reviewers.*
- *The process of publishing could be improved and sped up by using web platforms for publishing which can enable collaboration of any kind and number of contributors in the process of creation of research and knowledge including open and transparent review process.*

There are some open research publishing platforms which contribute to open science and speed up the process of publishing research objects. An example of such a platform is F1000Research[98]. F1000Research is an open research publishing platform for scholars offering rapid publication of articles and other research outputs. Articles are published first and transparent peer reviews with visible reviewer's names and comments are organised after an article has been published. Moreover, the data behind each article are also published and are downloadable.

The other, perhaps less popular, open research publishing platforms are The Winnower and Qeios:

- Winnower[99] is trying to revolutionise science by breaking down the barriers to scientific communication through cost-effective and

[95] https://www.go-fair.org/go-fair-initiative/
[96] https://www.go-fair.org/resources/internet-fair-data-services/
[97] https://ropensci.org/
[98] https://f1000research.com/
[99] https://thewinnower.com/

transparent publishing for scholars and the platform is free of charge for usage. Winnower platform does not support publishing other research objects than textual digital publications.

- Qeios[100] is not completely free of charge for usage, although part of its functionalities are. Moreover, this platform provides additional tools related to writing research outputs, including easy-to-use online text editor enabling seamless collaboration with colleagues on the same document. The Qeios platform does not support publishing other research objects than textual digital publications.

Qeios is an example of a web tool which can be used for collaborative creation of research objects. Besides usage of desktop or web tools for creation of research objects, some research objects (datasets) can be the result of crowdsourcing, e.g. citizen science. Researchers can analyse their information more quickly and accurately with the help of volunteers and citizen science platforms such as Zooniverse[101], SciStarter[102], and EU-Citizen.science[103].

arXiv[104] is a popular open-access repository of digital preprints[105]. arXiv is a free distribution service and an open-access archive for more than million and seven hundreds thousands scholarly articles which set arXiv as one of the top ten global hosts of green open access and a significant promoter of open science. Moreover, there are generic global repositories of open access academic objects which also fosters further development of open science such as Figshare[106]and Zenodo[107].

The aim of the traditional anonymous peer review process used in academic publishing is to ensure quality and objectivity. There are at least two drawbacks of anonymous peer review: reviewers give up their time for very little reward, and a lack of transparency which can lead to flawed and fraudulent research making it into prestigious publications. Those issues can be resolved by application of open peer review (Görögh, 2019) or platforms for recording peer

100 https://www.qeios.com
101 https://www.zooniverse.org/
102 https://scistarter.org/
103 https://eu-citizen.science/)
104 https://arxiv.org/
105 https://en.wikipedia.org/wiki/List_of_preprint_repositories
106 https://figshare.com
107 https://zenodo.org

reviews such as Publons[108]. The Publons platform enables giving credit to peer reviewers. Publons Peer Review Awards are recognitions for top peer reviewers and editors.

Research objects and open science - Good practices and repository certification

Summary:

- *Adoption of recommendations and good practices for the implementation of a data repository can lead to discoverable data in understandable format by humans and machines, recorded usage of data and recognized efforts of data creators.*
- *The list of criteria for estimation of quality of research data repositories prescribed in the form of certificate can be used as a benchmark for comparison and helps to determine the strengths and weaknesses of a repository, as well as for research data assessment based on quality of publication channel.*

Adoption of recommendations and good practices for implementation of data repository can lead to discoverable data in understandable format by humans and machines, recorded usage of data and recognised efforts of data creators.

The Data on the Web Best Practices[109] prescribed by a W3C working group defined recommendations for the development of data publication infrastructures and usage of formats and vocabularies in accordance with open science and theFAIR principles. Key features which should be reached by implementing those best practices are discoverable data in an understandable format by humans and machines, recorded usage of data and recognizing efforts of data creators.

The list of criteria for estimation of the quality of research data repositories prescribed in the form of a certificate can be used as a benchmark for comparison and help to determine the strengths and weaknesses of a repository, as well as for research data assessment based on quality of publication channel.

Different organisations can maintain different and not aligned lists of recommended repositories for data sharing. DataCite and FAIRsharing are

[108] https://publons.com/
[109] https://www.w3.org/TR/dwbp

carrying out a joint project to identify a set of criteria for the recommendation of research data repositories for the benefit of the broader research community[110]. This set of criteria can be used for creating a common list of recommended data repositories. In future, an evaluation rulebook based on a common list could be used to qualify research data for research assessments.

The Core Trust Seal[111] offers a core level certification for data repositories based on the DSA–WDS Core Trustworthy Data Repositories Requirements catalogue and procedures. This universal catalogue of 16 requirements reflects the core characteristics of trustworthy data repositories. Moreover, it can be used as a benchmark for comparison and helps to determine the strengths and weaknesses of a repository. Some evaluation rulebooks in the future could count research datasets deposited in data repositories with Core Trust Seal certificates. These certificates could provide proof that a repository is sustainable and trustworthy; repositories could also fulfil some other set of criteria such as one defined in Data Repository Selection: Criteria That Matter (Sansone et al., 2020).

The ethical and legal situation for publishing research data is not always clarified and people hesitate to invest time in addressing these complex issues (Stehouwer and Wittenburg, 2018). Therefore, members of the Harvard Privacy Tools project have developed DataTags[112], a suite of tools to help researchers share and use sensitive research data in a standardized and responsible way. DataTags can help and encourage researchers who are not legal or technical experts to share their research data under appropriate tags (and licenses). The DataTags are human-readable and machine-actionable labels that express conditions under which research data can be stored, transmitted, or used.

3.5.3. AGGREGATION OF SYSTEMS AND LINKING OF RESEARCH ENTITIES

Summary:

- *Research entities isolated in numerous systems usually can't be used for FAIR and comprehensive assessment of a researcher; thus integration of those systems and linking their research entities instances are necessary for the*

[110] https://blog.datacite.org/mou-between-datacite-and-fairsharing/
[111] https://www.coretrustseal.org
[112] http://datatags.org

purpose of assessment. This integration is necessary also for increasing discoverability of research objects.

Integration of research eInfrastructures

Summary:

- *Defined mappings between widely used formats for representation of research information and for description of research objects make possible integration of research eInfrastructures.*

CERIF is one of the three widely used standards for ensuring the interchangeability of research information according to the final report of EUNIS–EUROCRIS joint survey on CRIS and IRa (Ribeiro et al., 2015). Moreover, metadata mappings from the other models to CERIF have been implemented to foster interoperability:

- Lezcano and colleagues (2012) defined mapping between VIVO-ISF and CERIF to enable clients to integrate data coming from heterogeneous sources. They concluded that the majority of mapping problems have risen from higher classification granularity in VIVO-ISF, the representation of VIVO-ISF sub-hierarchies in CERIF as well as from the representation of CERIF attributes in VIVO-ISF. Moreover, VIVO-ISF does not support multilingual features which are supported in CERIF. Besides all previously listed mapping issues, the most significant research information such as information about persons (researchers), projects, and organisation units can be successfully converted from one representation to the other and vice versa.

- The "Core Data Set Research" for the German science (KDSF)[113] has been mapped to VIVO-ISF (Walther, 2019) and CERIF[114]. Due to different purposes of KDSF, the number of the KDSF data model entities is much smaller compared to CERIF. Although CERIF is a more comprehensive data model and a large part of the basic elements mentioned in the KDSF is available in CERIF, there are some not mapped parts of KDSF in CERIF such as information about structured doctoral programs and spin-offs.

---

[113] https://www.kerndatensatz-forschung.de/

[114] https://www.kerndatensatz-forschung.de/version1/technisches_datenmodell/Mapping.html

- The EPOS metadata model[115] is set to provide the required services on the Integrated Core Services for the European Research Infrastructure on Solid Earth. The model contains entities for Persons, Organizations, Services, Equipment, Data, Publication, Web Services, Facility, Software and Model code. The EPOS e-infrastructure[116] has been built on top of this model. The EPOS infrastructure aims at facilitating and promoting the integrated use of data, data products, services and facilities from internationally distributed research infrastructures for Solid Earth Science in Europe, and thus enhances multidisciplinary and collaborative research and promotes transparent open data. A set of EPOS metadata[117] can be mapped to CERIF as it has been described in (Bailo et al., 2017). The mapping was defined on top of the MERIL project representation of Research Infrastructure (RI) concept in CERIF.

The presented mappings enable exchanging metadata between numerous research information systems based on different models. Interoperability between those systems could help in creation of European central catalog of research information which could improve evaluation of European researchers and promotion of open science and open access researchers results.

Some mappings between previously listed formats for the description of research data have been defined to increase interoperability of research data repositories based on various models. For instance, DataCite to DCAT-AP Mapping has been defined[118]. This mapping enables sharing metadata about research data across DCAT-AP-enabled data catalogues and the DataCite infrastructure. Moreover, mappings of metadata formats for describing research objects to CERIF have been defined as a result of the VRE4EIC projects[119] for the sake of creation of a CERIF based central catalog of research objects (Remy et al., 2019).

---

[115] https://www.epos-ip.org/what-metadata-epos
[116] https://www.epos-ip.org/
[117] https://www.epos-ip.org/what-metadata-epos
[118] https://ec-jrc.github.io/datacite-to-dcat-ap/
[119] https://zenodo.org/record/2548732

Linking of research ecosystem entities

Summary:

- *Linking of research entities through the available eInfrastructures assists in making research reproducible and making it possible to give credits to research objects contributors. Linking of research entities is usually performed based on persistent identifiers and there are initiatives and projects for establishing seamless integration between articles, data, and researchers across the research eInfrastructures based on persistent identifiers.*
- *There are well-known persistent identifiers for certain research entity types, but not for all.*

Scholix[120] stands for "(a Framework for) Scholarly Link Exchange". It is a consensus achieved by journal publishers, data centres, and global service providers to create an open global information ecosystem by collecting and exchanging links between research data and literature. The Scholix initiative offers link information packages containing information about the two objects (data and literature) and information about the nature of the link and the link package itself (date, issuer, rights, etc.). ResearchGraph[121] and OpenAIRE Research Graph[122] are similar initiatives. All three initiatives are based on the idea of PID graphs (Fenner and Aryani, 2019).

Links between literature and data (Colavizza et al., 2020) could be a source for defining a metric for data assessment: how many links, weight of the links - for example based on number of citations for literature, etc. However, the link from data and publications to the authors and other contributors should be also established using the other infrastructures and initiatives (such as ORCID) in order to give credits to researchers.

ORCID[123] is a well-known global unique persistent identifier which is a part of the wider digital infrastructure needed for researchers to share information on a global scale. ORCID enables transparent and trustworthy connections between researchers, their contributions, and affiliations. The EOSC-Pillar 'National Initiatives' Survey (Boldos et al., 2020) shows that on average, 42% of the repositories in the five EOSC-Pillar countries (Austria, Belgium, France, Germany and Italy) use unique identifiers for researchers in metadata, most frequently

[120] http://www.scholix.org/
[121] https://researchgraph.org
[122] https://graph.openaire.eu/
[123] https://orcid.org/

ORCIDs. Records stored in the ORCID platform[124] could also be a source for evaluation of researchers and their achievements. ORCID records can be manually entered or by linking data from other sources. Herzog and Radford (2015) describes using ORCID and the tool ÜberWizard for ORCID to simplify filling ORCID records, linking ORCID records with funding information. Together with efforts by funding organisations to integrate ORCID identifiers into grant applications and post-award reporting workflows, these tools could enable automatic harvesting of funding records into ORCID profiles.

The THOR H2020 project[125] aimed to establish seamless integration between articles, data, and researchers across research eInfrastructures. THOR established interoperability between the existing ORCID and DataCite infrastructures to enable establishing links between contributors, their roles, the organisations to which they are connected and the data they produce. The project is now complete; its successor is the ongoing FREYA H2020 project.

The FREYA project[126] aims to build an infrastructure for persistent identifiers as an important component of open science. EOSC-hub and FREYA signed a Memorandum of Understanding (MoU) in July 2020. The European Open Science Cloud (EOSC) will bring open science to Europe's researchers through the development of open and seamless services for storage, management, analysis and re-use of research data and other research objects, while FREYA will support the EOSC by developing a PID infrastructure that will facilitate and boost this ecosystem. Establishing links between publications, data and researchers using federated PID eInfrastructures is necessary for researchers' evaluation and giving credits/rewards to the authors of research objects - publications and data (Haak et al., 2018).

Crossref is a non-profit organisation with the goal to make research outputs easy to find, cite, link, assess, and reuse[127]. Crossref is an official Digital Object Identifier (DOI) Registration Agency. The DOI is a well known persistent identifier which can be assigned to research data and other research digital objects (software code, publications, etc.). Assigning DOIs to research data is in

---

124 https://members.orcid.org/api/tutorial/reading-xml
125 https://project-thor.eu/
126 https://www.project-freya.eu/en
127 https://www.crossref.org/

accordance with FAIR principles. Moreover, it enables traceability[128] and thus makes it easier to assign metrics to research data and other objects. One approach of linking research data with scientific publications using DOIs has been presented by Novacescu and colleagues (2018). Authors reported on a pilot project carried out in collaboration with the AAS publisher[129] and a data repository. They found that during a 1.5-year period, over 75% of submitting authors opted to use the integrated DOI service to clearly identify data analysed during their research project when prompted at the time of paper submission. COCI[130], OpenCitations Index of Crossref open DOI-to-DOI citations, is an RDF dataset containing details of all the citations that are specified by the open references to DOI-identified works present in Crossref.

Discovery platforms

Summary:

- *There are local, European and global, type specific and type agnostic aggregators of digital repositories. Those aggregators promote open science, improve discoverability of research objects and enable giving credits to authors based on new metrics for assessment of popularity and impact of research objects.*
- *There are also initiatives for building registers for discovering available infrastructures and services.*

OpenAIRE[131] is an aggregator for European digital repositories. The OpenAIRE platform enables search over numerous institutional repositories. Moreover, it also harvests metadata from CRIS systems based on a CERIF profile developed for the OpenAIRE Guideline for CRIS managers[132]. In addition to this guideline, guidelines for Literature Repositories and Data Archives have been also published[133], while guidelines for software repositories and other research product repositories are in preparation.

A research data repository aggregator improves discoverability of research data stored in data providers. re3data[134] is a service managed by DataCite for

---

[128] https://www.crossref.org/services/reference-linking/; https://www.crossref.org/services/cited-by/
[129] https://journals.aas.org/
[130] https://opencitations.net/index/coci
[131] https://www.openaire.eu/
[132] https://guidelines.openaire.eu/en/latest/cris/index.html
[133] https://guidelines.openaire.eu/en/latest/
[134] https://www.re3data.org/

providing detailed information on more than 2,000 research data repositories and search of data from those repositories. Icons are assigned to data providers making identification of key characteristics of repositories easier. Characteristics represented by icons include open access, certification of data repositories, and assigning of DOIs to data. re3data supports the FAIR Principles by making its information openly accessible and machine-actionable using CC0 (no copyright reserved) through its API.

Google Dataset Search[135] supports the discovery of datasets wherever they are hosted, whether it is a publisher's site, a digital library, or an author's personal web page. The only precondition for being indexed and searchable through this platform is definition of dataset types in accordance with schema.org. DOI and License can be attached to a dataset, as well as citation information including authors names and affiliations.

Polleres (2021) presents some issues with making the FAIR data ecosystem, and suggests a vision of a decentralized system for moving from Open Data Portals to Open Data Ecosystems. Moreover, Pollers stresses that knowledge graphs/Linked Data can help to strengthen open data and thinks that closed data should be included in considerations for building a broader data ecosystem.

DataONE is a community driven program providing search and access to earth and environmental data across multiple member repositories. DataONE launched metrics visualizations for datasets through its search and discovery platform[136] in 2018. Metrics in DataOne include live counts of citations, downloads, and views for each dataset in the network could be a source for assessment of a researcher and his/her research data. ORCIDs can be assigned to researchers in DataONE which makes possible the creation of researchers profiles and assigning research data to those profiles.

Wittenburg and colleagues (2020) described integration of three data sources. They discussed the typical integration challenges which have to be overcome to integrate data from different data sources such as fragmentation, bad quality and also social differences. They concluded that the use of PIDs is key in the process of integration.

---

135 https://datasetsearch.research.google.com/
136 https://search.dataone.org

The majority of research data aggregator platforms use protocols for harvesting metadata including link to data in the source repository. However, there are also approaches for packaging research data and associated metadata. RO-Crate (Research Object Crate)[137] specifies a method of organizing file-based data with associated metadata, using linked data principles, in both human and machine-readable formats, with the ability to include additional domain-specific metadata. It is a lightweight approach for packaging research objects with their metadata based on schema.org annotations in JSON-LD. RO-Crate can be used as a protocol for the implementation of research objects aggregators for the purpose of enhancing discoverability of research objects.

ScienceOpen[138] is a discovery platform enhancing open science which aims to give credits to scholars for their research results. The features of the platform are advanced search and discovery functions, automatic recommendation, social sharing, and integration with ORCID.

This is not a publishing platform, it is a discovery platform, meaning it only increases discoverability of already published articles and enables post-publication discussion within the scientific community in the form of post-publication peer review. The database is expanding at a rate of more than one million records per month. As it is already written above only final published versions of articles are harvested, with the exception of papers in the arXiv. Adding ‘green’ versions and more pre-prints in the future are in the plans.

ResearchGate (a social network for scholars) can be also seen as a discovery platform for open research objects (Van Noorden, 2014). Although open science is promoted within ResearchGate, it is not fully under control, meaning some of the uploaded papers and data appear to infringe copyright, because the authors uploaded the publisher's version.

Kudos is a cloud-based platform, through which researchers can accelerate and broaden the positive impact of their research in the world by storytelling. ICT tools such as Kudos[139] can help researchers to make a story by bringing everything about research together into one place and explaining what it’s about and why it’s important - in non-technical language that makes it accessible to a

[137] https://www.researchobject.org/ro-crate/
[138] https://www.scienceopen.com/
[139] https://info.growkudos.com/

broad audience. Moreover, the platform aggregates relevant metrics about research, and maps outreach activities against those metrics.

There are also initiatives for building registers for discovering available infrastructures and services.

The MERIL (Mapping of the European Research Infrastructure Landscape) portal provides access to a database that stores information about openly accessible research infrastructures (RIs) in Europe. The MERIL portal[140] could help to develop open science by giving researchers access to information on research infrastructures in Europe, their facilities, and areas of research. The MERIL project (H2020) ended in 2019; users are currently redirected to the CatRIS project (H2020).

CatRIS[141] is an open H2020 project. The main result is a trusted and user-friendly portal to a harmonised and aggregated catalogue of services and resources provided by research infrastructures and core facilities across Europe. It is a bottom-up initiative that is meant to be populated and run by research infrastructures and core facilities service providers at European, national, regional and institutional levels. CatRIS will be complementary to and interoperable with the EOSC catalogue. The data in the MERIL and CatRIS portals could help to ensure effective planning for future research infrastructures, by pinpointing gaps and identifying opportunities for collaboration at the European level - for instance to find missing infrastructures for the development of open science or research assessment.

EuroRIs-Net+[142] builds on the Network of National Contact Points for the Research Infrastructures programme (RIs NCPs). The EuroRIs-Net+ project (FP7) has been over and the successor is the RICH H2020 project[143]. Building of RIs NCPs facilitates cooperation between NCPs, promotes the effective implementation of the research infrastructure programme, supports transnational and virtual access to RIs and highlights the opportunities offered by research infrastructures. RIs NCPs can gather, organise and provide access to information on RI projects and calls, which will facilitate promotion of projects

---

140 https://portal.meril.eu/meril/
141 https://project.catris.eu/
142 http://www.euroris-net.eu/
143 http://www.rich2020.eu/about

and funding calls, increase competitiveness of the call and the quality of RI projects including the project and calls related to the open science and EOSC.

3.5.4. CVs, METRICS AND INDICATORS

In this section, we analyze tools for making CVs/portfolios, metadata models for describing research objects' metrics and indicators, as well as indicator databases and tools for conducting research impact analyses.

Tools for making CVs/portfolios

- Tools for making CVs/portfolios make easier creation of a CV, introducing a standardized form of CVs which simplifies the process of researchers' assessments.
- Existing tools for making CVs/portfolios don't fully support diversity of research objects' types and scholarly communication channels.

Bhargava and colleagues (2015) presented a reference manager approach for the creation of a CV. Usage of a tool makes it easier to create a CV and eliminates some phases of error-prone manual processes. The approach was demonstrated on usage of the reference manager software called Papers. Researchers can create an academic portfolio by usage of the Papers tool that allows digital organisation of teaching, and research accomplishments in an indexed library enabling efficient updating, rapid retrieval, and easy sharing. However, there is no support for internet-based synchronization across different platforms such as web-based storage, desktop computers, and mobile devices, neither for adding references for research data which are crucial for the adoption of this tool in the open science era.

DeGóis[144] is a Portuguese academic CV platform based on ORCID as the central hub of scientific works (Sousa Pinto et al., 2017). It is being replaced with a new platform CIÊNCIAVITAE[145]. Usage of ORCIDs to identify a researcher solves many of the problems related to the correct identification of the research outputs' authors; this is important for giving rewards/recognition to researchers, as well as for researchers' assessments. However, the limitation of this approach is the fact that the ORCID platform doesn't support evidence of all research objects and other aspects of CVs at the moment.

---

[144] http://www.degois.pt/globalindex.jsp#
[145] https://cienciavitae.pt/

Cataloguing research outputs only once for the purposes of dissemination and creation of researchers' profiles reduces researchers cataloguing efforts and incentivizes researchers to collaborate with research infrastructure developers, administrators, and librarians. Takaku and Tanifuji (2009) presented an integrated approach to a Researcher Portfolio and Institutional Repository in NIMS eSciDoc. A researcher profile page can be generated from the integrated system. Moreover, the integrated data can be a baseline for researchers' assessment. Although the authors only discussed storage of articles, this approach could be extensible with support for research data and other research objects.

Previously presented tools for making CVs/portfolios do not fully support diversity of research objects' types and scholarly communication channels.

Models for metrics and indicators

Summary:

- *Any Snowball Metric representing institution specific evaluation perspectives in accordance with evaluation recipes and good practices can be represented in the well-known format for representation of research information - CERIF.*

As mentioned previously, Snowball Metrics provide the opportunity for institutions to approach evaluation from their own perspective by reusing existing standards - similar to next generation metrics. An example of reusing existing standards is the CERIFication of Snowball Metrics performed in partnership with euroCRIS[146]. Snowball Metrics can be represented in CERIF XML as the "Measurements" entity instances, which allows machines-readable transfer of data (values) in the Snowball Metrics framework. Expressing Snowball Metrics in the CERIF data model enhances the possibility of using a CERIF based system for responsible research assessment. A Snowball Metrics "message" in CERIF XML contains two elements:

- Definition of the submitting institution, as a <cfOrgUnit>,
- A number of Measurements – the individual Snowball Metrics – as <cfMeas>.

[146] https://snowballmetrics.com/eurocris-cerif-xml-for-snowball-metrics/

Tools based on traditional metrics and indicators

Summary:

- *There is a set of software tools for citation analysis which can be used for research assessment based on traditional approaches.*
- *The common issues of tools for citation analysis are comprehensiveness in data, limitations to textual publications citations and not distinguishing disputing from supporting citations.*

There are several well-established and emerging international data sources for publication data, such as WOS, Scopus, Google Scholar, Microsoft Academic, and Dimensions (Martín-Martín et al., 2020; Visser, van Eck, & Waltman, 2020).

The traditional approach for measuring impact of research outputs, researchers and institutions is based on citation analysis. Two popular tools for this approach are InCites[147] and SciVal[148]. The first one is a product of Clarivate Analytics and citation analysis is based on the Web of Science database, while the second one is a product of Elsevier and citation analysis is based on the Scopus database. The limitations for both tools are comprehensiveness of analysed data including the number of records and types of records. Moreover, the tools are not promoting open science and new methods and indicators for measuring open research objects impacts. Analyses are publication based, impact of other research objects are not taken into account.

Besides those two comercial solutions for citation based research assessment and reporting, there are some tools for citation and co-authorship network analysis for the purpose of conducting bibliometric studies, such as VosViewer[149], CitNetExplorer[150], CiteSpace[151]. Those tools can be used to find patterns and trends in a field, highly impactful publications, institutions and researchers. Although those tools enable import of data from numerous sources, combination of data sources, deduplication, resolving identification of different persistent entity identifiers are not automatised, thus it is quite complicated and a lot of manual work is needed to make a comprehensive study using the data from various sources. Moreover, all those tools are publication based, meaning other open research objects are not taken into account.

[147] https://clarivate.com/webofsciencegroup/solutions/incites/
[148] https://www.elsevier.com/solutions/scival
[149] https://www.vosviewer.com/
[150] https://www.citnetexplorer.nl/
[151] http://cluster.cis.drexel.edu/~cchen/citespace/

There are also examples of research analytics platforms developed for the needs of an institution or region such as the DTU (Technical University of Denmark) Research Analytics Platform[152]. The platform presents data and calculations from Web of Science and InCites in a fast and simple way, adapted to DTU needs and preferences. This is a tool for generating collaboration reports between two institutions based on collaborative publications. Establishing relations with other institutions is important for further development of an institution and thus some credit could be given to researchers who established collaboration with other institutions through publishing collaborative papers. However, collaboration could be established also through mobility programs and projects which is not taken into account in this tool.

The common issues of tools for citation analysis are comprehensiveness in data, limitations to textual publications citations and not distinguishing disputing from supporting citations.

All citations are not positive; there are also disputing citations. Those citations should not be treated in the same way in a citation analysis. Scite[153] is a platform for discovering and evaluating scientific articles via Smart Citations. Smart Citations give context of a citation, i.e. allow users to see how a scientific paper has been cited by providing a citation classification describing whether it provides supporting or disputing evidence for the cited claim. Scite classifies citations in three categories: disputing, supporting, and others (without evidence for its validity). However, the deep learning based algorithm is not perfect, and mis-classifications are possible. Moreover, users can see where the citation appeared - Introduction, methodology, results, discussion, etc. There are more than 690 million smart citations in the database. There is an API implemented as part of the Scite platform enabling integration of Scite services into third-party applications such as Zotero[154]. This platform provides classification of citations which could enable building research assessment policies not only based on pure number of citations, i.e. disputing citations could be treated in a proper way. Unfortunately, only citations of textual publications have been analysed. The platform could be extended with support for tracking data and other research objects citations.

---

[152] http://rap.adm.dtu.dk/
[153] https://scite.ai/
[154] https://github.com/scitedotai/scite-zotero-plugin

OpenCitations, DataCite, PLOS, eLIFE, wikimedia foundation, and Center for Culture and Technology run the Initiative for Open Citations (I4OC)[155] to promote the unrestricted availability of scholarly citation data. Citation data are usually not machine-readable, and the aim of this initiative is to change practices of cataloguing data on citations to be structured, separable, and open. The open citation data will enable building new services over those data such as creation of a public citation graph to explore connections between knowledge fields, to follow the evolution of ideas and scholarly disciplines, and to perform researchers' assessment based on citation graphs. Usually citations in publications are linked to other research publications, but it is expected that citation of research data, software and other research objects will become a common practice in the near future.

Tools based on new metrics and indicators

Summary:

- *There are tools which can provide new metrics and indicators, usually called altmetrics, and which can be easily integrated in a researcher web page, as well as journal, publisher, institution web page.*
- *Some tools enable building the new indicators based on linked research entities.*

Altmetrics include metrics for usage, mentions, captures and social media popularity. There are tools which can provide those indicators and which can be easily integrated in a researcher web page, as well as journal, publisher, institution web page. Two popular tools of this type are Altmetric and PlumX.

Altmetric[156] is a part of Digital Science company and develops a set of tools for tracking where published research is mentioned online. The provided tools and services can be used by institutions, publishers, researchers, funders and other organisations to monitor their impact based on altmetrics. The easy-to-embed Altmetric widget called Altmetric Badge provides an instantly recognisable visualisation to help showcase the wider influence and dissemination of a published research object. Besides that easily integrable widget, there is an API for fetching all metrics values assigned to a research object.

[155] https://i4oc.org/
[156] https://www.altmetric.com/

Plum Analytics[157] is a part of Elsevier company dedicated to measuring the influence of scientific research objects. Easily embeddable PlumX widgets make it easy to display PlumX metrics on a website of researcher, organisation, or publisher. Similar as for Altmetric, besides easily integrable widget, metrics are also fetchable through an API.

Dimensions[158] is a linked research data platform making it easy to navigate through data using links between grants, publications, clinical trials, patents, datasets and policy documents. Dimensions maps the entire research lifecycle, meaning someone can follow research from funding through research output to impact. Users are encouraged to use the broad range of connected data freely available in the Dimensions database to develop the next generation of useful indicators for research objects assessment, such as Co-Citation Percentile Rank (Seppänen et al., 2020). Moreover, those connected data could be used to promote open research objects and link them to awards/recognitions[159].

The European Research Infrastructure for Science, Technology and Innovation policy Studies (RISIS)[160] is a H2020 project aiming at building data and services infrastructure supporting the development of a new generation of analyses and indicators. The RISIS-KNOWMAK tool, which is a result of this project, enables the analysis and download of a number of relevant integrated indicators on knowledge production in Europe. These indicators and analysis of these indicators could have an impact on strategy and vision for development of open science in Europe. Moreover, the tool contributes to Open evaluation.

ResearchGate[161] is a social network for researchers, promoting cooperation between more than 17 millions registered researchers, as well as promoting open debates and open science (Van Noorden, 2014). Moreover, it defines the RG score metric which can be used as an altmetric, although it is not a reliable researcher impact measure. RG score is an author-level metric which has been criticised as having questionable reliability and an unknown calculation methodology.

[157] https://plumanalytics.com/
[158] https://www.dimensions.ai/
[159] For example https://www.dimensions.ai/blog/new-platform-for-open-science/
[160] https://www.risis2.eu/
[161] https://www.researchgate.net/

4. REFERENCES

ACUMEN Consortium (2014) Guidelines for Good Evaluation Practice with the ACUMEN Portfolio, http://research-acumen.eu/wp-content/uploads/D6.14-Good-Evaluation-Practices.pdf

Adam P, Ovseiko PV, Grant J, et al. (2018) ISRIA statement: ten-point guidelines for an effective process of research impact assessment. Health Res Policy Sys 16, 8. https://doi.org/10.1186/s12961-018-0281-5

Agate N, Kennison R, Konkiel S. et al. (2020)The transformative power of values-enacted scholarship. Humanit Soc Sci Commun 7, 165. https://doi.org/10.1057/s41599-020-00647-z

Algra A, Koopman I, Snoek R (2020) How young researchers can re-shape the evaluation of their work. Looking beyond bibliometrics to evaluate success. Nature index 31 March 2020

Ali-Khan SE, Harris LW, Gold ER (2017) Point of view: Motivating participation in open science by examining researcher incentives. eLife 2017;6:e29319 doi: 10.7554/eLife.29319

Allen C, Mehler DMA (2019) Open science challenges, benefits and tips in early career and beyond. PLoS Biol 17(5): e3000246. https://doi.org/10.1371/journal.pbio.3000246

Allen R, Hartland D (2018) FAIR in practice - Jisc report on the Findable Accessible Interoperable and Reusable Data Principles (Version 1). Zenodo. http://doi.org/10.5281/zenodo.1245568

Aroeira R, Castanho M (2020) Can citation metrics predict the true impact of scientific papers?. The FEBS Journal. 287. 10.1111/febs.15255.

Ayris P, López de San Román A, Maes K, Labastida I (2018) Open Science and its role in universities: A roadmap for cultural change. Leuven: LERU Office. https://www.leru.org/files/LERU-AP24-Open-Science-full-paper.pdf

Bailo D, Ulbricht D, Nayembil ML, Trani L, Spinuso A, Jeffery KG (2017) Mapping solid earth Data and Research Infrastructures to CERIF. *Procedia Computer Science*, *106*, pp.112-121.

Beigel F, et al. (2020) TOWARDS A TRANSFORMATION OF SCIENTIFIC RESEARCH ASSESSMENT IN LATIN AMERICA AND THE CARIBBEAN Diagnosis and Proposals for a Regional Initiative - Report of Latin American Forum for Research Assessment. https://www.clacso.org/en/diagnostico-y-propuestas-para-una-iniciativa-regional/

Benedictus R, Miedema F (2016) Fewer numbers, better science. Nature 538, 453–455 (27 October 2016) doi:10.1038/538453a

Bhargava P, Patel VB, Iyer RS, et al. (2015) Academic Portfolio in the Digital Era: Organizing and Maintaining a Portfolio Using Reference Managers. J Digit Imaging 28, 10–17. https://doi.org/10.1007/s10278-014-9713-0

Bierer BE, Crosas M, Pierce HH (2017) Data Authorship as an Incentive to Data Sharing. N Engl J Med 2017; 376:1684-1687 DOI: 10.1056/NEJMsb1616595

Bishop BW, Hank C (2018) Measuring FAIR principles to inform fitness for use. International Journal of Digital Curation 13(19. DOI: https://doi.org/10.2218/ijdc.v13i1.630

Bloemers M, Montesanti A (2020) The FAIR funding model: Providing a framework for research funders to drive the transition toward FAIR data management and stewardship practices. Data Intelligence 2(2020), 171–180. doi: 10.1162/dint_a_00039

Bodlos A, Hönegger L, Kaczmirek L, Beckmann V, Breton V, Romier G, van Wezel J, Streit A, Stevanovic U, Galeazzi F, Tanlongo F, van Nieuwerburgh I (2020) Summary report of the EOSC-Pillar \National Initiatives Survey. Version 1.1, Zenodo. doi: 10.5281/zenodo.393731

Borgman CL (2016) Data citation as a bibliometric oxymoron. In C. R. Sugimoto (Ed.), Theories of Informetrics and Scholarly Communication (pp. 93–115). Berlin, Boston: Walter de Gruyter GmbH & Co KG.

Bryant R, Clements A, de Castro P, Cantrell J, Dortmund A, Fransen J, Gallagher P, Mennielli M (2018) Practices and Patterns in Research Information Management: Findings from a Global Survey. Dublin, OH: OCLC Research. https://doi.org/10.25333/BGFG-D241.

Burgelman J-C, Pascu C, Szkuta K, Von Schomberg R, Karalopoulos A, Repanas K, Schouppe M (2019) Open Science, Open Data, and Open Scholarship: European

Policies to Make Science Fit for the Twenty-First Century. Frontiers in Big Data 2: 43. DOI=10.3389/fdata.2019.00043

Casola L (2020) Planning for Long-Term Use of Biomedical Data: Proceedings of a Workshop.

Cohen J, Katz DS, Barker M, Chue Hong N, Haines R, Jay C (2021) The Four Pillars of Research Software Engineering. IEEE Software, vol. 38, no. 1, pp. 97-105, Jan.-Feb. 2021, doi: 10.1109/MS.2020.2973362.

Colavizza G, Hrynaszkiewicz I, Staden I, Whitaker K, McGillivray B (2020) The citation advantage of linking publications to research data. PLoS ONE 15(4): e0230416. https://doi.org/10.1371/journal.pone.0230416

Collins S, et al. (2018) Turning Fair into Reality. Final Report and Action Plan from the European Commission Expert Group on FAIR Data. https://ec.europa.eu/info/sites/info/files/turning_fair_into_reality_1.pdf

Corti L, Van den Eynden V, Bishop L, Woollard M (2019) *Managing and sharing research data: a guide to good practice*. SAGE Publications Limited.

Curry S, de Rijcke S, Hatch A, Dorsamy P, van der Weijden I, Wilsdon J (2020) The changing role of funders in responsible research assessment: progress, obstacles and the way ahead. Research on Research Institute. Report. https://doi.org/10.6084/m9.figshare.13227914.v1

Curtiss J, Gramatté AM (2018) Publishing models, assessment, and open science - Report and outcomes from a workshop held by the Global Young Academy. https://globalyoungacademy.net/activities/optimising-assessment-promoting-excellence/

Data Citation Synthesis Group (2014) Joint Declaration of Data Citation Principles. https://doi.org/10.25490/a97f-egyk

Davidson J, Engelhardt C, Proudman V, Stoy L, Whyte A (2019) D3.1 FAIR Policy Landscape Analysis (Version v1.0_draft). FAIRsFAIR.

Davidson J, Grootveld M, Whyte A, Herterich P, Engelhardt C, Stoy L, Proudman V (2020) D3.3 Policy Enhancement Recommendations (Version 1.0). Zenodo.

del Carmen Calatrava Moreno M, Warta K, Arnold E, Tiefenthaler B, Kolarz P, Skok S (2019) Science Europe Study on Research Assessment Practices - Final

Report. https://www.scienceeurope.org/our-resources/science-europe-study-on-research-assessment-practices/

Devaraju A, Huber R, Mokrane M, Cepinskas L, Davidson J, Herterich P, et al. (2020) FAIRsFAIR Data Object Assessment Metrics (Version 0.3). Zenodo. http://doi.org/10.5281/zenodo.3934401

Duarte K, Weber R, Pachero RCS (2016) Purpose-oriented metrics to assess researcher quality. International Conference on Science and Technology Indicators. Valencia (Spain) September 14-16 2016.

Dutch Ministry of Education, Culture and Science (2016, April 4). Amsterdam call for action on open science. https://www.government.nl/documents/reports/2016/04/04/amsterdam-call-for-action-on-open-science

Eisfeld-Reschke J, Herb U, Wenzlaff K (2014) Research Funding in Open Science. In: Bartling S., Friesike S. (eds) Opening Science. Springer, Cham. https://doi.org/10.1007/978-3-319-00026-8_16

European Commission (2010) Assessing Europe's University-Based Research. https://web.archive.org/web/20190317171523/http://ec.europa.eu/research/science-society/document_library/pdf_06/assessing-europe-university-based-research_en.pdf

European Commission (2018a) Mutual Learning Exercise (MLE) on Open Science: Altmetrics and Rewards. https://rio.jrc.ec.europa.eu/sites/default/files/report/MLE%20OS_Final%20Report_0.pdf

European Commission (2018b). Open Science Policy Platform Recommendations.https://www.zbw-mediatalk.eu/wp-content/plugins/pdfjs-viewer-shortcode/pdfjs/web/viewer.php?file=/wp-content/uploads/2018/07/integrated_advice_opspp_recommendations.pdf&dButton=true&pButton=true&oButton=false&sButton=true&v=1.5.4

European Commission (2019, January 16). Cost-benefit analysis for FAIR research data: Policy recommendations. [Website]. http://op.europa.eu/en/publication-detail/-/publication/d3766478-1a09-11e9-8d04-01aa75ed71a1

European University Association (2018) EUA Roadmap on Research Assessment in the Transition to Open Science. https://eua.eu/component/attachments/attachments.html?id=348

Fane B, Ayris P, Hahnel M, Hrynaszkiewicz I, Baynes G, Farrell E (2019) The state of open data report 2019 (p. 1320193 Bytes). Digital Science. https://doi.org/10.6084/M9.FIGSHARE.9980783.V1

Fecher B, Friesike S (2014) Open science: one term, five schools of thought. Opening science, 17-47. https://library.oapen.org/bitstream/handle/20.500.12657/28008/2014_Book_OpeningScience.pdf?sequence=1#page=24

Fecher B, Friesike S, Hebing M et al. (2017) A reputation economy: how individual reward considerations trump systemic arguments for open access to data. Palgrave Commun 3, 17051. https://doi.org/10.1057/palcomms.2017.51

Federer L (2019) Who, what, when, where and why? Quantifying and understanding biomedical data reuse. PhD thesis. University of Maryland.

Fenner M (2014) Altmetrics and Other Novel Measures for Scientific Impact. In: Bartling S., Friesike S. (eds) Opening Science. Springer, Cham. https://doi.org/10.1007/978-3-319-00026-8_12

Fenner M, Aryani A (2019) Introducing the PID Graph. https://doi.org/10.5438/JWVF-8A66

Ficarra V, Fosci M, Chiarelli A, Kramer B, Proudman V (2020) Scoping the Open Science Infrastructure Landscape in Europe. Zenodo. http://doi.org/10.5281/zenodo.4159838

Friesike S, Schildhauer T (2015) Open science: many good resolutions, very few incentives, yet. In *Incentives and Performance* (pp. 277-289). Springer, Cham.

Forsström PL, Lilja E, Ala-Mantila M (2020) Atlas of open science and research in Finland 2019: Evaluation of openness in the activities of higher education institutions, research institutes, research-funding organisations, Finnish academic and cultural institutes abroad and learned societies and academies. Final report. https://julkaisut.valtioneuvosto.fi/handle/10024/161990

Fosci M, Richens E, Johnson R (2019) Insights into European research funder Open policies and practices. SPARC Europe Report. http://doi.org/10.5281/zenodo.3401278

Foster D, Locker A, Maldonado A, O'Dea E, Saez Llorente J, Sõgel J, Tapanainen H, Thomas S, Tirian A, Richardson J, Marsic I, Pintado C, Czomba B, Ringwald F, Gabbi S (2019) Publication of scientific data from EU-coordinated monitoring programmes and surveys. EFSA supporting publication 2019: 16(1):EN-1544. 43 pp. doi: 10.2903/sp.efsa.2019.EN-1544

Gadd E (2019) Influencing the Changing World of Research Evaluation. Insights 32 (1): 6. DOI: http://doi.org/10.1629/uksg.456

Global Research Council (2018) STATEMENT OF PRINCIPLES ON PEER/MERIT REVIEW. https://www.globalresearchcouncil.org/fileadmin/documents/GRC_Publications/Statement_of_Principles_on_Peer-Merit_Review_2018.pdf

Goble C, Cohen-Boulakia S, Soiland-Reyes S, Garijo D, Gil Y, Crusoe MR, Peters K, Schober D (2020) FAIR Computational Workflows. Data Intelligence (2020) 2 (1-2): 108–121. https://doi.org/10.1162/dint_a_00033

Görögh E (2019) Open peer review. Conference on Open Research Data in Slovenia, Workshop on Open Peer Review, November 15, 2019, Maribor

Groth P, Cousijn H, Clark T, Goble C (2020) FAIR Data Reuse – the Path through Data Citation. Data Intelligence 2020; 2 (1-2): 78–86. doi: https://doi.org/10.1162/dint_a_00030

Guédon JC, et al. (2019) Future of scholarly publishing and scholarly communication: Report of the Expert Group to the European Commission. https://data.europa.eu/doi/10.2777/836532

Guy M, Ploeger L (2015) PASTEUR4OA Briefing Paper: Open Access to Research Data. http://pasteur4oa.eu/resources/153#.X2iGAFUzZaR

Haak LL, Meadows A, Brown J (2018) Using ORCID, DOI, and Other Open Identifiers in Research Evaluation. Front. Res. Metr. Anal. 3:28. doi: 10.3389/frma.2018.00028

Hatch A, Curry S (2020) Research Culture: Changing how we evaluate research is difficult, but not impossible. eLife 2020;9:e58654 doi: 10.7554/eLife.58654

Haustein S (2016) Grand challenges in altmetrics: heterogeneity, data quality and dependencies. Scientometrics 108, 413–423. https://doi.org/10.1007/s11192-016-1910-9

Helmer S, Blumenthal DB, Paschen K (2020) What is meaningful research and how should we measure it?. Scientometrics 125, 153–169. https://doi.org/10.1007/s11192-020-03649-5

Herb U (2016) Impact Messung, Transparenz & Open Science. Young Information Scientist, 2016, vol. 1. [Journal article (Unpaginated)]

Herterich P, Davidson J, Whyte A, Molloy L, Matthews B, Kayumbi Kabeya G (2019) D6.1 Overview of needs for competence centres. FAIRsFAIR.

Herzog C, Radford G (2015) ORCID for funders: Who's who - and what are they doing? - ORCID IDs as identifiers for researchers and flexible article based classifications to understand the collective researcher portfolio [version 1; peer review: 1 not approved]. F1000Research 2015, 4:122 (https://doi.org/10.12688/f1000research.6504.1)

Hicks D, Wouters P, Waltman L, de Rijcke S, Rafols I (2015) Bibliometrics: The Leiden Manifesto for research metrics. Nature 520: 429–431 doi:10.1038/520429a

Himanen L, Gadd E (2019) Introducing SCOPE – a process for evaluating responsibly. The Bibliomagician, December 11 2019.

Ilva J (2017) Towards Reliable Data – Counting the Finnish Open Access Publications, Procedia Computer Science, Volume 106, Pages 299-304, https://doi.org/10.1016/j.procs.2017.03.029.

Jacobsen A, de Miranda Azevedo R, Juty N, Batista D, Coles S, Cornet R et al. (2020) FAIR Principles: Interpretations and Implementation Considerations. Data Intelligence (2020) 2 (1-2): 10–29. https://doi.org/10.1162/dint_r_00024

Jamali HR, Nicholas D, Herman E (2016) Scholarly reputation in the digital age and the role of emerging platforms and mechanisms, Research Evaluation, Volume 25, Issue 1, January 2016, Pages 37–49, https://doi.org/10.1093/reseval/rvv032

Jappe A (2020) Professional standards in bibliometric research evaluation? A meta-evaluation of European assessment practice 2005–2019. PLoS ONE 15(4): e0231735. https://doi.org/10.1371/journal.pone.0231735

Kidwell MC, Lazarević LB, Baranski E, Hardwicke TE, Piechowski S, Falkenberg L-S, et al. (2016) Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency. PLoS Biol 14(5): e1002456. https://doi.org/10.1371/journal.pbio.1002456

Koers H, Bangert D, Hermans E, van Horik R, de Jong M, Mokrane M (2020) Recommendations for Services in a FAIR Data Ecosystem. Patterns 1(5) DOI:https://doi.org/10.1016/j.patter.2020.100058

Konkiel S (2020) Assessing the Impact and Quality of Research Data Using Altmetrics and Other Indicators. Scholarly Assessment Reports, 2(1), p.13. DOI: http://doi.org/10.29024/sar.13

Kraus J (2014) Cash, carrots, and sticks: Open Access incentives for researchers. *F1000Research*, *3*(263), p.263.

Kucharski AJ, Funk S, Eggo RM (2020) The COVID-19 response illustrates that traditional academic reward structures and metrics do not reflect crucial contributions to modern science. PLoS Biol 18(10): e3000913. https://doi.org/10.1371/journal.pbio.3000913

Laine H (2018) Open science and codes of conduct on research integrity. Informaatiotutkimus, 37(4). https://doi.org/10.23978/inf.77414

Lawrence R, Mendez E, et al. (2020) Progress on open science: Towards a shared research knowledge system : final report of the open science policy platform. https://data.europa.eu/doi/10.2777/00139

Leonelli S, Spichtinger D, Prainsack B (2015) Sticks and carrots: encouraging open science at its source. Geography and Environment 2: 12-16.

Lezcano L, Jörg B, Sicilia MA (2012) Modeling the Context of Scientific Information: Mapping VIVO and CERIF. In: Bajec M., Eder J. (eds) Advanced Information Systems Engineering Workshops. CAiSE 2012. Lecture Notes in

Business Information Processing, vol 112. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-31069-0_11

Lin D, Crabtree J, Dillo I, et al. (2020) The TRUST Principles for digital repositories. Sci Data 7, 144. https://doi.org/10.1038/s41597-020-0486-7

Lowenberg D, Chodacki J, Fenner M, Kemp J, Jones MB (2019) Open Data Metrics: Lighting the Fire (Version 1). Zenodo. http://doi.org/10.5281/zenodo.3525349

L'Hours H, von Stein I (2019) FAIR Ecosystem Components: Vision (Version 01.00). Zenodo. http://doi.org/10.5281/zenodo.3565428

Mahieu B, Arnold E, Kolarz P (2014) Measuring Scientific Performance for Improved Policy Making. Brussels: European Parliamentary Research Service. https://web.archive.org/web/20190527064957/http://www.europarl.europa.eu/RegData/etudes/etudes/join/2014/527383/IPOL-JOIN_ET(2014)527383(SUM01)_EN.pdf

Martín-Martín A, Thelwall M, Orduna-Malea E, Delgado López-Cózar E (2020) Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: a multidisciplinary comparison of coverage via citations. Scientometrics 126, 871–906. https://doi.org/10.1007/s11192-020-03690-4

Materska K (2019) Research Information Management in the Context of Open Science. ZIN 2019, 57(1A), 37–45

McKiernan E, Bourne PE, Brown CT, Buck S, Kenall A, Lin J, McDougall D, Nosek BA, Ram K, Soderberg CK, Spies JR, Thaney K, Updegrove A, Woo KH, Yarkoni T (2016) Point of view: How open science helps researchers succeed eLife 2016;5:e16800 doi: 10.7554/eLife.16800

McKiernan EC, Schimanski LA, Muñoz Nieves C, Matthias L, Niles MT, Alperin JP (2019) Meta-Research: Use of the Journal Impact Factor in academic review, promotion, and tenure evaluations. eLife 2019;8:e47338 doi: 10.7554/eLife.47338

McNut M (2014) The measure of research merit. Science 05 Dec 2014: Vol. 346, Issue 6214, pp. 1155 DOI: 10.1126/science.aaa3796

Mirowski P (2018) The future(s) of open science. Social Studies of Science. 48(2): 171-203. doi:10.1177/0306312718772085

Moher D, Naudet F, Cristea IA, Miedema F, Ioannidis JPA, Goodman SN (2018) Assessing scientists for hiring, promotion, and tenure. PLoS Biol 16(3): e2004089. https://doi.org/10.1371/journal.pbio.2004089

Moher D, Bouter L, Kleinert S, Glasziou P, Sham MH, Barbour V, et al. (2020) The Hong Kong Principles for assessing researchers: Fostering research integrity. PLoS Biol 18(7): e3000737. https://doi.org/10.1371/journal.pbio.3000737

Molloy L, Nordling J, Grootveld M, van Horik R, Whyte A, Davidson J, et al. (2020) D3.4 Recommendations on practice to support FAIR data principles (Version 1.1). Zenodo.

Mons B, et al. (2017) Cloudy, Increasingly FAIR; Revisiting the FAIR Data Guiding Principles for the European Open Science Cloud. 1 Jan. 2017 : 49 – 56.

Morais R, Borrell-Damián L (2018) 2017-2018 EUA Open Access Survey Results. https://eua.eu/downloads/publications/2017-2018%20open%20access%20survey%20results.pdf

Munafò M, Nosek B, Bishop D, et al. (2017) A manifesto for reproducible science. Nat Hum Behav 1, 0021. https://doi.org/10.1038/s41562-016-0021

Nisonger TE (2004) The benefits and drawbacks of impact factor for journal collection management in libraries. *The Serials Librarian*, *47*(1-2), 57-75. https://doi.org/10.1300/J123v47n01_04

Nosek BA, Alter G, Banks GC, Borsboom D, Bowman SD, et al. (2015) Promoting an open research culture. Science 26 Jun 2015: Vol. 348, Issue 6242, pp. 1422-1425 DOI: 10.1126/science.aab2374

Nosek B (2018) Changing a research culture. https://assets.pubpub.org/5nv701md/01521405455055.pdf

Novacescu J, Peek JE, Weissman S, Fleming SW, Levay K, Fraser E (2018) A model for data citation in astronomical research using digital object identifiers (DOIs). *The Astrophysical Journal Supplement Series*, *236*(1), p.20.

O'Carrol C, et al. (2017, November 14). Evaluation of research careers fully acknowledging Open Science practices: Rewards, incentives and/or recognition for researchers practicing Open Science. [Website].

http://op.europa.eu/en/publication-detail/-/publication/47a3a330-c9cb-11e7-8e69-01aa75ed71a1

O'Carroll C, Kamerlin S, Kohl U, Brennan N, Hyllseth B, O'Neill G, Berg R (2017) Providing researchers with the skills and competencies they need to practise Open Science. Report of the Working Group on Education and Skills under Open Science. 10.2777/121253.

Pampel H, Dallmeier-Tiessen S (2014) Open Research Data: From Vision to Practice. In: Bartling S., Friesike S. (eds) Opening Science. Springer, Cham. https://doi.org/10.1007/978-3-319-00026-8_14

Pergl R, Hooft R, Suchánek M, Knaisl V, Slifka J (2019) "Data Stewardship Wizard": A Tool Bringing Together Researchers, Data Stewards, and Data Experts around Data Management Planning. Data Science Journal, 18(1), p.59. DOI: http://doi.org/10.5334/dsj-2019-059

Perrier L, Blondal E, MacDonald H (2020) The views, perspectives, and experiences of academic researchers with data sharing and reuse: A meta-synthesis. PLoS ONE 15(2): e0229182. https://doi.org/10.1371/journal.pone.0229182

Pierce HH, Dev A, Statham E, Bierer BE (2019) Credit data generators for data reuse. Nature 570, 30-32. doi: https://doi.org/10.1038/d41586-019-01715-4

Piwowar H, Priem J (2013) The power of altmetrics on a CV. *Bulletin of the American Society for Information Science and Technology*, *39*(4), pp.10-13.

Polleres A (2021) From Open Data Portals to Open Data-Ecosystems Experiences & thoughts on how to proceed from monitoring and analyzing Open Data Portals. Presentation at the Future of Open Data Portals webinar organized by European Data Portal. https://aic.ai.wu.ac.at/~polleres/presentations/20210112OpenDataEcosystems_webinar_future-open-data-portals.pdf

Puuska H-M, Guns R, Pölönen J, Sivertsen G, Mañana-Rodríguez J, Engels T (2018) Proof of concept of a European database for social sciences and humanities publications: Description of the VIRTA-ENRESSH pilot. CSC & ENRESSH. https://doi.org/10.6084/m9.figshare.5993506

Puuska H-M, Nikkanen J, Engels T, Guns R, Ivanović D, Pölönen J (2020) Integration of national publication databases — towards a high-quality and comprehensive information base on scholarly publications in Europe. Proceedings of the International Conference on ICT enhanced Social Sciences and Humanities 2020. https://doi.org/10.1051/itmconf/20203302001

Pölönen J, Laakso M, Guns R, Kulczycki E, Sivertsen G (2020) Open access at the national level: A comprehensive analysis of publications by Finnish researchers. Quantitative Science Studies. https://doi.org/10.1162/qss_a_00084

Ravenscroft J, Liakata M, Clare A, Duma D (2017) Measuring scientific impact beyond academia: An assessment of existing impact metrics and proposed improvements. PLoS ONE 12(3): e0173152. https://doi.org/10.1371/journal.pone.0173152

RDA FAIR Data Maturity Model Working Group (2020) FAIR Data Maturity Model: specification and guidelines. Research Data Alliance. DOI: 10.15497/RDA00050

Remy L, Ivanović D, Theodoridou M, Kritsotaki A, Martin P, Bailo D, Sbarra M, Zhao Z, Jeffery K (2019) Building an integrated enhanced virtual research environment metadata catalogue. The Electronic Library.

Ribeiro L, De-Castro P, Mennielli M (2016) Final report: EUNIS–EUROCRIS joint survey on CRIS and IR. ERAI (Eunis Research and Analysis Initiative).

Robinson-Garcia N, Costas R, van Leeuwen TN (2020) Open Access uptake by universities worldwide. PeerJ 8:e9410 https://doi.org/10.7717/peerj.9410

Saenen B, Morais R, Gaillard V, Borrell-Damián L (2019) Research Assessment in the Transition to Open Science. 2019 EUA Open Science and Access Survey Results. https://eua.eu/resources/publications/888:research-assessment-in-the-transition-to-open-science.html

Sansone S-A, McQuilton P, Cousijn H, Cannon M, Chan WM, Callaghan S, et al. (2020) Data Repository Selection: Criteria That Matter. Zenodo. http://doi.org/10.5281/zenodo.4084763

Schmidt R (2020) Rethinking Research Assessment: Addressing Institutional Biases in Review, Promotion, and Tenure Decision-Making. The Official PLOS Blog. September 29, 2020.

Schöpfel J, Prost H, Rebouillat V (2017) Research Data in Current Research Information Systems. Procedia Computer Science 106: 305-320. https://doi.org/10.1016/j.procs.2017.03.030.

Schöpfel J, Azeroual O (2021) Current research information systems and institutional repositories: From data ingestion to convergence and merger. Editor(s): David Baker, Lucy Ellis, In Chandos Digital Information Review, Future Directions in Digital Information, Chandos Publishing, 2021, Pages 19-37, https://doi.org/10.1016/B978-0-12-822144-0.00002-1.

Schwardmann U, Fenner M, et al. (2020) Pid architecture for the EOSC: Report from the EOSC executive board working group (Wg) architecture PID task force (Tf). https://data.europa.eu/doi/10.2777/525581

Science Europe (2016) Position Statement on Research Information Systems. https://www.scienceeurope.org/media/qbziuyj2/se_positionstatement_ris_web.pdf

Science Europe (2019) Workshop Report on Digital Transformation in Scholarly Communication. https://www.scienceeurope.org/our-resources/workshop-report-digital-transformation/

Science Europe (2020) Position Statement and Recommendations on Research Assessment Processes. https://www.scienceeurope.org/our-resources/position-statement-research-assessment-processes/

Seibold H (2019) Research software and Open Science. https://www.fosteropenscience.eu/node/2637

Seppänen J-T, Värri H, Ylönen I (2020) Co-Citation Percentile Rank and JYUcite: a new network-standardized output-level citation influence metric and its implementation using Dimensions API. Reprint. bioRxiv https://doi.org/10.1101/2020.09.23.310052

Shrestha AK, Vassileva J (2018) Blockchain-Based Research Data Sharing Framework for Incentivizing the Data Owners. In: Chen S., Wang H., Zhang LJ. (eds) Blockchain – ICBC 2018. ICBC 2018. Lecture Notes in Computer Science, vol 10974. Springer, Cham. https://doi.org/10.1007/978-3-319-94478-4_19

Shrestha AK, Vassileva J (2019) User Data Sharing Frameworks: A Blockchain-Based Incentive Solution, 2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 2019, pp. 0360-0366, doi: 10.1109/IEMCON.2019.8936137.

Sile L, et al. (2017) European Databases and Repositories for Social Sciences and Humanities Research Output. Antwerp: ECOOM & ENRESSH.DOI:10.6084/m9.figshare.5172322

Smederevac S, Dejan Pajić, Radovanović S, Gilezan S, Čolović P, et al. (2020) Otvorena nauka: praksa i perspektive. https://open.uns.ac.rs/handle/123456789/16297

Sousa Pinto CS, Lopes PL, Valério LP, DeGóis LAA (2017) The Portuguese National Academic CV Platform, Procedia Computer Science,Volume 106, Pages 168-175, https://doi.org/10.1016/j.procs.2017.03.013.

Stall S, Yarmey L, Cutcher-Gershenfeld J, Hanson B, Lehnert K, Nosek B, Parsons M, Robinson E, Wyborn L (2019) Make scientific data FAIR. Nature 570, 27-29. doi: https://doi.org/10.1038/d41586-019-01720-7

Stehouwer H, Wittenburg P (2018) RDA Europe: Data Practices Analysis. http://hdl.handle.net/11304/926fd0c8-4737-4097-b027-ec4234f9f54d

Stuart D (2017) Data bibliometrics: metrics before norms. Online Information Review, Vol. 41 No. 3, pp. 428-435. https://doi.org/10.1108/OIR-01-2017-0008

Sveinsdottir,T, Proudman V, Davidson J (2020). An Analysis of Open Science Policies in Europe v6. http://doi.org/10.5281/zenodo.4005612

Sugimoto C, Work S, Larivière V, Haustein S (2016) Scholarly use of social media and altmetrics: A review of the literature. Journal of the Association for Information Science and Technology. 68. 10.1002/asi.23833.

Takaku M, Tanifuji M (2009) An integrated approach to Researcher Portfolio and Institutional Repository in NIMS eSciDoc. DRF International Conference 2009

Tanhua T, Pouliquen S, Hausman J, O'Brien K, Bricher P, de Bruin T, Buck JJH, Burger EF, Carval T, Casey KS, et al. (2019) Ocean FAIR Data Services. Frontiers in Marine Science 6: 440 DOI=10.3389/fmars.2019.00440

The EOSC executive board (2020) Six Recommendations for implementation of FAIR practice. Report of the European open science cloud FAIR working group.

https://op.europa.eu/en-GB/publication-detail/-/publication/4630fa57-1348-11eb-9a54-01aa75ed71a1/language-en

The Open Science and Research Initiative (2016) Framework for Open Science and Research. https://www.doria.fi/handle/10024/130363

The PASTEUR4OA consortium (2016) Open Access policy guidelines for research funders. http://pasteur4oa.eu/resources/149#.X2iGf1UzZaQ

The Serbian Ministry of Education, Science and Technological Development (2018) Serbian platform for open science (Platforma za otvorenu nauku). http://www.mpn.gov.rs/wp-content/uploads/2018/07/Platforma-za-otvorenu-nauku.pdf

Van Noorden R (2014) Online collaboration: Scientists and the social network. Nature. 512. 126-9. 10.1038/512126a.

Visser, M., van Eck, N. J., & Waltman, L. (2020). Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic. Arxiv. https://arxiv.org/ftp/arxiv/papers/2005/2005.10732.pdf

Vogel AL, Hall KL, Falk-Krzesinski HJ, Klein JT (2019) Broadening our Understanding of Scientific Work for the Era of Team Science: Implications for Recognition and Rewards. In: Hall K., Vogel A., Croyle R. (eds) Strategies for Team Science Success. Springer, Cham. https://doi.org/10.1007/978-3-030-20992-6_39

VSNU, KNAW, NWO (2020) Strategy Evaluation Protocol 2021–2027 (for Netherlands). https://www.vsnu.nl/files/documenten/Domeinen/Onderzoek/SEP_2021-2027.pdf

VSNU, NFU, KNAW, NWO, ZonMw (2019) Room for everyone's talent. https://www.vsnu.nl/recognitionandrewards/wp-content/uploads/2019/11/Position-paper-Room-for-everyone%E2%80%99s-talent.pdf

Walther T, Hauschke C, Kasprzik A (2019) The Research Core Dataset (KDSF) in the Linked Data context. Procedia Computer Science, 146, 29-38.

Wellcome (2020) What Researchers Think About the Culture They Work In. https://wellcome.org/sites/default/files/what-researchers-think-about-the-culture-they-work-in.pdf

Whyte A, Engelhart C, Bangert D, Kayumbi-Kabeya G, Lambert S, Thorle M, et al. (2019) D3.2 FAIR Data Practice Analysis (Version 1.0). Zenodo.

-Wildgaard LE (2015) Measure Up! The extent author-level bibliometric indicators are appropriate measures of individual researcher performance. Det Humanistiske Fakultet, Københavns Universitet

Wilkinson MD, et al. (2016) The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, *3*(1), 1-9.

Wilkinson MD et al. (2018) A design framework and exemplar metrics for FAIRness. Sci. Data 5:180118 doi: 10.1038/sdata.2018.118.

Wilkinson MD, Dumontier M, Sansone SA, et al. (2019) Evaluating FAIR maturity through a scalable, automated, community-governed framework. Sci Data 6, 174 (2019). https://doi.org/10.1038/s41597-019-0184-5

Wilsdon J, Allen L, Belfiore E, Campbell P, Curry S, Hill S, Jones R, Kain R, Kerridge S, Thelwall M, Tinkler J, Viney I, Wouters P, Hill J, Johnson B (2015) The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management. 10.13140/RG.2.1.4929.1363.

Wilsdon J, Bar-Ilan J, Frodeman R, Lex E, Peters I, Wouters P (2017) Next-generation metrics: responsible metrics and evaluation for open science. 10.2777/337729.

Wittenburg P, Lautenschlager M, Thiemann H, Baldauf C,Trilsbeek P (2020) FAIR practices in Europe. Data Intelligence 2(2020), 257–263. doi: 10.1162/dint_a_00048

Working group for responsible evaluation of a researcher (2020) Good practice in researcher evaluation. Recommendation for the responsible evaluation of a researcher in Finland. Responsible Research Series 7:2020. Open Science Coordination in Finland. TSV & TJNK: Helsinki.

Wouters P, Ràfols I, Oancea A, Kamerlin SCL, Holbrook JB, Jacob M (2019) Indicator frameworks for fostering open knowledge practices in science and scholarship. In: von Schomberg, R. (Ed.). Brussels: European Union. https://doi.org/10.2777/445286

ANNEX 3

SURVEY REPORT, METHODOLOGY AND QUESTIONNAIRE

ANNEX 3

SURVEY REPORT, METHODOLOGY AND QUESTIONNAIRE

Annex 3 dataset: *Survey for Academic Assessment Systems for Open Science & Research Data (anonymized).* Kathleen Gregory, Dragan Ivanović, Elina Koivisto, Henriikka Mustajoki, Elina Pylvänäinen, Janne Pölönen. doi: 10.5281/zenodo.4704393.

SURVEY FOR ACADEMIC ASSESSMENT SYSTEMS FOR OPEN SCIENCE & RESEARCH DATA – SURVEY REPORT

Key takeaways

- Local eInfrastructures used for assessments offer only very limited support for recording information about the diverse Open Science outputs and activities. Open access publications and self-archiving of publications is a relatively well-covered area; however, other aspects of Open Science activities and outputs require much further development of information systems and platforms.
- Web of Science and Scopus remain the major information sources and platforms (in addition to local eInfrastructures) to support research assessment, despite the criticism related to their lack of coverage of diversity of research outputs.
- Harvesting data from global resources and infrastructures suffers from missing and/or inconsistent use of PIDs for some entities, and requires additional effort by librarians and other personnel to check, consolidate and enrich information.
- Data used in the academic assessment and results of assessment are rarely made open, and even if they are, that is usually related to the group assessment.

1. BACKGROUND

Federation of Finnish Learned Societies (TSV) conducted this survey during December 2020 and January 2021 as a part of a project funded by the European Open Science Cloud Co-Creation program. The survey will serve as a reference for the project members to overview current state-of-the-art and identify gaps in responsible academic assessment of research and researchers throughout Europe.

2. PURPOSE

The purpose of this survey is to understand how current academic assessment systems and infrastructures support evaluating open science practices, particularly those related to research data. Responses from this survey will be used in writing the overview to understand the current state of information used in researcher assessment, and specifically how FAIR data figures in the picture. Moreover, it could be a starting point for creation of a vision of what and how academic assessment practices and available infrastructures should be changed in the next five years.

3. METHODOLOGY

The survey was conducted online using the LimeSurvey tool. Our target audience for the survey were individuals at institutions in Europe who are:

- *Technicians* - those responsible for developing and maintaining the technical systems used in assessments. Examples of these systems include current research information systems (CRIS), researcher profile systems, personnel management systems, etc.

  AND/OR who are

- *Data collectors* - those responsible for creating or collecting data used in assessments. Such individuals may work in libraries, human resource departments or in research support offices, for example.

The respondents were guaranteed confidentiality of the shared information.

The 28 survey questions were divided into the following main sections:

- Academic assessment systems (general questions)

- Open science in academic assessment systems (a selection of OS-CAM criteria)
- Research data in academic assessment systems
- Persistent identifiers in academic assessment systems

Questions related to demographics were also included to help contextualize responses. Moreover, at the end of the survey, participants were asked to identify and describe in free form any missing feature of their academic assessment infrastructure.

The following [Limesurvey question types]() were used: multiple choice, single choice, free text, array number checkboxes. The complete list of questions can be found at https://docs.google.com/document/d/15JoWBCrf-gYbicuKAd45zIHMIT3iaqjSLh8yk1l9S_g/edit#heading=h.74av9zwg92le.

The following terms used in the survey were defined and shown to the survey participants before responding to questions:

- **Academic assessment** - processes of evaluation completed within research & academia. These processes occur at different levels (i.e. for researchers, departments and institutions) and for different purposes. The terms "research(er) evaluation" or "research(er) assessment" are often used to convey the same idea.

- **Academic assessment systems** - technical systems and infrastructures used during academic assessments. Examples of these systems include current research information systems (CRIS), researcher profile systems, personnel management systems, etc. These systems may be used by applicants (i.e., researchers applying for a job or promotion); by data collectors (i.e., people collecting data for academic assessments); or by individuals conducting the assessments.

- **Open science/open research** - FOSTER defines open science as "the practice of science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods."

- **Research data** - representations of observations, objects, or other entities used as evidence for the purposes of research or scholarship (Borgman, 2015)

- **FAIR data** - data which meet the data management principles of findability, accessibility, interoperability, and reusability, as defined by Wilkinson and colleagues here.

- **PID (persistent identifier)** - a long-lasting reference to a digital object that is accessible over the Internet.

The survey was distributed via the following networks:

- YERUN working group on Open Science
- ENRESSH members
- euroCRIS members
- LIBER Europe members
- The PID Forum (https://www.pidforum.org)
- DLF-ANNOUNCE@LISTS.CLIR.ORG, RADICALOPENACCESS@JISCMAIL.AC.UK, JISC-REPOSITORIES@JISCMAIL.AC.UK, LIS-BIBLIOMETRICS@JISCMAIL.AC.UK mailing lists' members

4. RESULTS

We have collected 24 responses in total. Twenty-one of these were complete responses; the remaining three respondents completed 70%-92% of the survey questions. We include answers from these partially-complete responses in our below analysis. Five among participants are responsible for maintaining/developing the technical systems used in academic assessments, four participants are responsible for creating or collecting data used in assessments, while four work in another role but have knowledge of the technical systems used in academic assessments. Majority of respondents are affiliated with universities (17), while there are 4 affiliated with research institutions (4), 2 with other types of institutions, and there is one retired participant not affiliated with any institution. Respondents are employed in 16 European countries: the Netherlands (5), France (2), and Armenia, Austria, Belgium, Bulgaria, Czech Republic, Estonia, Finland, Germany, Latvia, Norway, Spain, Sweden, Turkey, UK (1 per country). Moreover, one participant is affiliated with a non-European (Australian) institution. Those institutions are performing assessment

- at different levels:
  - Research groups or departments are assessed (19)
  - Individual researchers are assessed (17)
  - Applications for funding are assessed (17)
  - Assessments of the institution as a whole are performed (16)
  - Research projects are assessed at particular points in the project (i.e. mid-term evaluation, when a project is completed) (10)
- for different purposes:
  - For career promotion or advancement (i.e., for seeking tenure) (17)
  - For funding allocations (i.e., based on grant applications) (15)
  - When hiring new employees (12)
  - For awarding prizes (11)
  - To incentivise desired behaviour (11)
- and for different disciplines:
  - Social sciences and humanities (19)
  - Engineering and technological sciences (18)
  - Natural sciences and mathematics (16)
  - (Bio)medical sciences (16)

A. General Questions about Academic Assessment Systems

Various platforms are used to support the academic assessment:

- Current research information systems (CRIS) (13)
  - Pure (6), Converis (1), Sympletic Elements (1), in-house solution (5)
- Personnel management systems (8)
- Researcher profile systems or tools to create online CVs and academic profiles (7)
- Other local platforms (9)

Data for the academic assessment have been submitted or uploaded using CRIS or other local platforms (15) or other widely available resources such as:

- Using academic databases (17)
  - Web of Science (12),
  - Scopus (10),

    - ORCID (1),
    - EBSCO (1),
    - Scimago Journal Rank (1)
- Using altmetrics databases (4)
    - PlumX (1),
    - Altmetric.com (3),
    - SciVal (2)

Therefore, we can conclude that besides all criticism related to the comprehensiveness and coverage of academic and altmetric databases, those systems are used in majority of academic assessment policies at least as a supplement to the data preserved in local eInfrastructures. By analysing free form descriptions of data submission processes, we found out that some institutions are using librarian and other support professionals to check, consolidate and enrich data harvested from global resources.

Moreover, we found that the data submission process in some respondents' institutions depends on the purpose of academic assessment (hiring, promotion, funding, etc.) and organization units (departments, faculties), and can be guided by uploading data in prescribed format in the local eInfrastructures. Data are submitted/uploaded in structured formats and in unstructured formats in the assessment process at 15 and 5 respondents' institutions, respectively. In order to motivate researchers to catalogue their results using the local eInfrastructure for the needs of group assessment (department and institution), some institutions are using salary complements defined by published results which have to be catalogued in the local eInfrastructures.

Academic assessment processes at 10 institutions also include ways to provide narrative descriptions of research or impact; 9 institutions do not and 5 respondents did not know whether or not this was possible. Those narrative descriptions are usually linked with individual assessment, while are missing for group assessment (institutional). Qualitative data is collected via free text fields or by researchers creating reports of varying lengths to describe their work. For some, this type of qualitative data is optional, while for others it is required. Peer-reviewers of both promotion and tenure applications as well as grant proposals also provide comments on applications. However, there are also assessments which only require quantitative data.

Besides storing data used in the assessment (12) and making it available for participants of the assessment process (7), some institutions also store results of assessment (10) and even make it open after the completion of the assessment process (6). However, those publishing of assessments' results is usually related to group assessment, while results of individual assessment are usually not published.

The assessment process has been enhanced by usage of information communication technologies using the following techniques:

- Some entities are automatically linked (i.e., linking researchers to publications using ORCIDs) (13)
- Some data and indicators are automatically collected using available APIs (11)
- Collected data are exported to global and local platforms (i.e., ORCID, national database) (6)
- The fulfilment of some evaluation criteria are automatically determined, based on assessment data and any assigned indicators (i.e., only candidates with a certain number of citations will be considered) (5)

Four participants responded their institutions are not using any automated approaches.

B. Open Science in Academic Assessment Systems

To gain an overview of the extent to which assessment systems support evaluation of Open Science outputs, practices, and activities, a selection of OS-CAM (Open Science Career Assessment Matrix) criteria was used (or slightly modified). As shown in Figure 1, the combined results from all 24 academic assessment systems included in the survey show a great variety in enabling the recording the information (fully, somewhat or not at all) about various research outputs and processes, as well as services and leadership activities, and teaching and professional development activities related to open science.

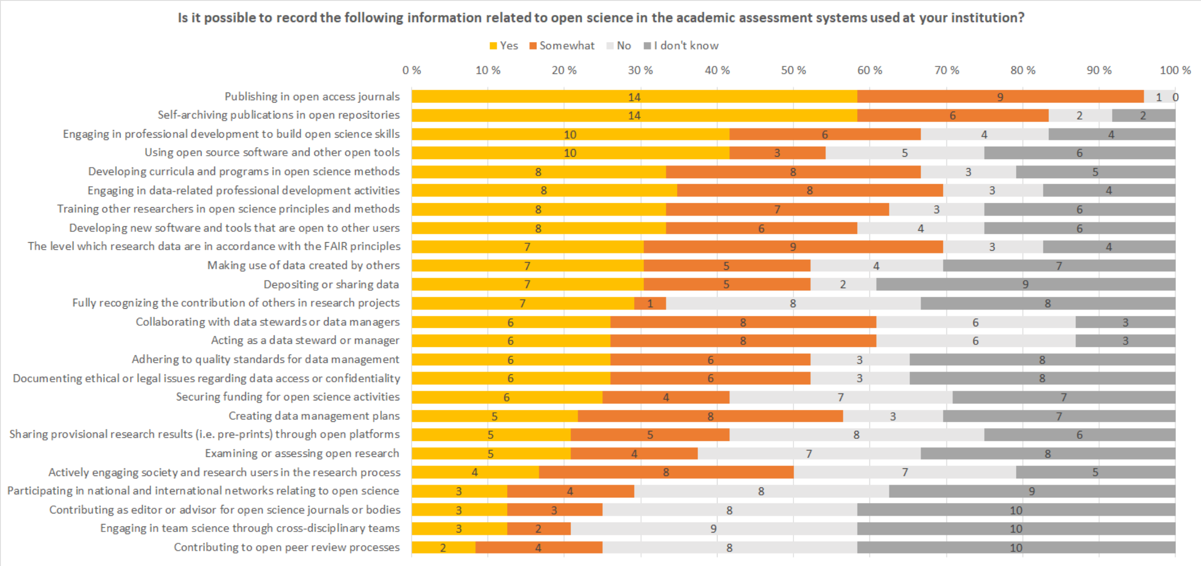


Figure 1. Storing of Open Science related research outputs, processes, service and leadership activities, and teaching and professional development activities in academic assessment systems.

Moreover, some respondents indicated in the free text comment that some of that information can be recorded in local eInfrastructures (particularly in the Pure CRIS), but that there is no obligation to record that information in the database. Similarly, some respondents stated that information about open science outputs and activities can be shared using existing text-fields; these text-fields do not specifically ask for information related to open science, however.

> “Researchers have the ability to record some of these open science practices in Pure, but not the obligation to. Open access publication is something we record for all output."

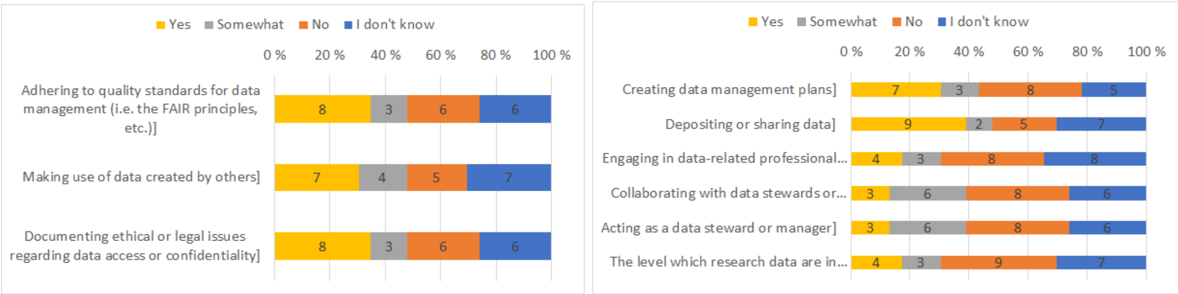
> “Again, anything can be shared using free-text boxes, but that’s not to say its specifically asked for.”

One respondent described the approach to assessing the open science activities of research units or groups in the Netherlands. Here, the overall open science and research strategy and policy, as well as plans for establishing open science guidelines, are taken into account.

Furthermore, some respondents stated the support for recording this information is in the development.

C. Research Data in Academic Assessment Systems

Figure 3 shows supporting storage of the information about research data, as well as additional information about activities related to research data by the academic assessment systems used at the survey participants' institutions.



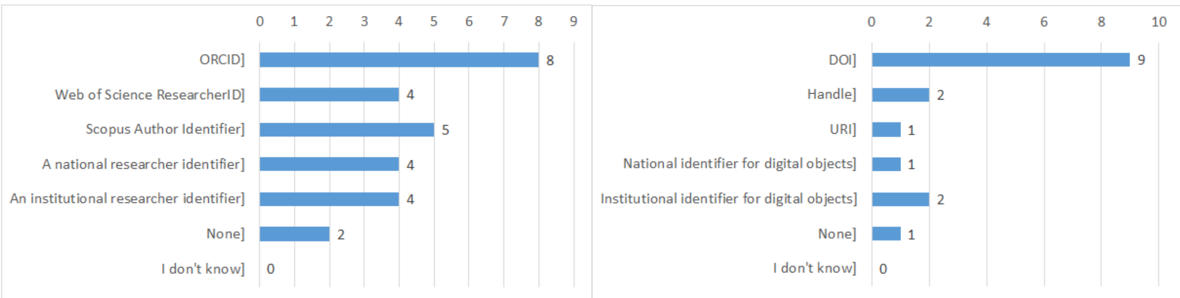
(a) (b)

Figure 3. Storing of information about research data (a) and additional information about activities related to research data in academic assessment systems.

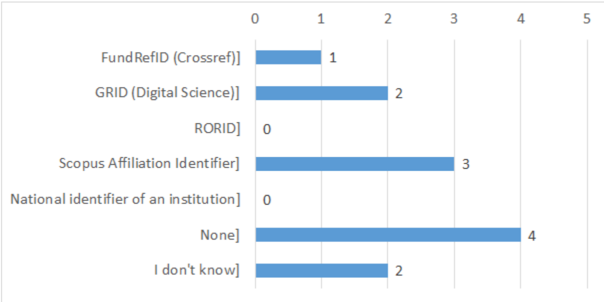
Similarly, as above, one respondent indicated in a free-text comment that some of that information can be recorded in local eInfrastructures (particularly in the Pure CRIS), but are not used in assessment at the moment. Furthermore, a few respondents stated there are plans to support storage of those information, but the development is in the early stage. Again, some respondents emphasize that free-text (narrative) descriptions can be used for describing published research data and related activities.

D. Persistent Identifiers (PIDs) used in Academic Assessment Systems

Figure 4 shows usage of persistent identifiers (PIDs) for researchers, digital research outputs, and institutions in academic assessment systems at respondents' institutions. The results reflect the growing establishment of using ORCIDs for individual researchers and DOIs for research outputs documented elsewhere.



(a) (b)



(c)

Figure 4. Usage of researchers' persistent identifiers (a), digital research outputs' persistent identifiers (b), and institutions' persistent identifiers.

Moreover, one participant stated a local or Cordis identifiers have been assigned to the projects, as well as to funders. On the other side one participant stated the standardized PID for the project entity is missing at the moment. Furthermore, introducing PIDs for peer review and editorial activities have been suggested by a survey participant, as well as PIDs for anything that is not electronically published (for instance, books).

5. MISSING FEATURES

At the end of the survey, participants were asked to identify and describe any missing feature of their academic assessment infrastructure. One participant complained that data from sources for the evaluation of book publishers (SPI) are not harvested by their platforms for the needs of academic assessment. Moreover, there were a couple more identified issues for academic assessment which are more related to policies than to academic assessment systems:

- One of the identified issues was that criteria for assessment of large and small organization units shouldn't be the same, as well as for young and senior researchers, meaning criteria should depend on the size of the group for institutional assessment, and should depend on seniority for individual assessment.
- Extension of assessment criteria which includes not-publication based results (project applications, software, supervising, etc.) are needed.
- Furthermore, although open science policies have been introduced a couple of years ago at some institutions, the effects are still not visible, meaning scholars haven't changed their old (closed science) practices yet.

6. CONCLUSION

Twenty-four people responsible for maintaining/developing the technical systems used in academic assessments, or for creating or collecting data used in assessments participated in the survey. Those people are affiliated with institutions which perform academic assessment at various levels, and for various purposes and disciplines. The survey's analysis shows that local eInfrastructures are used to support the academic assessment process. However, those infrastructures are not yet well developed, especially in the domain of supporting recording information about Open Science outputs and activities including research dataset and activities related to dataset. Even if some aspects of Open Science outputs are supported by some academic assessment systems, it is usually related to open access publications and self-archiving of publications. However, there is the recognition of needs for recording other aspects of Open Science activities and outputs and its inclusion in the academic assessment process, and a few respondents emphasized that this extension is under development at their institutions, but still in the early stage.

Besides all criticism related to the comprehensiveness and coverage of global platforms and its databases (e.g., Web of Science, Scopus), those systems are used in majority of academic assessment policies at least as a supplement to the data preserved in local eInfrastructures. However, additional effort by librarians and other officers are needed in order to check, consolidate and enrich data harvested from global resources. Missing of PIDs for some entities or inconsistent usage of some PIDs are one of the roots of those harvesting problems.

Data used in the academic assessment and results of assessment are rarely made open, and even if they are, that is usually related to the group assessment.

Therefore, this survey confirmed the following conclusions and gaps previously identified by performing comprehensive review of state of the art in the field of researcher assessment movement in Open Science (see the section 2):

- The importance of accounting for other types of research output in researchers' assessment has been recognized, but not yet fully adopted in policies or in practice.

- Guidelines and policies for encouraging open access publications are much better developed/adopted than for open data (and other research objects).
- Technologies do not bring change on their own, but eInfrastructures should support Open Science and evaluation practices for Open Science.
- Research entities isolated in numerous systems usually can't be used for fair and comprehensive assessment of a researcher, thus integration of those systems and linking their research entities instances are necessary for the purpose of assessment, as well as for increasing discoverability and reusing of open research objects.
- The focus in openness/FAIRness of data used in research evaluations is usually on making the criteria and methods used in evaluations open, including transparency in indicator development. Fewer documents call for making the data used in evaluations open.
- There is a need to develop a scholarly infrastructure of interoperable, linked, transparent systems to support open science and robust evaluations which can be based on persistent identifiers assigned to each research ecosystem entity to enable tracking indicators for Open Science results.
- There are well-known persistent identifiers for some research entity types, but not for all.

QUESTIONNAIRE – SURVEY FOR ACADEMIC ASSESSMENT SYSTEMS FOR OPEN SCIENCE & RESEARCH DATA

PART 1. DEMOGRAPHICS AND CONTEXT

Definition of term used in the questions:

Academic assessment - processes of evaluation completed within research & academia. These processes occur at different levels (i.e. for researchers, departments and institutions) and for different purposes. The terms "research(er) evaluation" or "research(er) assessment" are often used to convey the same idea.

*** Question 1. What are your professional responsibilities?**

Please choose **all** that apply:

- ☐ I am responsible for maintaining/developing the technical systems used in academic assessments.
- ☐ I am responsible for creating or collecting data used in assessments.
- ☐ I work in another role but have knowledge of the technical systems used in academic assessments.

*** Question 2. Please state your professional role (i.e. librarian, professor, systems engineer, etc.)**

Please write your answer here:

*** Question 3. At which type of institution do you work?**

Please choose **only one** of the following:

- ◯ University
- ◯ Research institution
- ◯ Governmental organization
- ◯ Research funding organization
- ◯ Other

Question 3a. What is the name of your institution?

Please write your answer here:

*** Question 4. In which country is your institution located?**

Please write your answer here:

*** Question 5. At which level are academic assessments performed at your institution?**

Please choose **all** that apply:

☐ Individual researchers are assessed

☐ Applications for funding are assessed

☐ Research projects are assessed at particular points in the project (i.e. mid-term evaluation, when a project is completed)

☐ Research groups or departments are assessed

☐ Assessments of the institution as a whole are performed

☐ I do not know which types of assessments are performed

☐Other:

*** Question 6. For which of the below purposes are assessments made at your institution?**

Please choose **all** that apply:

☐ For career promotion or advancement (i.e. for seeking tenure)

☐ For funding allocations (i.e. based on grant applications)

☐ When hiring new employees

☐ For awarding prizes

☐ To incentivise desired behaviour

☐ None of the above applies

☐ I don’t know

☐Other:

*** Question 7. For which disciplines do you perform assessments?**

Please choose **all** that apply:

- ☐ Natural sciences and mathematics
- ☐ Engineering and technological sciences
- ☐ (Bio)medical sciences
- ☐ Social sciences and humanities
- ☐ I don’t know
- ☐ Other:

PART 2.1. GENERAL QUESTIONS ABOUT ACADEMIC ASSESSMENT SYSTEMS

Definition of term used in the question:

Academic assessment - processes of evaluation completed within research & academia. These processes occur at different levels (i.e. for researchers, departments and institutions) and for different purposes. The terms “research(er) evaluation” or “research(er) assessment” are often used to convey the same idea.

*** Question 8. Which local platforms are used in academic assessment? (Local platforms are systems running internally at your institution).**

Please choose **all** that apply:

- ☐ Current Research Information Systems (CRIS), (i.e. PURE, Converis, an in-house solution, etc.).
- ☐ Personnel management systems
- ☐ Researcher profile systems or tools to create online CVs and academic profiles (i.e. Sympletic Elements Platform, an in-house solution, etc.)
- ☐ Other local platforms

☐ None. We do not use local platforms in assessment.

☐ I don’t know

Question 8a. If known, please also indicate the names of local platforms used (i.e. PURE, Converis, Sympletic Elements Platform, an in-house solution, etc.).

Please write your answer here:

*** Question 9. How are data (i.e. outputs, citations, funding) for academic assessments captured?**

Please choose **all** that apply:

☐ Data are submitted/uploaded using CRIS or other local platforms such as those described in Question 8.

☐ Using global academic databases (i.e. Web of Science, Scopus, ORCID, CrossRef, etc.)

☐ Using altmetrics databases (i.e. PlumX, Altmetric.com, etc.)

☐ Data are submitted/uploaded in structured formats (i.e. templates are provided, structured web formats are specified, detailed instructions are given, etc.)

☐ Data are submitted/uploaded in unstructured formats (no templates are provided, no structured web formats are specified, no detailed instructions are given, etc.)

☐ I don’t know

☐Other:

Question 9a. If known, please write the name of the resources and databases (Web of Science, Scopus, PlumX, Altmetric, etc.) used in academic assessments.

Please write your answer here:

Question 9b. Please briefly describe the submission process for uploading/collecting the data used in assessments. Please include any information about prescribed templates or specific formats used for submitting data.

Please write your answer here:

PART 2.2. GENERAL QUESTIONS ABOUT ACADEMIC ASSESSMENT SYSTEMS

Definition of term used in the question:

Academic assessment - processes of evaluation completed within research & academia. These processes occur at different levels (i.e. for researchers, departments and institutions) and for different purposes. The terms "research(er) evaluation" or "research(er) assessment" are often used to convey the same idea.

*** Question 10. Do the systems used in academic assessments at your institution include ways to provide narrative descriptions of research or impact?**

Please choose **only one** of the following:

- ○ Yes. Please describe how narrative descriptions are captured in the system.
- ○ No.
- ○ I don't know.

Make a comment on your choice here:

PART 2.3. GENERAL QUESTIONS ABOUT ACADEMIC ASSESSEMNT SYSTEMS

Definition of term used in the question:

Academic assessment systems - technical systems and infrastructures used during academic assessments. Examples of these systems include current research information systems (CRIS), researcher profile systems, personnel management systems, etc. These systems may be used by applicants (i.e. researchers applying for a job or promotion); by data collectors (i.e. people collecting data for academic assessments); or by individuals conducting the assessments.

*** Question 11. Which of the following describe how assessment data are stored and made available at your institution?**

Please choose **all** that apply:

☐ There is a platform for storing data used in assessments

☐ There is a platform for storing the results of assessments

☐ Data used in assessments are openly available to all involved in the assessment process (evaluators, applicants, etc.)

☐ Assessment results are made openly available once assessments are complete

☐ I don't know

☐Other:

Question 11a. If necessary, please describe any of your responses further.

Please write your answer here:

PART 2.4. GENERAL QUESTIONS ABOUT ACADEMIC ASSESSMENT SYSTEMS

*** Question 12. Which automated approaches for assessment does your institution use?**

Please choose **all** that apply:

☐ Some entities are automatically linked (i.e. linking researchers to publications using ORCIDs)

☐ Some data and indicators are automatically collected using available APIs

☐ Collected data are exported to global and local platforms (i.e. ORCID, national database)

☐ The fulfillment of some evaluation criteria are automatically determined, based on assessment data and any assigned indicators (i.e. only candidates with a certain number of citations will be considered)

☐ None. We do not use automated approaches.

☐ I don't know.

☐ Other:

PART 3.1. OPEN SCIENCE IN ACADEMIC ASSESSMENT SYSTEMS

These questions were formulated according to particular sections of the Open Science Career Assessment Matrix (OS-CAM).

Definitions of terms used in the questions:

Academic assessment - processes of evaluation completed within research & academia. These processes occur at different levels (i.e. for researchers, departments and institutions) and for different purposes. The terms “research(er) evaluation” or “research(er) assessment” are often used to convey the same idea.

Open science/open research - FOSTER defines open science as “the practice of science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods.”

*** Question 13. Is it possible to record the following information about research outputs related to open science in the academic assessment systems used at your institution?**

Please choose the appropriate response for each item:

	Yes	Somewhat	No	I don't know
Publishing in open access journals	☐	☐	☐	☐
Self-archiving publications in open repositories	☐	☐	☐	☐
Using open source software and other open tools	☐	☐	☐	☐

Developing new software and tools that are open to other users	☐	☐	☐	☐
Securing funding for open science activities	☐	☐	☐	☐

Question 13a. If necessary, please describe any of your responses further.

Please write your answer here:

PART 3.2. OPEN SCIENCE IN ACADEMIC ASSESSMENT SYSTEMS

Definitions of terms used in the questions:

Academic assessment systems - technical systems and infrastructures used during academic assessments. Examples of these systems include current research information systems (CRIS), researcher profile systems, personnel management systems, etc. These systems may be used by applicants (i.e. researchers applying for a job or promotion); by data collectors (i.e. people collecting data for academic assessments); or by individuals conducting the assessments.

Open science/open research - FOSTER defines open science as "the practice of science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods."

*** Question 14. Is it possible to record the following information about research processes related to open science in the academic assessment systems used at your institution?**

Please choose the appropriate response for each item:

	Yes	Somewhat	No	I don't know
Actively engaging society and research users in the research process (i.e. citizen science efforts)	☐	☐	☐	☐
Sharing provisional research results (i.e. pre-prints) through open platforms (e.g. Arxiv, Figshare)	☐	☐	☐	☐
Engaging in team science through cross-disciplinary teams	☐	☐	☐	☐

Fully recognizing the contribution of others in research projects, including collaborators, co-authors, citizens, open data providers	☐	☐	☐	☐

Question 14a. If necessary, please describe any of your responses further.

Please write your answer here:

PART 3.3. OPEN SCIENCE IN ACADEMIC ASSESSMENT SYSTEMS

Definitions of terms used in the questions:

Academic assessment systems - technical systems and infrastructures used during academic assessments. Examples of these systems include current research information systems (CRIS), researcher profile systems, personnel management systems, etc. These systems may be used by applicants (i.e. researchers applying for a job or promotion); by data collectors (i.e. people collecting data for academic assessments); or by individuals conducting the assessments.

Open science/open research - FOSTER defines open science as “the practice of science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods.”

*** Question 15. Is it possible to record the following information about service and leadership activities related to open science in the academic assessment systems used at your institution?**

Please choose the appropriate response for each item:

	Yes	Somewhat	No	I don't know
Contributing as editor or advisor for open science journals or bodies	☐	☐	☐	☐
Contributing to open peer review processes	☐	☐	☐	☐
Examining or assessing open research	☐	☐	☐	☐
Participating in national and international networks relating to open science	☐	☐	☐	☐

Question 15a. If necessary, please describe any of your responses further.

Please write your answer here:

PART 3.4. OPEN SCIENCE IN ACADEMIC ASSESSMENT SYSTEMS

Definitions of terms used in the questions:

Academic assessment systems - technical systems and infrastructures used during academic assessments. Examples of these systems include current research information systems (CRIS), researcher profile systems, personnel management systems, etc. These systems may be used by applicants (i.e. researchers applying for a job or promotion); by data collectors (i.e. people collecting data for academic assessments); or by individuals conducting the assessments.

Open science/open research - FOSTER defines open science as "the practice of science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods."

*** Question 16. Is it possible to record the following information about teaching and professional development activities related to open science in the academic assessment systems used at your institution?**

Please choose the appropriate response for each item:

	Yes	Somewhat	No	I don't know
Training other researchers in open science principles and methods	☐	☐	☐	☐

Developing curricula and programs in open science methods, including open science data management	☐	☐	☐	☐
Engaging in professional development to build open science skills	☐	☐	☐	☐

Question 16a. If necessary, please describe any of your responses further.

Please write your answer here:

PART 4.1. RESEARCH DATA IN ACADEMIC ASSESSMENT SYSTEMS

Definitions of terms used in the questions:

Academic assessment systems - technical systems and infrastructures used during academic assessments. Examples of these systems include current research information systems (CRIS), researcher profile systems, personnel management systems, etc. These systems may be used by applicants (i.e. researchers applying for a job or promotion); by data collectors (i.e. people collecting data for academic assessments); or by individuals conducting the assessments.

Research data - representations of observations, objects, or other entities used as evidence for the purposes of research or scholarship (Borgman, 2015)

FAIR data - data which meet the data management principles of findability, accessibility, interoperability, and reusability, as defined by Wilkinson and colleagues here.

*** Question 17. Is it possible to record the following information about research data in the academic assessment systems used at your institution?**

Please choose the appropriate response for each item:

	Yes	Somewhat	No	I don't know
Adhering to quality standards for data management (i.e. the FAIR principles, etc.)	☐	☐	☐	☐
Making use of data created by others	☐	☐	☐	☐
Documenting ethical or legal issues regarding data access or confidentiality	☐	☐	☐	☐

Question 17a. If necessary, please describe any of your responses further.

Please write your answer here:

PART 4.2. RESEARCH DATA IN ACADEMIC ASSESSMENT SYSTEMS

Definitions of terms used in the questions:

Research data - representations of observations, objects, or other entities used as evidence for the purposes of research or scholarship (Borgman, 2015)

Fair data - data which meet the data management principles of findability, accessibility, interoperability, and reusability, as defined by Wilkinson and colleagues here.

*** Question 18. Is it possible to record information about the following additional information about activities related to research data?**

Please choose the appropriate response for each item:

	Yes	Somewhat	No	I don't know
Creating data management plans	☐	☐	☐	☐
Depositing or sharing data	☐	☐	☐	☐
Engaging in data-related professional development activities (i.e. classes on data management, sharing or reuse)	☐	☐	☐	☐

Collaborating with data stewards or data managers (i.e. through consultations)	☐	☐	☐	☐
Acting as a data steward or manager	☐	☐	☐	☐
The level which research data are in accordance with the FAIR principles. **Please explain in the question 18a how these levels of FAIRness are determined.**	☐	☐	☐	☐

Question 18a. If necessary, please describe any of your responses further.

Please write your answer here:

PART 5. PERSISTENT IDENTIFIERS (PIDs) USED IN ACADEMIC ASSESSMENT SYSTEMS

Definitions of terms used in the questions:

PID (persistent identifier) - a long-lasting reference to a digital object that is accessible over the Internet.

Academic assessment - processes of evaluation completed within research & academia. These processes occur at different levels (i.e. for researchers, departments and institutions) and for different purposes. The terms “research(er) evaluation” or “research(er) assessment” are often used to convey the same idea.

Academic assessment systems - technical systems and infrastructures used during academic assessments. Examples of these systems include current research information systems (CRIS), researcher profile systems, personnel management systems, etc. These systems may be used by applicants (i.e. researchers applying for a job or promotion); by data collectors (i.e. people collecting data for academic assessments); or by individuals conducting the assessments.

*** Question 19. Which persistent identifiers (PIDs) are used for researchers in academic assessment systems at your institution?**

Please choose **all** that apply:

☐ ORCID

☐ Web of Science ResearcherID

☐ Scopus Author Identifier

☐ A national researcher identifier

☐ An institutional researcher identifier

☐ None

☐ I don't know

☐Other:

*** Question 20. Which persistent identifiers (PIDs) are used for digital research outputs in academic assessment systems at your institution?**

Please choose **all** that apply:

☐ DOI

☐ Handle

☐ URI

☐ National identifier for digital objects

☐ Institutional identifier for digital objects

☐ None

☐ I don't know

☐Other:

*** Question 21. Which persistent identifiers (PIDs) are used for institutions in academic assessment systems at your institution?**

Please choose **all** that apply:

☐ FundRefID (Crossref)

☐ GRID (Digital Science)

☐ RORID

☐ Scopus Affiliation Identifier

☐ National identifier of an institution

☐ None

☐ I don't know

☐Other:

Question 22. Which persistent identifiers (PIDs) are used for other entities (projects, services, indicators, etc) in academic assessment systems at your institution, if any?

Please write your answer here:

Question 23. Which entities of interest to academic assessments do not have a well-adopted PID?

Please write your answer here:

CLOSING

Definition of term used in the question:

Academic assessment systems - technical systems and infrastructures used during academic assessments. Examples of these systems include current research information systems (CRIS), researcher profile systems, personnel management systems, etc. These systems may be used by applicants (i.e. researchers applying for a job or promotion); by data collectors (i.e. people collecting data for academic assessments); or by individuals conducting the assessments.

Question 24. In your opinion, which gaps or missing features do academic assessment systems at your institutions have?

Please write your answer here:

Question 25. Do you have anything else that you would like to add?

Please write your answer here:

Question 26. Would you like to:

Please choose **all** that apply:

- ☐ Receive information about the results of this survey
- ☐ Be contacted for follow up research

Question 27. Please enter email address so we may contact you regarding your above choice.

Please write your answer here:

ANNEX 4

OS-CAM CASE STUDIES

ANNEX 4

OS-CAM CASE STUDIES

Annex 4 dataset: *OS-CAM Case Studies*. Dragan Ivanović, Valerie Brasse, Joonas Kesäniemi, Janne Pölönen. doi: 10.5281/zenodo.4704425.

This report presents five case studies of usage of research databases, models and platforms for responsible academic career assessment based on Open Science outputs, activities and expertise. The case studies are part of the EOSC co-creation project “Study Proposal #16: European overview of career merit systems” (https://avointiede.fi/en/networks/eosc-co-creation).

The Open Science Career Advancement Matrix (OS-CAM) is used as a well-established framework of assessment criteria for different types of open science outputs and activities. The five case studies have different contexts: one is a standardized data model and format (CERIF), while the second one is a global platform for researchers’ profiles (ORCID), and the other three case studies cover national research information platforms (Research.fi - Finland, NARCIS - The Netherlands, CRIStin - Norway).

The OS-CAM criteria, as well as the evaluation methodology and sources analysed in the five case studies, are described in section 1, followed by section 2 on results. Conclusions and recommendations are provided in section 3. These include recommendations for further refinement of the OS-CAM criteria, as well as for improving the coverage of OS-CAM criteria in assessment infrastructures.

Key takeaways

1. The OS-CAM criteria are often very difficult to apply in practice. The OS-CAM criteria should be described in much more detail, and they should be joined with a set of examples about the evidence of fulfilling those criteria.
2. The OS-CAM criteria related to the well-established and popular ways of acquiring academic merit such as journal articles and other traditional research outputs are much better covered by analysed infrastructures

than more holistic ones that include for example social impact, teaching and professional experience related information.

3. Missing Open Science related terms and vocabularies are the main obstacle to supporting the OS-CAM criteria in assessment infrastructures

1. AIM, METHODS AND DATA

Rewarding researchers for diverse open science practices requires reliable, comprehensive, well-structured and comparable data and metrics to support the assessment process. Various infrastructures are needed for and involved in integrating qualitative and quantitative data from, and facilitating interoperability between, international, national and institutional research information systems, databases and data models. The aim of this study is to analyse to what extent an OS-CAM recommendations-based academic career assessment could be built (based on automated check and/or requiring manual work) on top of five international and national research information platforms:

- CERIF (Common European Research Information Format)
- ORCID (Open Researchers and Contributor ID)
- Research.fi (Finland)
- NARCIS (The Netherlands)
- CRIStin (Norway)

OS-CAM

The Open Science Career Advancement Matrix (OS-CAM) was developed in 2017 by a Working Group on Rewards under Open Science, which was created with the mandate from the European Commission's Open Science Policy Platform (OSPP). The OS-CAM was published in the OSPP report *Evaluation of Research Careers fully acknowledging Open Science practices*. The aim of the working group was to “make recommendations in order that all researchers in Europe are recognised and rewarded for practising Open Science”. Therefore, the OS-CAM could serve as one toolkit, or stand on its own as an evaluative framework for researchers. The OS-CAM, as stated in the report, “represents a possible, practical move towards a more comprehensive approach to evaluating researchers through the lens of Open Science”. The framework provides concrete examples of assessment’s criteria for different types of open science outputs and activities across phases of research. The OS-CAM criteria could be used in various contexts such as grant evaluation/job applications/funding models, and at different levels, for example “learning about open science” for first stage researchers or “doing open science” for recognised researchers. The OS-CAM outlines six categories of Open Science activities, divided into 23 Open Science activities and 42 examples of evaluation criteria (the list of all OS-CAM criteria is provided in Appendix 1).

EVALUATION METHODOLOGY

Coverage of each of the 42 OS-CAM criteria in each of the five systems was analysed and classified using the following five levels:

- Fully - If an automatic check of fulfilling the OS-CAM criterion in the system is possible, it is classified as fully supported
- Very - an automatic check of fulfilling the OS-CAM criterion is almost fully supported, and just an administrative check is needed using the provided metadata (e.g., title, url, etc.)
- Partially - an expert should estimate whether the criterion is reached (e.g., whether it is related to open science, level of engagement, etc.)
- Slightly - the provided information is used only as a basis for further looking for information on other sources, and a complex assessment by an expert is required (e.g., using FAIR principles, adopting quality standards, etc.)
- Not at all - if there is only free-text biography (narrative field), it is classified as not at all

The evaluation was performed in three rounds:

- First, each system or data model under evaluation was assessed by an individual project participant who also assigned preliminary levels for each criteria.
- Next, the preliminary results involving national services, i.e., research.fi, NARCIS and CRIStin, were communicated to people involved with day to day operations of each system and meetings were organized between project participants and domain experts. Feedback was solicited for each criteria and changes to first results were discussed. In order to minimise

the chance of underrepresentation, the focus of these discussions was on the criteria where preliminary results have shown little or no support.
- Finally, project participants involved in the evaluation went through all the criteria for all the systems and data models together, in an effort to harmonise the use of different levels in respect to each assessment criteria.

Although project participants involved in evaluation have strived for a common understanding about the meaning of the OS-CAM criteria, they have also restricted themselves to a fairly narrow interpretation of each criteria. For example, for the OS-CAM evaluation criterion “Widening participation in research through open collaborative projects”, one or two examples on how the criterion in question could manifest itself were shared and the evaluation was based on those examples. This was done mainly in an effort to have something concrete to work with, which can facilitate common understanding. Coming up with examples for the more ambiguous criteria examples, such as “Pushing forward the boundaries of open science as a research topic”, was not always easy.

Conceptual differences, system vs. data models, and availability of publicly available information about the systems under evaluation have had some influence on evaluation results. Participants also have different background knowledge about the systems and models, but have tried to minimise possible biases by requiring a full quorum for the final evaluation level decisions.

The scope of the evaluation as a whole is also quite narrow. Focus on the OS-CAM was solely made with a very strict career assessment viewpoint. The Open Science theme could and should be evaluated also from other perspectives.

Detailed results of the evaluation are available in a Google Sheet document at https://docs.google.com/spreadsheets/d/1igbR5NgZzG070QN2DPJiKl6GJ_5nbX2WU9wyYC5macM/edit?usp=sharing.

SYSTEMS AND MODELS

CERIF

CERIF (the Common European Research Information Format) is both a data model covering all aspects of research information and an information format for the exchange of information between CRISs (Current Research Information Systems) and other information systems (such as OA Repositories or Research Data systems), on a local, national or international level. CERIF is a European Union initiative, the development of which has been entrusted since 2002 to euroCRIS, a non-profit organisation registered in the Netherlands.

CERIF is being recommended by the European Commission to the Member States bodies responsible for collecting and preserving information on Research and Technological Development activities at national level (https://op.europa.eu/s/oUUA, 6 May 1991).

- As stated in a 2014 study for the European Parliament (https://op.europa.eu/s/oUUE), at that date “19 out of the 28 Member States have developed or are in the course of developing national research information systems, close to all CERIF compliant [and] an additional 5 Member States are considering it”.
- The OpenAIRE and the EOSC share a common vision on building Open Science infrastructure (https://www.openaire.eu/a-common-vision-for-eosc-white-paper). The OpenAIRE Guidelines for CRIS managers are built as a CERIF-XML profile (https://openaire-guidelines-for-cris-managers.readthedocs.io/en/v1.1.1/introduction.htm).

As a conceptual model, CERIF is very flexible and its implementation into information systems is different from one system to another. An analysis was made of the OS-CAM to find out which criteria could be stored and retrieved with a system whose data model relies on the CERIF model. For each criterion, one or several implementation recommendations have been described: within which entities and properties a value should be stored and retrieved, if the value is one of a pre-defined list or free text, and when recommending the use of a list, is this list an existing vocabulary or one to create.

ORCID

The ORCID (Open Researchers and Contributor ID) denotes both an alphanumeric code to uniquely identify authors (ORCID iD) and contributors of scholarly communication, as well as the ORCID web site and services to look up authors and their bibliographic output. Moreover, ORCID is designed to be an infrastructure that organisations can then use to build tools on top of, or use data from ORCID in addition to other data sources as part of their assessment of researchers, etc. As an example, there is a relatively new platform called Rescognito: https://rescognito.com/ - which is based on ORCID and geared toward providing recognition to individuals for various roles they have played in open research, etc. ORCID is operated by a global, not-for-profit organisation of the same name sustained by fees from their member organisations.

In this study, we analysed how the OSCAM criteria could be stored and retrieved with a system whose data model relies on the ORCID researcher account. A researcher account and its available features were analysed through the ORCID GUI. Moreover, the available API and the following documentation sources have been analysed as well:

- ORCID record schema https://support.orcid.org/hc/en-us/categories/360000663114-Building-your-ORCID-record-connecting-your-iD
- Building your ORCID record & connecting your iD https://info.orcid.org/documentation/integration-guide/orcid-record/
- ORCID identifier types https://pub.orcid.org/v3.0/identifiers

Research.fi

Research Information Hub (RIH) is the national service that aggregates, links and distributes metadata related to the research activities and outputs in Finland. Data sources for RIH include for example university CRIS systems, data archives, and research funders. RIH does not include any facilities for data production aside from what can be inferred from the existing data. More information about the Research Information Hub can be found in RIH's wiki. Research.fi is a service provided by the ministry of Education and Culture, and the technical solution is developed by CSC - IT Center for Science, a company partly owned by the state of Finland (70%) and Finnish higher education institutions (30%), Research Information Hub is the so-called tip of the iceberg when it comes to producing and maintaining an up-to-date national view across all fields of science. Interoperability platform from Digital and Population Data Services Agency of Finland is another service that plays an essential role when it comes to putting open science "on the map" in RIH. CSC has adopted an interoperability platform as the tool for maintaining and publishing technical documentation related to RIH. This way the interoperability platform provides a public shared view to different stakeholders of RIH for both data producers and users.

Analysis was based on a valid version of documentation for Research Information Hub's data model and the reference data and terminology published via Interoperability Platform. The most important reference data from the OS-CAM analysis point of view is the "Activity and price types and roles" code list. It should be pointed out that analysis was done purely based on the published data model without taking into account the current state of the implementation of the Research Information Hub site.

NARCIS

NARCIS (National Academic Research and Collaborations Information System) is the main national portal in the Netherlands for those looking for information about Dutch researchers and their work. The NARCIS platform harvests scientific information from various institutional repositories and institutes in the Netherlands using an API. The information provided via NARCIS reflects how research information is collected and documented at these various institutes. As such, this case provides a glimpse of the open science landscape in the Netherlands, as viewed through the NARCIS portal.

NARCIS provides information about (open access) publications from the repositories of Dutch universities and research institutes, datasets from some data archives and descriptions of research projects, researchers and research institutes. NARCIS is a core service of the Dutch national centre of expertise and repository for research data (DANS), which is an institute of KNAW (Royal Netherlands Academy of Arts and Sciences) and NWO (Dutch Research Council).

Although NARCIS **cannot** be used as an entry point to access complete overviews of publications of researchers in the Netherlands, there are more and more institutions that make all their scientific publications accessible via NARCIS.

The Dutch research institutions are planning to create an Open Knowledge Base (OKB), which in future could extend or supersede NARCIS.

In this study, we analysed information (entities and properties) available through the GUI of the NARCIS search platform. Moreover, available documentation at the web site (https://www.narcis.nl/about/Language/en), as well as documentation provided by a contact person in DANS, were analysed too:

- Data model of NARCIS platform available in the MySQL Workbench model format (extension is mwb)
- Description of the platform and data model in the Dutch language

CRIStin

Norway has a national current research information system called CRIStin. It works as a centralised CRIS system for universities and research institutions. The main goals of CRIStin are the collection and distribution of information about Norwegian research, and support administrative and reporting tasks through data aggregation and reuse. In the future, development of CRIStin will be organised as part of the Nasjonalt vitenarkiv (NVA) services with a goal of providing an even more comprehensive overview of research in Norway. The CRIStin portal has well-documented APIs for read and write operations. CRIStin is owned by the Royal Ministry of Education and Research and is developed by Unit - The Norwegian Directorate of ICT and Joint Services in Higher Education and Research.

This analysis is based on information publicly available from the API documentation site and through the user interface of the portal itself. Most of the information was extracted from JSON schema documents. There are slight mismatches between schema documentation and the actual documents returned by API (e.g., https://api.cristin.no/v2/results/356583), so schema-based observations have been double checked with API requests of representative documents when necessary. Also, the older version of the CRIStin application was used for certain parts of the evaluation. Since the old CRIStin system does not provide a public portal, local domain experts assistance was required to gather information.

2. RESULTS

In this section are gathered the results of the evaluation of the 5 models, related to their coverage of the OS-CAM criteria, following the methodology outlined in the previous section.
Using a harmonised quantitative scale allows the creation of statistics and their graphical representation.

The OS-CAM divides criteria into six main open science activity categories: "research output", "research process", "service and leadership", "research impact", "teaching and supervision", and "professional experience". The main categories are further divided into open science activities, such as “publications”, “societal impact”, and “mentoring”, with one or more possible evaluation criteria. For example, “project management” activity has “Successfully delivering open science projects involving diverse research” as one of its possible evaluation criteria.

All five cases

Figure 1 shows the overall results of evaluation in main open science categories. We can see that the possible OS-CAM evaluation criteria related to research outputs are best covered in the five models. The results deteriorate as we move from established and popular ways of acquiring academic merit, such as research outputs, to more holistic ones that include for example teaching and professional experience related information. All in all, the coverage is quite low in all categories, with no or minimal support in about 50% or more of the evaluated criteria.

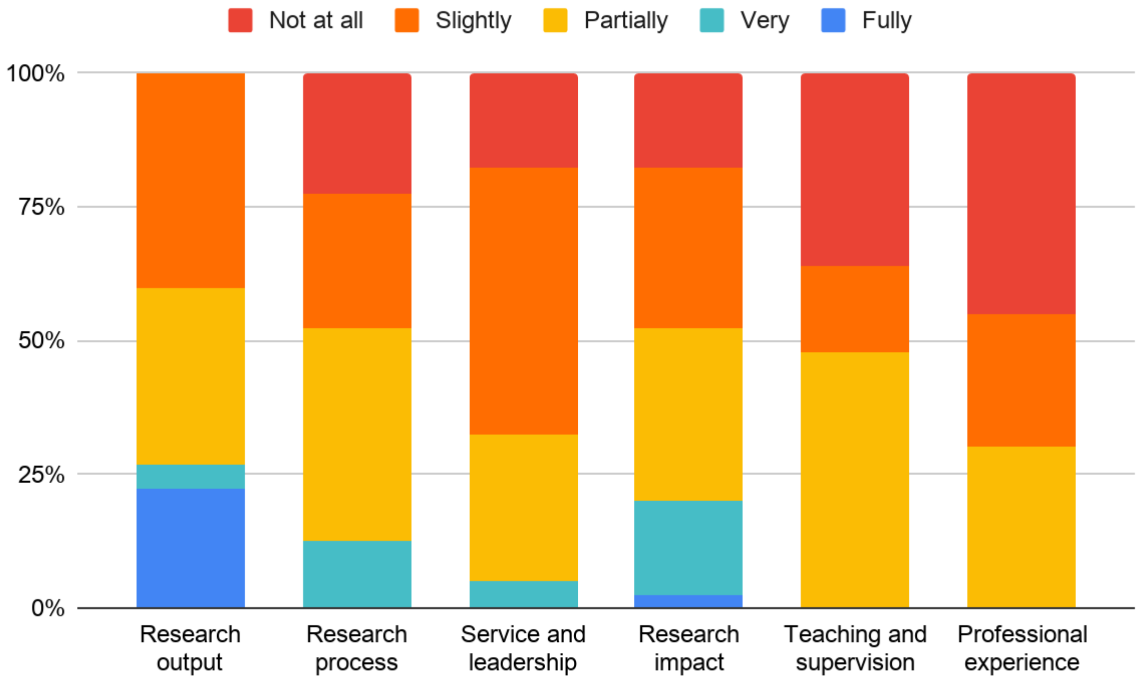


Figure 1: Coverage of the OS-CAM six main categories by analysed infrastructures.

Low coverage of the “Teaching and supervision” category highlights the fact that current systems focus on the research merits at the expense of educational merits. Both ways of accumulating merit can play an important part of many researchers' careers, but that is currently not reflected in the systems under evaluation.

According to our analysis of the five models, there is a great variation in the coverage of the 42 OS-CAM criteria (as ordered based on their weighted sum values calculated with the following weights: fully = 5, very = 3, partially = 1, slightly= 0.5 and not at all = 0). Figure 2 shows the results on the ten best covered OS-CAM criteria (the results on all OS-CAM criteria are provided in Appendix 2).

Unsurprisingly, publication related criteria (“self-archiving in open access repositories” and “publishing in open access journals”) are the ones with the best coverage. Open access has already made its way successfully into the current research information systems. Even through software development outputs are ranked third, they are still notably less well covered that publication outputs. Some evaluated systems for example only had an implicit way of handling software as part of research datasets. On the other hand, there were also cases where software was handled explicitly by the data model.

When evaluating “Making use of open data from other researchers” criterion, we required that the system/model could handle datasets explicitly with a possibility of classifying them as being both openly available and usable with an open license. In addition to that, it must be possible to create relationships between datasets and other research outputs with explicit “uses” semantics. Most of the systems failed one or more of these requirements. For example, it might be only possible to create the “ispartof” type of links between datasets.

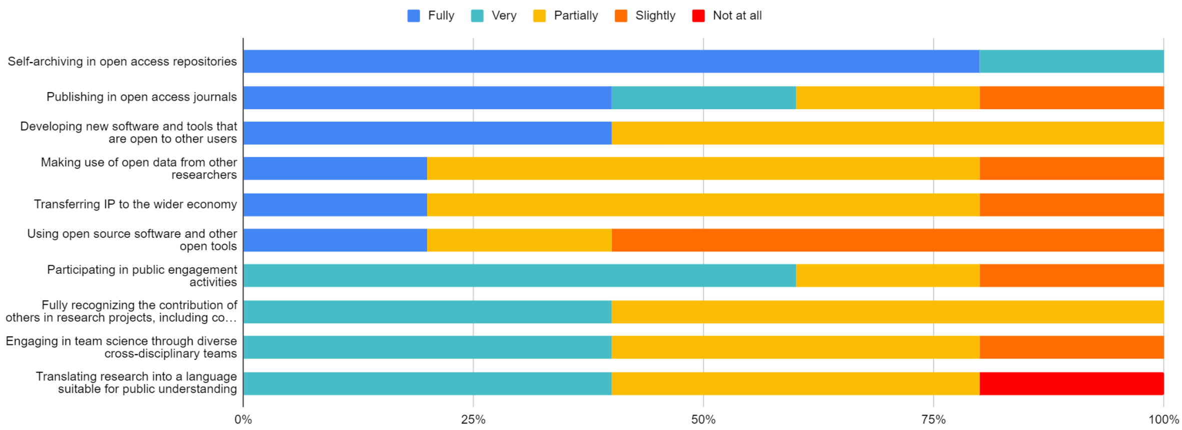


Figure 2: Top 10 best covered OS-CAM criteria according to this study.

Evaluation of “Fully recognizing the contributions of others...” and “Engaging in team science through diverse cross-disciplinary teams” relied heavily on the possibilities of linking people indirectly through labelled links between different outputs. It should be possible to add both internal and external actors, both persons and organizations, as contributors with appropriate roles to any type of research outputs. In order to do this in a flexible way, many models had opted for modelling participation link as a separate entity, which allows for adding metadata about the relationship such as role, date and affiliation. Even with a flexible data model in place, none of the evaluated systems offered vocabularies comprehensive enough to warrant for full support of the OS-CAM criteria.

The following sections provide more detailed information about the findings and analysis for each system and data model under evaluation.

CERIF

As is summarised in the following chart (Figure 3), out of 42 OS-CAM criteria, none is “Not at all” covered.

Six criteria are fully covered and seven are almost fully supported (Very), meaning only an additional check from an administrative person (not an academic assessment expert) is needed to assess whether a researcher fulfils the OS-CAM criterion.

The remaining 29 criteria have been classified as Partially (most of them) or Slightly (7 of them), meaning it is possible to assess the fulfilment of those criteria using information stored in a CERIF-based system, but enrichment of data from other sources (which links are stored in the CERIF-based system) is needed for further processing and assessment by evaluators.

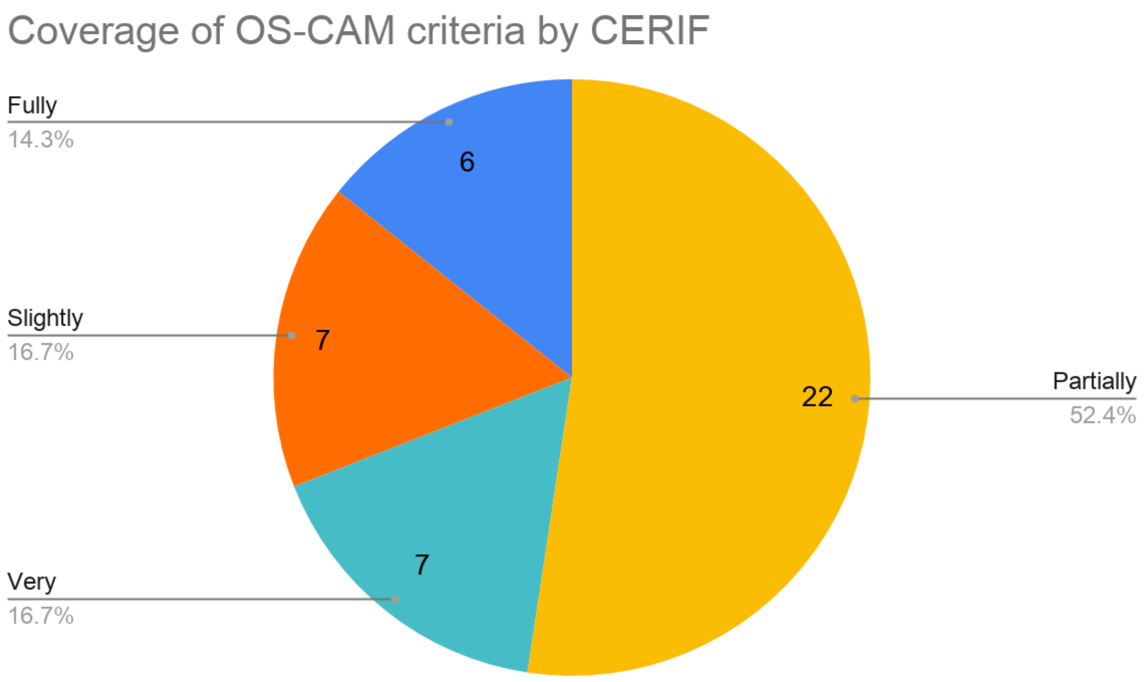


Figure 3: Level of coverage OS-CAM criteria by CERIF.

Overall, CERIF has the ability, as a model, to store or provide links to all the information needed by the OS-CAM for a researcher evaluation.

The ease of evaluation will however depend on the way this information is entered or linked, and this depends in turn on the way the system implements the data model: partly depending on the database model (tables and fields used to store the entities and properties), partly depending on the expected data format (free text or selection out of a pre-defined vocabulary), and partly depending on the UI (required and non-required fields).

The assessment of a researcher's career with the OS-CAM criteria requires to:

- MAP the researcher's career [description of a person skills and experience, related professional information], research outputs and activities [events in the research life cycle] and relations to others [through affiliations, memberships, teaching, mentoring...]
- MEASURE, in general, a level of FAIRness or quality of Open Science promotion (such an approach to define a measure of FAIRness is described, for example, in the article *Evaluating FAIR maturity through a scalable, automated, community-governed framework*[162])

For this, CERIF has dedicated entities (*person, org unit, publication, product, patent*, etc) and the ability to store relations between them, these relations being associated with roles or classification terms.

For these terms to be relevant in an OS-CAM based evaluation, there is a need for vocabularies:

- Terms describing Open Science: open access, open data, opens science projects, open peer review, open source in open science, altmetrics, open access policies, open data policies ...
  - For example, the FOSTER taxonomy: https://www.fosteropenscience.eu/foster-taxonomy/open-science-definition
- Terms describing the relationship between publications, data, etc to trace provenance : built on...
  - For example, the CERIF Inter-Output Relations vocabulary: https://cerif.eurocris.org/vocab/html/InterOutputRelations.html
- Terms describing participation of persons or org units in open innovations, collaborative projects,...: member, reviewer, group leader, contributor, reviewer, participant, stakeholder...
  - For example, the CERIF Person-Organisation Roles vocabulary: https://w3id.org/cerif/vocab/PersonOrganisationRoles
  - For example, the CERIF Organisation Project Engagements vocabulary:

[162] Evaluating FAIR maturity through a scalable, automated, community-governed framework https://www.nature.com/articles/s41597-019-0184-5

https://cerif.eurocris.org/vocab/html/OrganisationProjectEngagements.html

- Terms describing the diversity of org unit participating in open innovation, collaborative projects,...:
  - For example, adding a term "Citizen group" to the CERIF Organisation Types vocabulary: https://w3id.org/cerif/vocab/OrganisationTypes that already includes "Company", "SME, "Research Institute",...

Additionally, to the use of the previously mentioned entities, we can also make the parallel between the use of CERIF to define metrics related to a researcher profile, and the use of CERIF to define the Snowball metrics[163] used to measure research activities-related criteria (for example, the societal impact), and similarly make use of the *indicator* and *measurement* CERIF entities (see Appendix 3 on Snowball metrics).

ORCID

Out of 42 OS-CAM criteria, only 1 has been classified as fully supported by ORCID (Figure 4). Moreover, there is also 1 criterion almost fully supported (Very), meaning only an additional check from an administrative person (not an academic assessment expert) is needed to assess whether a researcher fulfils the OS-CAM criterion. On the other end, there are 7 OS-CAM criteria not at all supported by ORCID, meaning that evaluators can't use information stored in the ORCID record of a researcher to assess whether or not the researcher fulfils a criterion. The rest of 33 (78.6 %) criteria have been classified as Partially or Slightly, meaning it is possible to assess fulfilment of those criteria using information stored in the ORCID record, but enrichment of data from linked sources (via URLs in the ORCID record) is needed for further processing and assessment by evaluators.

---

[163] Snowball Metrics Recipe Book : https://www.elsevier.com/__data/assets/pdf_file/0006/53169/Snowball_Metrics_Recipe_Book.pdf

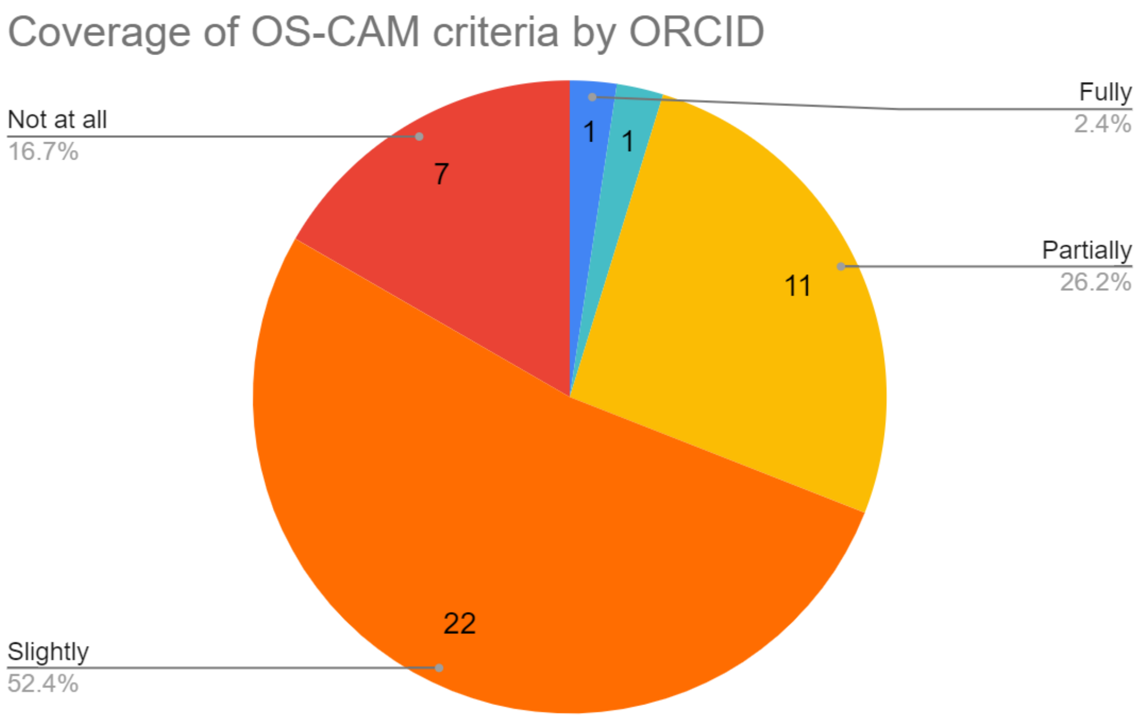


Figure 4: Level of coverage OS-CAM criteria by ORCID.

Although ORCID's web site and services have not been developed for the purpose of researchers' assessment, it could be even partially used for that purpose. However, there are the following major issues with assessment of Open Science outputs, activities and expertise:

- There is no flag for an output (Work) whether it is published under some open access license. There is a section for identifiers which include the fields: Identifier type, Identifier value, Identifier URL and Relationship. However, fields Identifier type and Relationship could have only values for predefined list of values (vocabularies). The list of Identifier types is available at https://pub.orcid.org/v2.0/identifiers, while Relationship can have only three values: Self, Part of, Version of. If the set of Identifier types would be extended with some licence type (for instance, Creative commons), and a Relationship type set with "has", it might be possible to assign license for any output (Work) and to assess whether some result is open-access or not.
- Publication or dataset can be linked with other results' identifiers including "Software", but there are only three relationships' types: self, part of, version of. Besides extension of Relationship type set with the term “has” (see previous point), there might be other terms representing relationship between outputs and activities (“use”, “contain”, “extend”, etc.)

- Funding can be catalogued in ORCID, but not the project (activity). Therefore, there is no list of participants and their roles in projects, meaning a part of the collaboration network is missing in an ORCID record and different roles in projects are not recorded.
- Work categories and types don't cover all Open Science outputs (e.g. strategies, visions).
- There is a limited set of roles/authorship types for a Work (output), and can be only specified through API (not available through GUI). However, adding the ability to enter co-authors and other collaborators when adding works to the registry manually is on ORCID longer term plans. Moreover, ORCID is planning to incorporate CRediT implementation for works on their 2021 roadmap: https://docs.google.com/document/d/17VVDljVZBkt0VQ0fkGY39mFbDFdB8HnT4xE1zT0O6rY/edit?ts=60479a74#.

Research.fi

Out of the 42 OS-CAM criteria under evaluation, five were assessed as being either fully or very supported. On the other hand, seven criteria were evaluated as not being supported at all. Support was lacking mainly in the "Teaching and supervision" and "Professional experience" main categories. Most of the categories were either slightly or partially supported, with these two levels comprised of 12 and 18 categories respectively. "Research output" and "Research impact" main categories were the only two categories with at least slight support for all criteria.

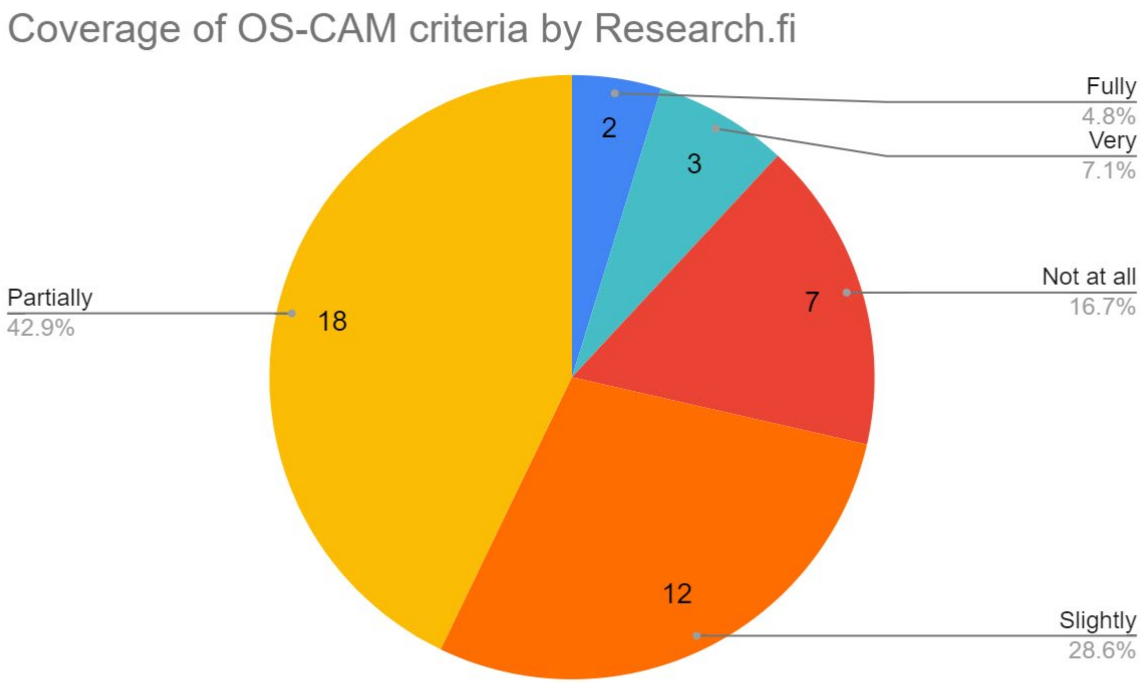


Figure 5: Level of coverage OS-CAM criteria by Research.fi

RIH's data model reflects the requirements stemming from RIH specific use cases as well as data models of the systems from which it aggregates. However, adoption and successful application of this Interoperability platform and especially the mindset that comes with it, has pushed the RIH's data model towards a more flexible and general direction and this work continues on terminology, vocabulary and modelling levels.

The following points are the summary of observations from RIH's data model from the perspective of OS-CAM:

- Model is missing patents as a research output.
- Software outputs are not represented as a class of its own, but are currently implemented as Research data of a certain type. This is due to practical reasons, as the data source for research dataset Metax contains also software "datasets".
- Activity and prize types and roles code list is created from the perspective of research merit. There is however a place for generic education related information, which can be "classified" with related competencies.
- All of the research outputs have a reference to keywords that can be identified with URIs and part of authoritative ontologies. This is good from OS-CAM's perspective, as they can be used to make unambiguous and

robust links to OS concepts from shared vocabularies, provided that those vocabularies exist of course.

- Research community and Role in the research community form a good basis for OS network descriptions. However, since many of these communities do not have persistent identifiers, hence linking across datasets is difficult.
- The easiest way to add more support for OS-CAM is to work on code lists and keyword vocabularies. For example KOKO, collection of Finnish core ontologies, currently contains concepts for “open access” and “open source code” but not for “open science”.

NARCIS

Out of 42 OS-CAM criteria, only one has been classified as fully supported by NARCIS (Figure 6). Moreover, there are also two criteria almost fully supported, meaning only additional check of an administrative person (not academic assessment expert) is needed to assess whether a researcher fulfils the OS-CAM criteria, and 11 criteria partially covered, meaning additional assessment by evaluator is needed. On the other side, there are 15 OS-CAM criteria not at all supported by NARCIS, meaning evaluators can’t use information available through the NARCIS platform to assess whether or not the researcher fulfils a criterion. The rest of 13 criteria have been classified as Slightly covered, meaning it is possible to assess fulfilment of those criteria using information available through the NARCIS platform, but enrichment of data from linked sources (via URLs) is needed for further processing and assessment by evaluators.

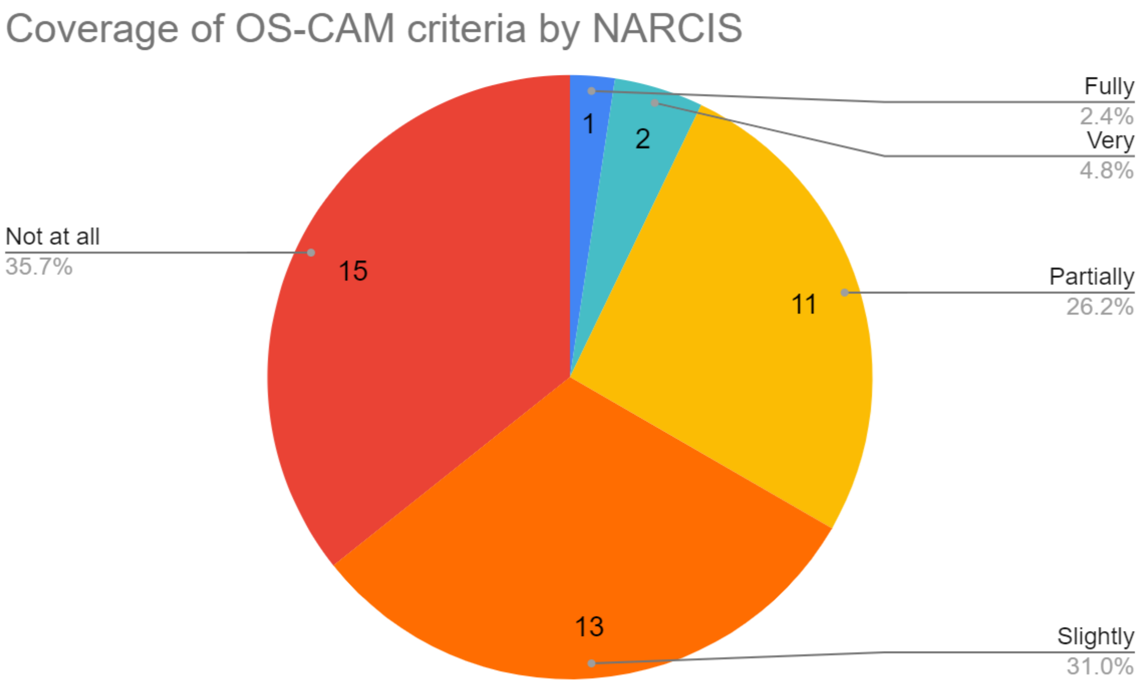


Figure 6: Level of coverage OS-CAM criteria by NARCIS.

Two thirds (28 of 42) of OS CAM criteria are only slightly or not at all covered by information available through the NARCIS portal.

Although the NARCIS portal has not been built for the purpose of researchers' assessment, it could be used for that purpose, but with significant limitations due to the following major issues with assessment of Open Science outputs, activities and expertise:

- Outputs - Metadata about outputs have been collected in numerous metadata formats (Dublin Core, METS, DCAT, etc.) with different levels of richness. This is likely due to the fact that NARCIS harvests information provided in local CRIS systems; the information provided in NARCIS is dependent on the information recorded at those institutions. There is a quite rich set of publication types, but it could still be extended with some types such as data/software paper - there is type software, dataset, article, but there is no type data paper, neither software paper. There is information about accessibility of output, meaning someone can find information for an output whether it is published under some open access license. However, there are only four types of accessibility, meaning it is not possible to distinguish all 6 types of creative commons licences. Outputs are linked with instances of researchers (people) preserved in the NARCIS database, but there are at least two issues with this linking. The first one is the fact that this link is not always well established with all contributors, and the second one is that the set of contributor types depend on the metadata format of harvested outputs' metadata records. Moreover, links between outputs (two different publications, dataset and publication, software and publication, etc) are also specific for metadata format and there are no standardised classifications of those links, meaning it is not unique throughout the complete collection.

- Activities - The NARCIS portal is much more output-centric than activity-centric, meaning there is no information about contribution to the organisation of the conference, peer reviewing, etc. Moreover, membership in international organizations can't be presented due to the limited set of types of relations between persons and organizations units (there is no "Member" type of relation). Although projects can be linked to the people, there are only three types of relations - Project leader, Researcher, Contact person. Moreover, the majority of projects have only information about the project leader. However, activity of theses' supervising and co-supervising can be recorded as well.
- Expertise - It is possible to record a researcher's expertise in the free-form of a set of expertises' terms. Moreover, prizes granted to a researcher are linked to the researcher entity, but there is no rich structure of the prizes' information. Furthermore, skills and qualifications obtained through some short courses can't be recorded at all.

CRIStin

Out of 42 evaluated OS-CAM criteria, only one has been classified as being fully supported. Three more criteria are evaluated as being "very"/well supported. All top evaluated criteria are either in research output or research impact categories. Criteria evaluated at "fully" or "very" can be used for OS assessment with minimal manual work from an administrative person. 12 of the OS-CAM criteria were labelled as being not supported at all. Categories with least support are focused on "Professional experience" categories where 75% or the criteria are unsupported. Whereas in the "Research output" category, all criteria are at least slightly supported. Slightly supported is the most common evaluation level with 15 criteria. For example, criteria in "Service and leadership" and "Research output" main categories are mainly (~63% and 67% respectively) only slightly supported. Partial support is evaluated to cover all in all 11 criteria and these evaluations are distributed quite evenly across the main categories, exception being the "Research process" category with 50% of the criteria being partially supported.

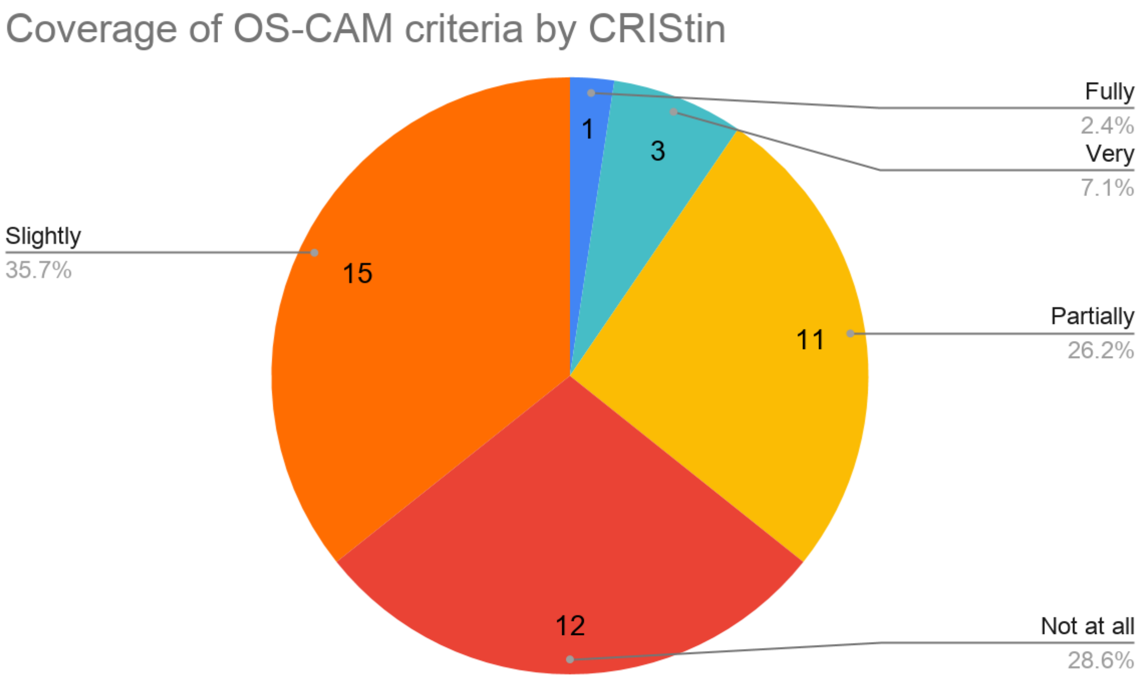


Figure 7: Level of coverage OS-CAM criteria by CRIStin.

CRIStin's data model does not include a concept for activity. Instead, similar outputs are modelled as contributions qualified with a role to a certain result. This makes records of research activities implicit in a way and puts the focus more on end results. For example, one can't use the portal to search for all persons that have supervised a doctoral thesis.

- Lack of explicit Activity type and the fact that contributions can only be linked to results makes it difficult to add more specific information to certain types of outputs. Sometimes it might be good to be able to make visible something that might not produce clear result artifacts.
- Having an "umbrella" type for basically all outputs both scientific and other makes the model straightforward to use.
- Restricting use of organization/institution related information only to ones known by the system can make it harder to input more niche data.
- CRIStin has a comprehensive set of result categories that cover many disciplines.
- Qualified contributions make for a very flexible model for connecting people to varied results, but there seems to be no publicly available

information about the possible restrictions on the contribution role's values.

3. CONCLUSION AND RECOMMENDATIONS

The OS-CAM criteria should be described in much more detail joined with a set of examples about the evidence of fulfilling those criteria. Some examples of OS-CAM criteria which might be understood in multiple ways are:

- Fully recognizing the contribution of others in research projects, including collaborators, co-authors, citizens, open data providers
  - Is it only in research projects, or any research in general?
  - Is it analysis of a researcher's practice to cite and acknowledge used resources appropriately in a publication?
- Translating research into a language suitable for public understanding
  - Does this mean translating a research publication from one language to another, or also transformation of a research result into a newspaper article or blog, which might be easy for understanding by citizens?
- Evidence of use of research by societal groups
  - Probably here should be mentioned altmetrics. All altmetrics or some of them? What about innovations and spin-of companies run by a researchers' group?
- Mentoring and encouraging others in developing their open science / Supporting early-stage researchers to adopt an open science approach
  - These are two quite similar criteria. Explanation of the distinction of those two criteria is needed.

One way to clarify the OS-CAM criteria could be similar as for the Snowball metrics (see the appendix 3).

Although analysed data models and systems have not been exclusively built for the purpose of assessment, they could also be used for that purpose. The analysis presented in this document shows that all systems currently have a limited coverage of the OS-CAM criteria. It confirms previously identified gaps in this project (EOSC Co-creation project - Study Proposal #16: European overview of career merit systems)

- Guidelines, policies and infrastructures for supporting open access publications are much better developed/adopted than for open data and other research objects.

- Although the importance of accounting for other types of research output and activities in researchers' assessment has been recognized, it has not yet been fully adopted in practice.

Although part of the OS-CAM criteria requests expert based analysis of qualitative aspects of researchers' achievements, the infrastructures can help in the assessment process and quite well semi-automatize assessment based on the OS-CAM criteria.
The list of six recommendations for good coverage of the OS-CAM criteria by an infrastructure:

1. License attached to any research output, as well as for publication channel should be catalogued. Moreover, mapping of licenses to the accessibility tag should be defined (open, close, restricted, embargo, etc.)
2. A rich hierarchy of
    a. outputs' types should be supported with possibility to extend with new types in the future. Those types should include traditional and modern scholarly communications - article, software, dataset, blog, post, newspapers, TV interview, etc.
    b. activities' types should be supported with possibility to extend with new types in the future. Those types should include traditional and modern activities - reviewing, participation in a project, teaching activities including supervising, chairing, organizing an event, social network activities, membership in international organizations, etc.
    c. achievements' types should be supported with possibility to extend with new types in the future. Those types should include traditional and modern scholarly achievements - competence, award, certificate, digital reward, metrics including altmetrics, social networks recommendations, etc.
3. Besides assigning type to any output, activity or achievement, the possibility of assigning additional classification representing subjects/keywords including Open science related terms would be very useful.
4. Establishing links based on PIDs between research outputs should be enabled. There should be an option to define the context of the link (type of the link - uses, extends, is based on; description of the link).
5. Establishing links based on PIDs between research outputs and projects on one side and actors on the other side (people and institutions) should

be enabled. These links should be modelled in a way that allows for metadata to be added to the link itself, meaning there should be an option to define the context of the link (type of the link representing role - author, supervisor, coordinator, creator, reviewer, etc; description of the link). Classification of the links can be based on some well known vocabularies (for instance https://casrai.org/credit/)

6. Altmetrics should be harvested and represented in the data model or integrated with a platform through altmetrics' web widgets such as https://plumanalytics.com/integrate/embed-metrics/ or https://www.altmetric.com/products/altmetric-badges/

Appendix 1 - The OS-CAM criteria

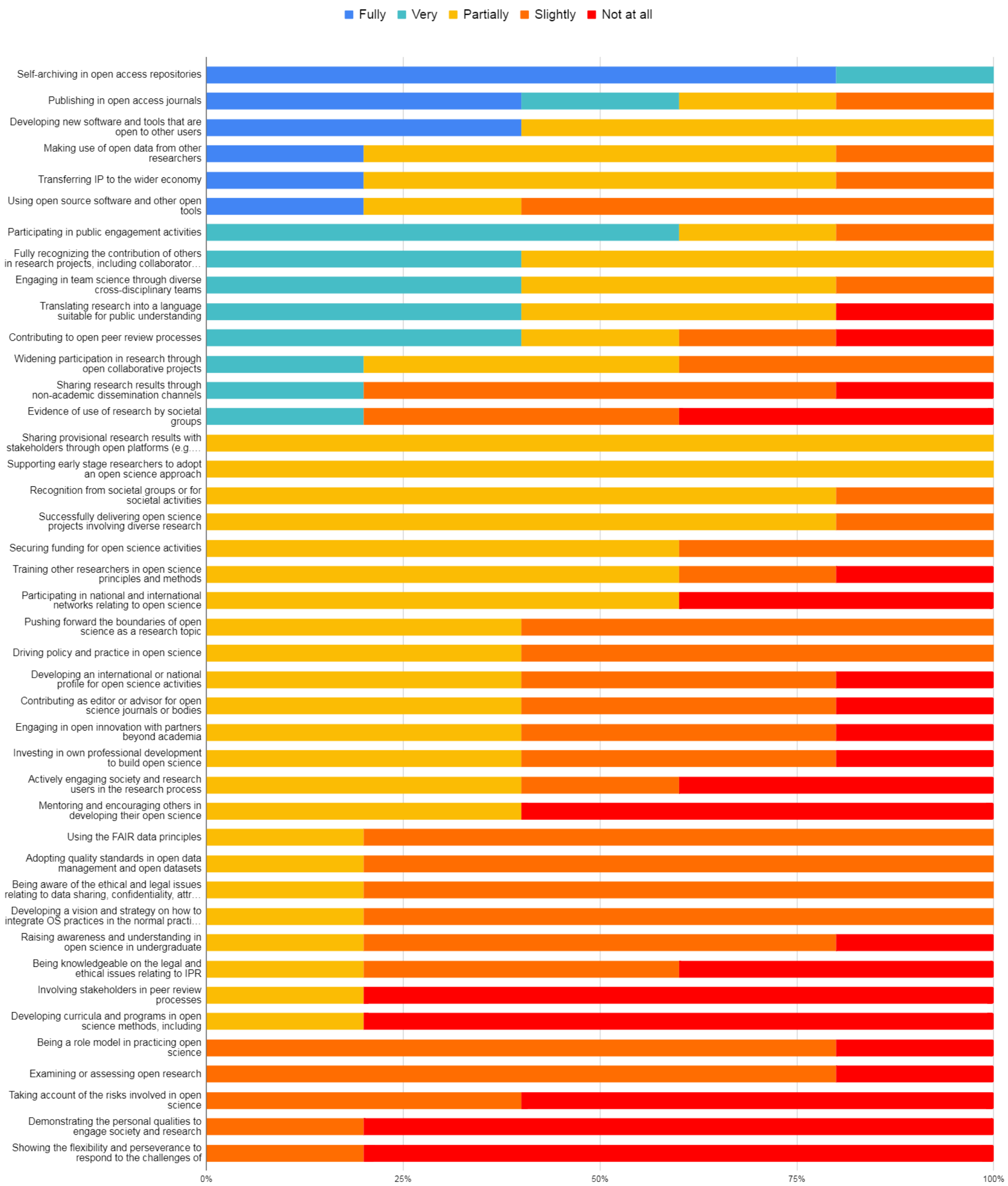
OS-CAM (Open Science Career Assessment Matrix)		
Category of Open science activity	**Open science activity**	**Evaluation criteria**
RESEARCH OUTPUT	Research activity	Pushing forward the boundaries of open science as a research topic
	Publications	Publishing in open access journals
		Self-archiving in open access repositories
	Datasets and research results	Using the FAIR data principles
		Adopting quality standards in open data management and open datasets
		Making use of open data from other researchers
	Open source	Using open source software and other open tools
		Developing new software and tools that are open to other users
	Funding	Securing funding for open science activities
RESEARCH PROCESS	Stakeholder engagement / citizen science	Actively engaging society and research users in the research process
		Sharing provisional research results with stakeholders through open platforms (e.g. Arxiv, Figshare)
		Involving stakeholders in peer review processes
	Collaboration and	Widening participation in research

	Interdisciplinarity	through open collaborative projects
		Engaging in team science through diverse cross-disciplinary teams
	Research integrity	Being aware of the ethical and legal issues relating to data sharing, confidentiality, attribution and environmental impact of open science activities
		Fully recognizing the contribution of others in research projects, including collaborators, co-authors, citizens, open data providers
	Risk management	Taking account of the risks involved in open science
SERVICE AND LEADERSHIP	Leadership	Developing a vision and strategy on how to integrate OS practices in the normal practice of doing research
		Driving policy and practice in open science
		Being a role model in practicing open science
	Academic standing	Developing an international or national profile for open science activities
		Contributing as editor or advisor for open science journals or bodies
	Peer review	Contributing to open peer review processes
		Examining or assessing open research
	Networking	Participating in national and international networks relating to open science
RESEARCH	Communication and	Participating in public engagement

IMPACT	Dissemination	activities
		Sharing research results through non-academic dissemination channels
		Translating research into a language suitable for public understanding
	IP (patents, licenses)	Being knowledgeable on the legal and ethical issues relating to IPR
		Transferring IP to the wider economy
	Societal impact	Evidence of use of research by societal groups
		Recognition from societal groups or for societal activities
	Knowledge exchange	Engaging in open innovation with partners beyond academia
TEACHING AND SUPERVISION	Teaching	Training other researchers in open science principles and methods
		Developing curricula and programs in open science methods, including open science data management
		Raising awareness and understanding in open science in undergraduate and masters' programs
	Mentoring	Mentoring and encouraging others in developing their open science capabilities
	Supervision	Supporting early stage researchers to adopt an open science approach
PROFESSIONAL EXPERIENCE	Continuing professional development	Investing in own professional development to build open science capabilities

	Project management	Successfully delivering open science projects involving diverse research teams
	Personal qualities	Demonstrating the personal qualities to engage society and research users with open science
		Showing the flexibility and perseverance to respond to the challenges of conducting open science

Appendix 2 - Coverage of the OS-CAM criteria by analysed infrastructures



Appendix 3 - Snowball metrics

The *Indicator* and *Measurement* CERIF entities were used to express Snowball metrics with CERIF ; something similar could then be used for the OS-CAM criteria similar to the highlighted ones.

Figure 2: Snowball Metrics recipes

	Research Inputs	Research Processes	Research Outputs and Outcomes
Research	• Applications Volume • Awards Volume • Success Rate	• Income Volume • Market Share	**Publications & citations** • Scholarly Output (enhanced) • Citation Count • Citations per Output • *h*-index • Field-Weighted Citation Impact • Outputs in Top Percentiles • Publications in Top Journal Percentiles **Collaboration** • Collaboration • Collaboration Publication Share • Collaboration Impact • Collaboration Field-Weighted Citation Impact • Academic-Corporate Collaboration • Academic-Corporate Collaboration Impact **Societal impact** • Altmetrics • Public Engagement • Academic Recognition
Enterprise Activities/ Economic Development	• Academic-Industry Leverage • Business Consultancy Activities	• Contract Research Volume	• Intellectual Property Volume • Intellectual Property Income • Sustainable Spin-Offs (enhanced) • Spin-Off-Related Finances (enhanced)
Post-Graduate Education	• Research Student Funding	• Research Student to Academic Staff Ratio	• Time to Award of Doctoral Degree • Destination of Research Student Leavers

- Snowball Metrics shared in original Recipe Book, November 2012
- Snowball Metrics shared in edition 2 of the Recipe Book, June 2014
- Snowball Metrics shared in this edition of the Recipe Book, November 2017

Examples of metrics in Snowball to measure Collaboration, Societal impact and IP:

Collaboration Impact	Citations per any collaborative output
	Citations per internationally collaborative output
	Citations per nationally collaborative output
Academic-Corporate Collaboration	Number of academic-corporate collaborative outputs
	Percentage of academic-corporate collaborative outputs
	Number of academic-corporate collaborative outputs per FTE
Academic-Corporate Collaboration Impact	Citations per academic-corporate collaborative output
Altmetrics	Scholarly Activity Count
	Scholarly Commentary Count
	Social Activity Count
	Mass Media Count
	Scholarly Activity Count per FTE
	Scholarly Activity Count per output
	Scholarly Commentary Count per FTE
	Scholarly Commentary Count per output
	Social Activity Count per FTE
	Social Activity Count per output
	Mass Media Count per FTE
	Mass Media Count per output
Public Engagement	Number of attendees
Intellectual Property Value	Number of patents filed
	Number of patents granted
	Number of active patents
	Number of licenses granted
	Number of patents filed per FTE
	Number of patents granted per FTE
	Number of active patents per FTE
	Number of licenses granted per FTE
Intellectual Property Income	Income
Sustainable Spin-Offs	Number of sustainable spin-offs

Some of those metrics were used as possible examples to understand some OS-CAM criteria.

ANNEX 5

BOOTCAMP METHODOLOGY AND OVERVIEW

ANNEX 5

BOOTCAMP METHODOLOGY AND OVERVIEW

The original project proposal was based on two face-to-face bootcamps with 20 participants from diverse backgrounds. Due to the Covid-19 restrictions, the project plan was adjusted. Instead of 2 longer face-to-face bootcamps, four half-day online bootcamps were organised.

The bootcamps were clustered:

- Bootcamps 1 and 2 focused on building FAIReR assessment vision.
- Bootcamps 3 and 4 focused on building a roadmap from vision to implementation.

Bootcamps were held via Zoom and used the Mural online collaboration platform. The bootcamps were facilitated by two professional facilitators Jani Turku and Kirsi Kaunissaari.

Bootcamp participants were recruited through the help of project steering committee networks, social media (including EOSC secretariat) as well as all other networks project team had.

Each bootcamp had preliminary material sent to the participants:

- Bootcamp 1 - vision draft for policy
- Bootcamp 2 - vision draft for infrastructure
- Bootcamp 3 - roadmap for policy development
- Bootcamp 4 - roadmap for infrastructure development

Bootcamps followed a structure to deliver an output for the project to support each of the vision and infrastructure elements. Each bootcamp produced large amounts of qualitative data in the Mural platform which was then utilised by the project team to develop project deliverables. Bootcamps also provided an opportunity for creating a community around developing FAIReR assessments in

Europe. Bootcamp participants had an opportunity to sign up to follow project developments and will receive a copy of the final report.

	Number of participants	Representing number of countries
Bootcamp 1 23.11.2020	18	9
Bootcamp 2 30.11.2020	18	9
Bootcamp 3 11.2.2021	35	13
Bootcamp 4 17.2.2021	27	11

ANNEX 6

PUBLIC CONSULTATION FOR VISION AND ROADMAP

ANNEX 6

PUBLIC CONSULTATION FOR VISION AND ROADMAP

Method and recruitment

We opened a draft of the vision and roadmap for public consultation for three weeks, from February 25 until March 15, 2021.

A PDF version of the draft was posted on the project website[164], along with a link to an online commentary form, designed using Lime survey. The commentary form was split into sections, corresponding to the structure of the vision/roadmap document. Respondents were asked to provide comments for each section individually; a final question solicited feedback about the document as a whole. It was also possible to comment either as an individual or on the behalf of an organisation.

Stakeholders and experts, drawn from the list of bootcamp participants and the professional networks of both project team members and the project's steering group, were invited to comment on the draft. The consultation was also widely advertised on social media.

Respondents and results

There was a high level of engagement with the consultation form (109 individuals accessed the form); we received five complete responses during the consultation period. Countries presented in the responses were the Netherlands (2), USA (1), Finland (1) and UK (1). Two of the responses were on behalf of a larger group; one respondent commented on behalf of a scientific publisher and another on behalf of a group of research support professionals.

Many responses clustered around the need to clarify language and terminology, e.g. the term "responsible" and "community-owned infrastructures."

---

[164] https://avointiede.fi/en/networks/eosc-co-creation

Respondents also encouraged recognizing the role of other stakeholders, e.g., publishers and commercial providers of metrics.

The importance of integrating qualitative assessments into the vision for the technical infrastructure was also mentioned. One respondent suggested accounting for privacy and security concerns in the development of the technical infrastructure and pointed out the development of other related systems and projects.

ANNEX 7

VISION OF THE FAIReR ASSESSMENT eINFRASTRUCTURE

ANNEX 7

VISION OF THE FAIReR ASSESSMENT eINFRASTRUCTURE

1. INTRODUCTION

The wide adoption of the Open Science paradigm as a common practice requests cultural changes, development of eInfrastructure, changes of policies and academic assessment practices. This document represents a vision of eInfrastructure which supports responsible academic assessment including assessment of Open Science outputs, activities and expertises. A research eInfrastructure and assessment policies should be developed taking care about cost and efficiency of the assessment process, and decreasing applicants' and evaluators' efforts.

The eInfrastructure should rely on meta-data linking every piece of scientific knowledge to a unique and persistent identifier that can become the basis for open, publicly available academic assessment data infrastructure [1]. By 2021, there are well-known persistent identifiers for some academic entity types, but unfortunately not for all. Anyway, the situation should be changed until 2025 taking into account a lot of initiatives and projects for PIDs (Freya, PID forum, PIDApalooza, etc.). It is expected that ORCID will be adopted as unique identifiers for researchers in the eInfrastructure, and an ORCID identifier will be mandatory for all applicants and participants in European funding programs [1]. Therefore, the eInfrastructure should be based on interoperable, linked, transparent systems to support robust evaluations which can be based on persistent identifiers assigned to each research ecosystem entity to enable tracking indicators for research outputs. The academic assessment indicators can reach their full potential only in such an eInfrastructure, i.e. only if they are underpinned with an open and interoperable academic data infrastructure [2]. The eInfrastructure should be open for an easy extension with new research object types, indicators, data providers. There are initiatives for building registers for discovering available data, publications, infrastructures and services, guidelines and policies, but there is no initiative for building a register of human and machine readable descriptions of indicators and (alt)metrics (gap 12 - Annex 2). Standardized semantic models for machine-readable representation of academic assessment criteria and rule based engines for automatic calculation of meeting those criteria should be built (gap 13 - Annex 2). This register should be developed in the next five years to make indicators easy to discover and reuse. Moreover, on top of this register the best practices for building indicators should be recognized and shared.

Furthermore, part of this eInfrastructure should enable automatization of collecting and formatting data for academic assessment. This will decrease applicants' efforts and simplify the process of collecting data. Moreover, the infrastructure could introduce standardization of type and format of data used for academic assessment, meaning could introduce standardized CV/Portfolio templates used in the academic assessment. Unfortunately, tools for making CVs/portfolios don't fully support diversity of research objects' types and scholarly communication channels at the moment (gap 9 - Annex 2), but it has to be changed by 2025 taking into account increased awareness of the importance of open science for further development of science and society.

At the end, the eInfrastructure for performing responsible academic assessment should be developed to decrease evaluators' efforts, and to make the assessment transparent and impartial. Complete insights into a researcher's achievements could be used as a supplement material by evaluators in qualitative assessment of researcher's career.

eINFRASTRUCTURE ARCHITECTURE FOR RESPONSIBLE ACADEMIC ASSESSMENT SUPPORTING OPEN SCIENCE PARADIGM

Figure 1 presents a vision for developing eInfrastructures which might make the academic assessment process responsible. The envisioned architecture is a result of

- overviewing available resources - platforms, initiatives, strategies, articles, etc. (see Annex 2 for more details),
- analysis of the survey conducted within this project (see Annex 3),
- five case studies of alignment data models and platforms across the Europe with the OS-CAM [3] criteria for assessment of Open Science outputs, activities and expertises (see Annex 4), and

- topics discussed at four bootcamps organized within the project.

Yellow rectangles are used for platforms/services which can be used for building research eInfrastructure ecosystem. Those platforms/services already exist across the world, but should be maintained and further extended in accordance with the Open Science paradigm. Orange cloud in the middle of the diagram represents integration of all those services and platforms under one umbrella. It will improve visibility/discoverability of platforms/services, and on the other side it will enable collecting complete achievement of a researcher or group. Moreover, three more orange rectangles should be developed as a part of this vision with the goal of making a basis for building local Academic assessment platforms (blue rectangles) which will support responsible academic assessment.

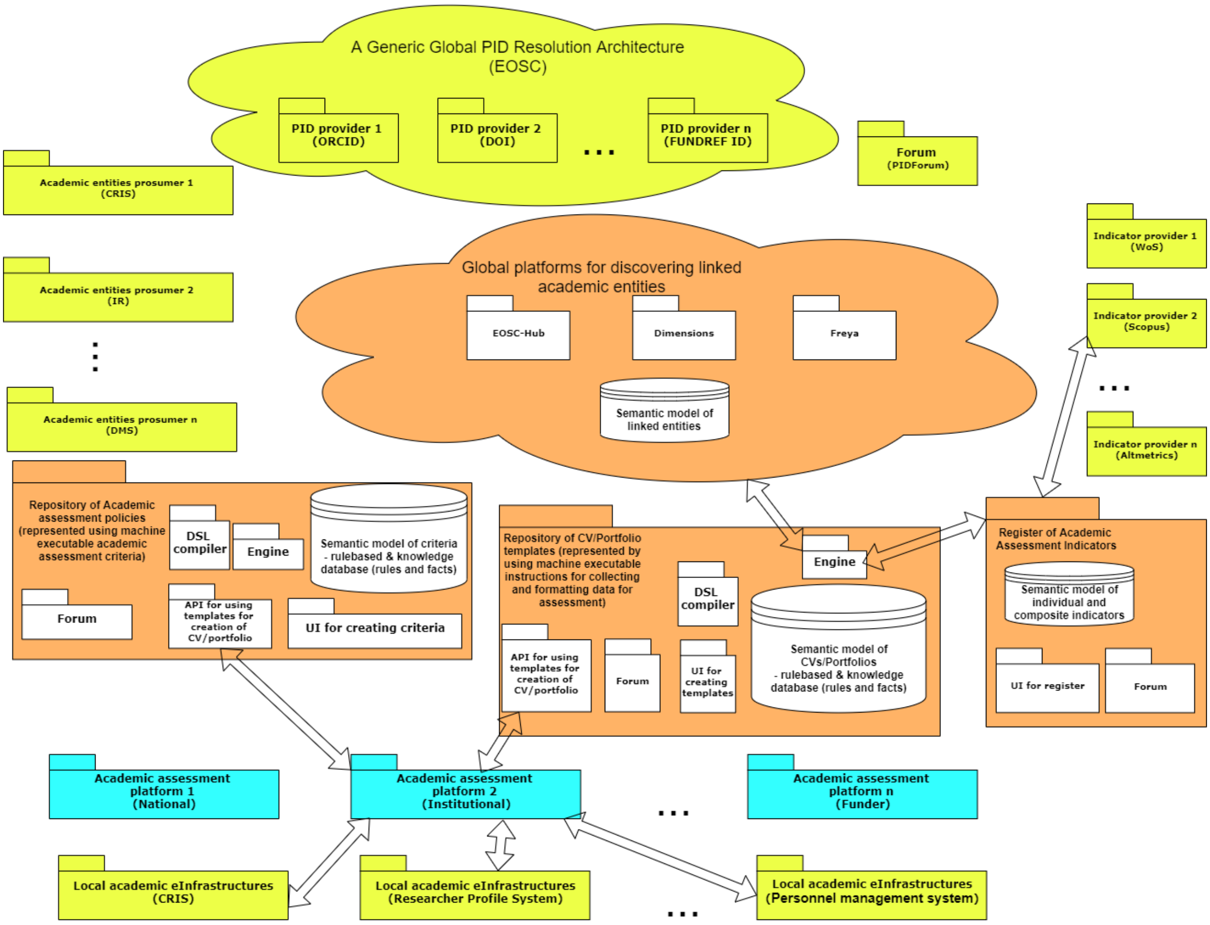


Figure 1. eInfrastructure architecture for responsible academic assessment.

ACADEMIC ENTITIES'PROSUMERS

What

Academic entities prosumers include platforms which produce and consume records representing information about academic entities such as Projects, Organizations, Researchers, Publications, Datasets, Conferences, etc. Those platforms could be based on different models, include different set of entities, and could be implemented for various purposes. Some examples are research information systems, publications' repositories, data management systems, researchers' profiles systems, publications platforms such as Open Research Europe Platform (https://open-research-europe.ec.europa.eu/), etc.

Why

Those platforms could be developed for various purposes, for instance to fulfil local requirements including management of scientific domain and reporting; or to disseminate local results to the global scientific community.

How

Academic entities' prosumers can be developed by adapting and customizing some comercial solution (Pure, Converis, DSpace, EPrints, CKAN, Dataverse, etc.). Moreover, the solution can be built as an in-house solution developed inside an institution from the scratch. The set of functionality could vary depending on the purpose of the platform and available resources (budget, technical support, etc.). Some of those platforms import data from global academic databases and platforms (e.g., Web of Science, Scopus, ORCID) in order to decrease users' efforts for acquiring data. Moreover, some platforms export data manually collected by users to global academic databases and platforms to make information discoverable for the global community. If it is possible, record entities' stored in a prosumer should be published in accordance with FAIR principles.

PID PROVIDERS

What

PID (persistent identifier) is a long-lasting reference to a digital object that is accessible over the Internet. PID providers are organizations responsible for handling requests for globally unique identifiers for certain types of entities.

Usually, there is a platform with an API for making a request for generating a new identifier. Moreover, there has to be an API for resolving an identifier, meaning returning URL to the object linked with the certain identifier. In the current scholarly communication environment, numerous types of PIDs can be recognized, PIDs for outputs (publications, data, software), people, institutions, etc.

Why

A persistent identifier assigned to a digital academic entity enables linking of information about the academic entity preserved in numerous databases. Persistent identifiers are necessary for the interoperability between academic platforms and for collecting and merging information about one academic entity including building inbound and outbound links for getting a comprehensive set of information needed for academic assessment. FAIR principles recommend assigning PID to a dataset.

How

There are some well known PIDs, such as DOI, ORCID, ROR, which might be adopted for the needs of building a global linked academic eInfrastructure ecosystem. However, there are two problems which still exist. The first one is missing appropriate and well-recognized PID for some academic entity types (for instance - indicators, projects). The second problem is the existence of more than one well-known PID for some academic entity types (institution - ROR ID, GRID, FundRef ID, etc.). In order to reach full linking of the academic eInfrastructure ecosystem, an agreement on adopting certain PIDs and assigning those PIDs for academic entities is necessary. PID forum (www.pidforum.org/) might help in making this agreement. Moreover, PID providers should be part of A Generic Global PID Resolution Architecture prescribed by EOSC (https://doi.org/10.2777/525581)

GLOBAL PLATFORMS FOR DISCOVERING LINKED ACADEMIC ENTITIES

What

Linked information enables large scale integration of, and reasoning on, information on the Internet. Academic ecosystem entities include information about researchers, projects, organizations, publications, data, equipment, etc.

Why

Academic assessment processes should take into account various information about research/researchers. Those information can be stored in numerous eInfrastructures with various purposes, numerous supported formats and different volumes of information. Exposing those information in standardised machine readable format might enable linking of those information and making platforms which help evaluators in the process of academic assessment. Even if the academic assessment is fully qualitative (narrative) based, evaluators need information about researchers or references used in the project proposal to get a comprehensive picture of the quality of a project proposal, team or a candidate for some position. Some of the information might be requested in the application process (to be collected by applicants) or expected to be checked by evaluators using available resources across the Internet. Some local repositories and platforms might be used as a source of information, but if those systems are isolated and not linked with other elements of the global academic ecosystem, a comprehensive picture of a researcher's achievements might be incomplete. One recommendation of Leiden Manifesto states that academic assessment should be efficient and decrease applicants' and evaluators' efforts.

How

Some Semantic web technologies (RDF, OWL, SKOS, SPARQL, etc.), persistent identifiers (ORCID, DOI, ROAR ID, etc.), and standardized vocabularies (CERIF vocabulary, CRediT, etc.) should be used for linking academic ecosystem entities. There are some initiatives (Scholix), scientific projects (EOSC-Hub) and platforms (Dimensions, ResearchGraph, OpenAIRE Research Graph) working on linking academic ecosystem entities. Some of them are working on the definition of PID for various academic entity types (Freya - www.project-freya.eu/en).

INDICATOR PROVIDERS

What

There are a few types of academic assessment indicators: quantity indicators, which measure the productivity of a particular researcher; quality indicators, which measure the quality (or "performance") of a researcher's output; and structural indicators, which measure connections between publications, authors, and areas of research. Some well-known indicators can be automatically calculated based on the information in some database. An indicator provider should offer an API for calculation of those indicators using available information in a database for the certain academic objects whose identifiers are provided as an input in the API call.

Why

Academic assessment should be based on a combination of quantitative measures based on indicators and quality assessment performed by evaluators. Indicator providers could make more efficient, fair and transparent calculation of quantitative measures, and thus make a more efficient, fair and transparent process of academic assessment.

How

Indicator providers should offer simple and standardized APIs based on well-known PIDs for identification of academic objects for which indicators should be calculated. Besides getting indicator for single academic objects, API should offer an end point for batch request for efficiency, meaning for requesting calculation of indicators for a set of academic objects (e.g. citations for list of outputs), as well as for cumulative calculations of an indicator for all linked objects to the certain academic object (researcher, organization, project team, etc.). For the time-dependent indicators, start date and end date should be parameters of API request as well.

REGISTER OF ACADEMIC ASSESSMENT INDICATORS

What

Academic assessment indicators and (alt)metrics include number of citations, number of views, h-index, etc. Register is a platform for making rich descriptions of those indicators and (alt)metrics FAIR. Description should also include a list of indicators providers, i.e. platforms which might calculate the value of the indicator (national and institutional CRIS, Crossref, Dimension, Microsoft Academic, WoS, Scopus, etc).

Why

Transparency is an important request for responsible academic assessment. Indicators and (alt)metrics used in academic assessment should be clearly and transparently defined. Moreover, some indicators are not popular, and we don't know for the existence of those indicators. Besides increasing transparency, discoverability and reusability of indicators, the register can be used as a forum, i.e. can be used for building community of interest to usage and building academic assessment indicators and (alt)metrics, discussing pros and cons of indicators, knowledge transfer, definition of good practices for building indicators, etc.

How

There are registers of data repositories (re3data), open-access repositories (OpenAIRE), open-access policies and mandates (ROARMAP, Overton.io, CHORUS Publisher Data Availability Policies Index), services and infrastructures (CatRIS). However, there is no initiative or project for building a register of human and machine readable descriptions of indicators and (alt)metrics at the moment for our best knowledge. Anyway, some of the existing platforms can be extended to become this register (e.g. https://scimeter.org/). A machine and human readable format for description of indicators and (alt)metrics should be defined. It can be part of some existing format (such as CERIF), or can be developed from the scratch. List of registered indicator providers (platforms which can calculate the indicator) and description of their APIs should be assigned to an indicator. An indicator can be a composite indicator, meaning it could be defined as a combination of existing indicators already described in the register. PID for indicators should be defined. The platform should also offer features for building community in form of some open review of described indicators in the register and for discussion about some indicators' issues (something similar as the PID forum - www.pidforum.org/).

REPOSITORY OF CV/PORTFOLIO TEMPLATES

What

CVs/Profiles Templates shape all applications in the same format. Machine readable and actionable CVs/Profiles Templates can be integrated with other infrastructure elements mentioned above. Repository should store CVs/Profiles Templates represented by using machine executable instructions for collecting and formatting data for assessment. Templates represented in this way enable making researchers' CVs/Portfolios by using those templates and collecting data from linked academic entities' platforms. Besides CVs/Profiles Templates in machine readable format, the repository preserves templates' descriptions in a rich metadata format and makes templates FAIR.

Why

Repository of CV/Portfolio templates using machine executable instructions for collecting and formatting data enables storage and discovering of good practices for building CVs/Profiles and makes easier the process of making a researcher's CV/Portfolio through integration of the repository with linked academic entities preserved in numerous platforms across the world, as well as with integration with indicator providers. Moreover, CV/Portfolio instances created using this machine readable template will be also machine readable and can be used as an input to machine executable academic assessment criteria (see the next section). Building of this element of infrastructure is in accordance with a recommendation of Leiden Manifesto that academic assessment should be efficient and decrease applicants' and evaluators' efforts. Moreover, the repository can be used for building the community, meaning there should be functionality of a forum for discussing pros and cons of templates, knowledge transfer, definition of good practices for building templates, etc.

How

A language specific for description of CV/portfolio should be defined (DSL - domain specific language). This language should have constructions/statements/elements for easy creation of common inputs expected in a CV/Portfolio including Open Science results and activities, as well as machine instructions how those inputs can be collected and from where (for instance, only DOI should be specified for a reference or GRID for institution where researcher has been employed). Furthermore, the language should also enable creation of non-common inputs which are complex for collecting from other systems. Moreover, those constructions/statements/elements should enable inputs representing quantitative indicators described in the “Register of academic assessment indicators”. A compiler for validating, parsing and compiling the instruction for collecting and formatting information described in the DSL into rules for some rule based engine should be implemented. The rule base engine can collect data from linked platforms and format it in the prescribed format for the purpose of creation of a CV/Portfolio based on the template. Furthermore, templates should enable adding qualitative (narrative) assessments. The platform should offer an API for integration with other systems which will store and use generated CV/Portfolio. Storing of generated CV/Portfolio in the repository should be optional due to privacy issues, and if it is stored it should be linked to the template. A format for description of metadata related to a template should be defined as well (who created it, when, for what purpose, etc). It can be part of some existing format or can be developed from the scratch. Moreover, a PID should be assigned to a template. The platform should also offer features for building community in the form of some open review of preserved templates in the repository and for discussion about issues related to usage of CV/Profiles templates in the academic assessment process.

REPOSITORY OF ACADEMIC ASSESSMENT POLICIES

What

An academic assessment policy prescribed by a pdf document can include a group of academic assessment criteria. Those criteria can be represented in machine executable format. Machine executable academic assessment criteria can automatically produce final classification based on complex rules built on top of input data provided by evaluators, applicants or indicator providers. Repository of academic assessment policies enables storage, discovering and execution of policies and its criteria. Besides academic assessment policies in machine executable format, the repository preserves pdf files and its descriptions in a rich metadata format making policies FAIR.

Why

Transparency is an important request of responsible academic assessment. Criteria should be transparent and fairly applied to all candidates. Moreover, manual application of complex criteria might be error-prone and time-prone. Automatisation of the application of criteria is in accordance with a recommendation of Leiden Manifesto that academic assessment should be efficient and decrease applicants' and evaluators' efforts. The evaluator can be more focused on simple elements they have to evaluate. If the process is automated it is possible to experiment with novel approaches, evaluate the assessment process, and make it more effective. Moreover, the repository can be used for building the community, meaning there should be functionality of a forum for discussing pros and cons of criteria, knowledge transfer, definition of good practices for building criteria, etc.

How

A language specific for description of academic assessment policies and criteria should be defined (DSL - domain specific language). This language should have constructions/statements/elements for easy creation of common academic assessment criteria (e.g. number of published research datasets), but should also enable creation of non-common and complex academic assessment (e.g. contribution to the community which might be met through some of listed activities). Moreover, those constructions/statements/elements should enable criteria based on quantitative indicators, as well as on results of qualitative assessment provided by evaluators. Furthermore, combining and reusing criteria should be enabled. A compiler for validating, parsing and compiling these criteria (described in the DSL) into rules for some rule based engine should be implemented. The rule base engine can execute criteria using input data. The platform should offer an API for integration with other systems which will store and use results of execution of criteria. Storing of results of assessment in the repository should be optional due to privacy issues, and if it is stored it should be linked to the policy. A format for metadata related to preserved policies and criteria should be defined as well (who created, when, for what purpose, etc). It can be part of some existing format (such as CERIF), or can be developed from the scratch. Moreover, PID for academic assessment policy and criteria should be defined. The platform should also offer features for building community in the form of some open review of preserved policies and criteria in the repository and for discussion about academic assessments issues.

LOCAL ACADEMIC eINFRASTRUCTURES

What

A local academic eInfrastructure could be academic entities' prosumers for global platforms (see the section "Academic entities' prosumers"). Those eInfrastructures could be developed for local specific needs and could preserve information which are not of interest for the global community or for privacy issues can't be exported to the global platforms, and therefore not exportable in the global platforms. However, those information preserved in local platforms might be useful in the institutional academic assessment process. Therefore, those platforms might be represented two times on the diagram, but offer different sets of functionalities for local users and local platforms for academic assessment compared to the set of public functionalities available for global scientific communities and platforms. Some examples are research information systems, personnel management systems, researchers' profiles systems, etc.

Why

Those platforms might preserve information of interest for the local academic assessment process.

How

Those eInfrastructures might be developed by adopting and customizing some comercial solution (Pure, Converis, DSpace, EPrints, CKAN, Dataverse, etc.), or built as an in-house solution developed inside an institution from scratch. It might be exchanging subset of data with global platforms in both directions. However, it should offer a secured API for integration with other local eInfrastructures including the platform for the academic assessment.

ACADEMIC ASSESSMENT PLATFORMS

What

Academic assessment platforms should implement the academic assessment process. It might be used by applicants and evaluators. An evaluator can be a committee board member or an individual evaluator. Moreover, assigning external reviewers for qualitative evaluation of an output should be supported.

Why

Those platforms should centralize academic assessment processes in one institution or region, and make those processes more efficient, fair and transparent.

How

An academic assessment platform might be implemented as a module within existing local academic eInfrastructure or as a new application integrated through secured API with local eInfrastructures. Moreover, it should be connected with "Repository of CVs/Profiles Templates" through its API for the purpose of collecting information available in global platforms which are not of interest or not possible to be collected in the local academic eInfrastructures, and for uniform formatting of data, in machine and human readable format, which should be assessed. Furthermore, it should be connected with "Repository of academic assessment policies" for partial automatization of the process of writing assessment reports.

2. USE CASE

PREPARATION OF AN ACADEMIC ASSESSMENT

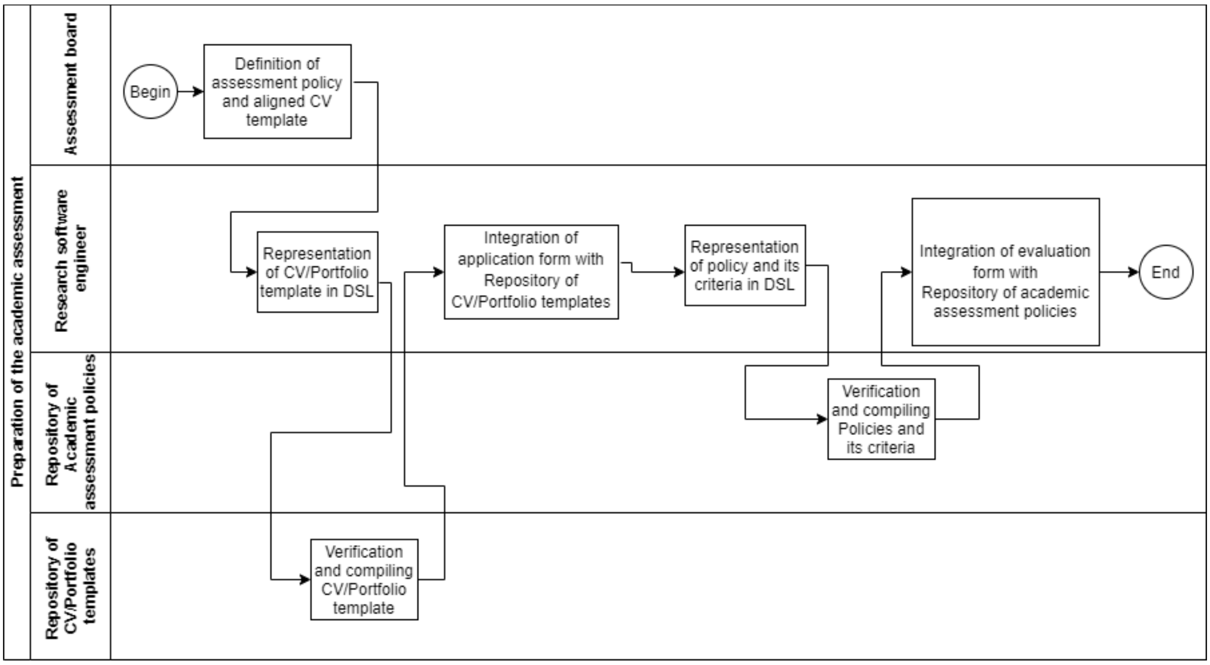


Figure 2. Workflow for preparation of an academic assessment.

Definition of assessment policy and an aligned CV template

Assessment board should create an academic assessment policy with clearly defined criteria in accordance with the purpose of assessment. For instance, for the use case presented in this document the policy could be the following one:

For the process of promotion to the higher position at a university besides qualitative assessment (peer review of committee board members or hired reviewers) of 5 published articles in the last 5 years, popularity of researcher is taken into account. Popularity of researcher is defined in the following way:

- High - number of views of all researcher’s outputs more than 1,000, and number of mentions at social networks more than 100; or an article cited more than 200 times and at least 50 papers presented at international conferences
- Medium - number of views of all research objects more than 300 and number of mentions at social networks more than 30; or an article cited more than 50 times and at least 15 papers presented at international

conferences

- Low - if none of above criteria is fulfilled

The academic board should define narrative description of a CV/Portfolio expected in the application process in accordance with defined policy for assessment.

Representation of CV/Portfolio template in DSL

A research software engineer should represent the narrative description of a CV/Portfolio in domain specific language (DSL) for defining CV/Portfolio template using the GUI of the platform “Repository of CV/Portfolio templates”.

Verification and compiling CV/Portfolio template

“Repository of CV/Portfolio templates” automatically perform verification of defined CV/Portfolio template, compile it to the set of rules for collecting information from global sources and for formatting information in the template, and generate and return PID of defined template.

Integration of application form with Repository of CV/Portfolio templates

The PID of the created template and API of the platform “Repository of CV/Portfolio templates” have been used for integration with the university platform. Moreover, the form for application at the university platform should be developed.

Representation of Policy and its criteria in DSL

A DSL definition of academic assessment policy including “popularity criteria” (see above) which should be used in the assessment process has been created using the GUI of the platform “Repository of academic assessment policies”.

Verification and compiling Policy and its criteria

“Repository of academic assessment policies” automatically perform verification of defined policy, compile it to the set of rules for automatically checking criteria, generate and return PID of defined policy.

Integration of evaluation form with Repository of academic assessment policies

The PID of this policy and API of the platform "Repository of academic assessment policies" have been used for integration with the university platform. Moreover, the form for assessment at the university platform has been developed. At the end of the preparation phase, there is a university platform (Academic assessment platform 2 in the Figure 1) for supporting assessment process which is integrated with "Repository of CV/Portfolio templates" and with "Repository of academic assessment policies" components through APIs.

APPLICATION FOR THE ACADEMIC ASSESSMENT

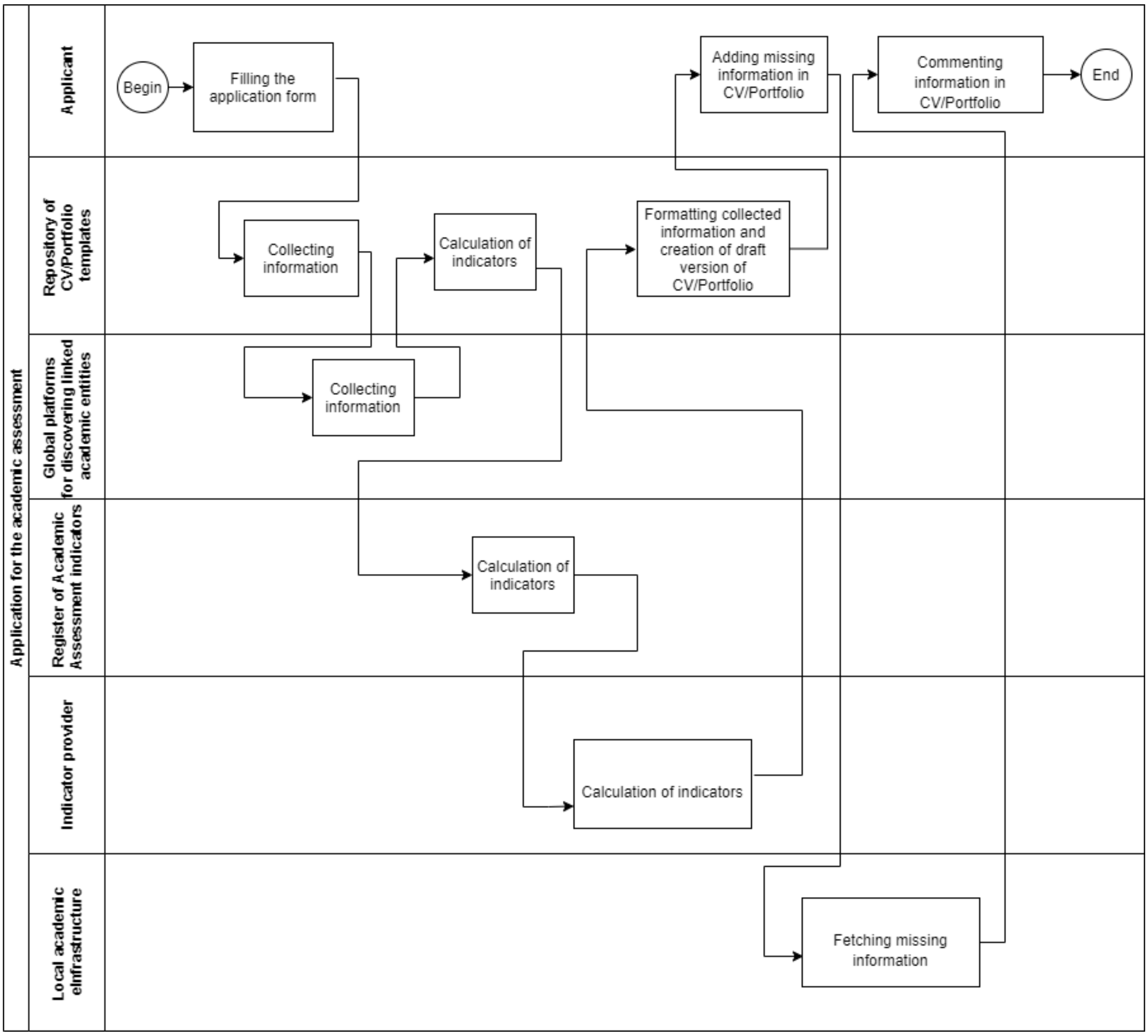


Figure 3. Workflow for the application for academic assessment.

Filling the application form

At the university platform there is a form for application with fields for basic information about a researcher including ORCID and social network accounts, and inputs of 5 DOIs for representative references. Moreover, there can be a narrative field about researcher career and background. Moreover, there is one hidden field - CV/Portfolio template PID.

Collecting information and creation of draft version of CV/Portfolio

When a candidate fills previously mentioned information, an API call to the platform “Repository of CV/Portfolio templates” is fired with arguments: CV/Portfolio template PID, ORCID, social networks accounts, 5 DOIs. As a result of this call the platform “Repository of CV/Portfolio templates” collects academic objects linked with the researcher from “Global platforms for discovering linked academic entities” and “Register of Academic Assessment indicators”, and collects cumulative numbers of views and mentions of the researcher and his/her research objects at social networks, number of citations for each article, and the list of papers presented at international conferences. All those information are shown to the candidate through the university platform. Next to each collected piece of information is information about the source (Scopus, WoS, Twitter, OpenAIRE, etc.).

Adding missing information in CV/Portfolio

The candidate can’t change collected records from global sources, but there is the option for adding missing records in each section (e.g. journal article reference). The university platform can be integrated with local software infrastructures (e.g. CRIS, institution repository, research profile system, etc.), and applicants can fetch records from those local platforms or manually add information. All those added records are visually highlighted in the user interface, and source of information is “local information”.

Commenting information in CV/Portfolio

Next to each information, manually added or automatically collected, there is a free text field where the candidate can discuss information, complain that something is incorrect and provide proof in the case it is manually added information, or state any request that should be taken into account for his/her evaluation. Besides commenting in the free text form, there is the option to upload the file as a supplement for the free-text claim (e.g. email for acceptance of a conference paper which is in the process of printing). All those information are stored on the university platform and will be visible to the evaluator. This is in accordance with Recommendation for the Responsible Evaluation of a Researcher in Finland [4]: "The researcher's self-evaluation is combined with the evaluation by giving an opportunity to express an understanding of the objectives, significance and effectiveness of their work."

THE ACADEMIC ASSESSMENT

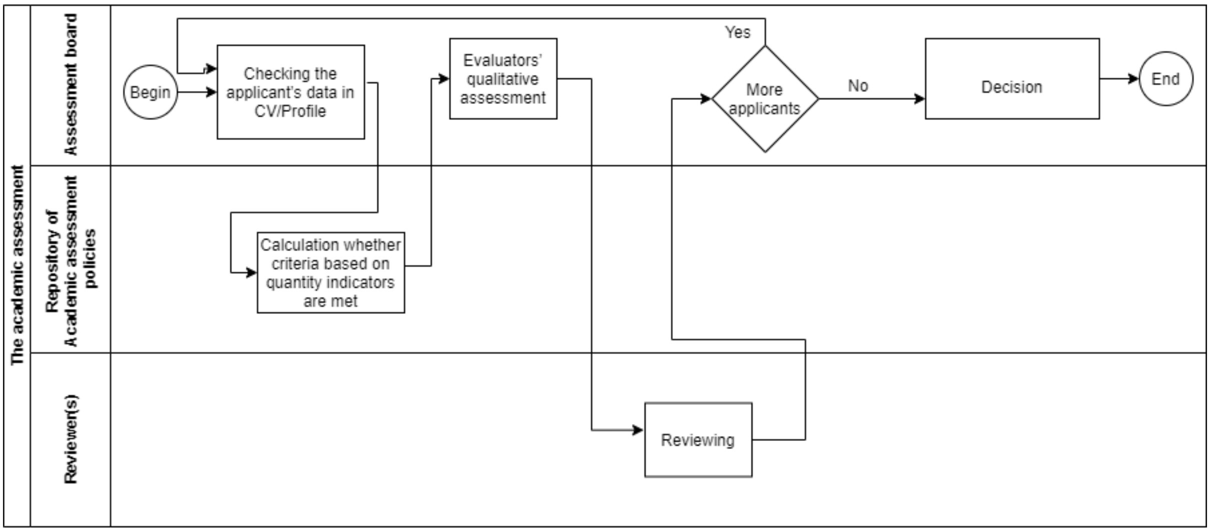


Figure 4. Workflow for performing academic assessment.

Checking the applicant's data in CV/Profile

All data from the application process are presented to an assessment board member in the assessment form. The board representative checks data, especially comments and manually added information which are highlighted in the user interface. Accepting some of the arguments in comments might have an impact on data which will be used for calculation of the popularity criteria.

Calculation whether criteria based on quantity indicators are met

An API call to the platform "Repository of academic assessment policies" is fired with arguments: policy PID, number of views of all research objects, number of mentions at social networks, number of citations for the mostly cited article, number of articles published in journals with impact factors. As a result of this call the platform "Repository of academic assessment policies" runs a rule based engine to classify the researcher's popularity as "High", "Medium", or "Low" based on provided inputs.

Evaluators' qualitative assessment

The assessment board assigns 5 representative references of the applicant to evaluators/reviewers. Next to 5 provided representative references in the application process, there are input fields for an evaluator's comments and scores. There can be more evaluators which don't see each other's comments and scores. Each evaluator reviews 5 representative references and provides final comments and results of the review process.

Decision

The assessment board member can see the report for each applicant, i.e. CV/Portfolio, measure of popularity (High, Medium, Low), and all reviewers' comments and recommendations. Based on those information, the assessment board makes a final decision and writes a short final report of the assessment process. The final report is deposited in the university platform.

3. SWOT ANALYSIS

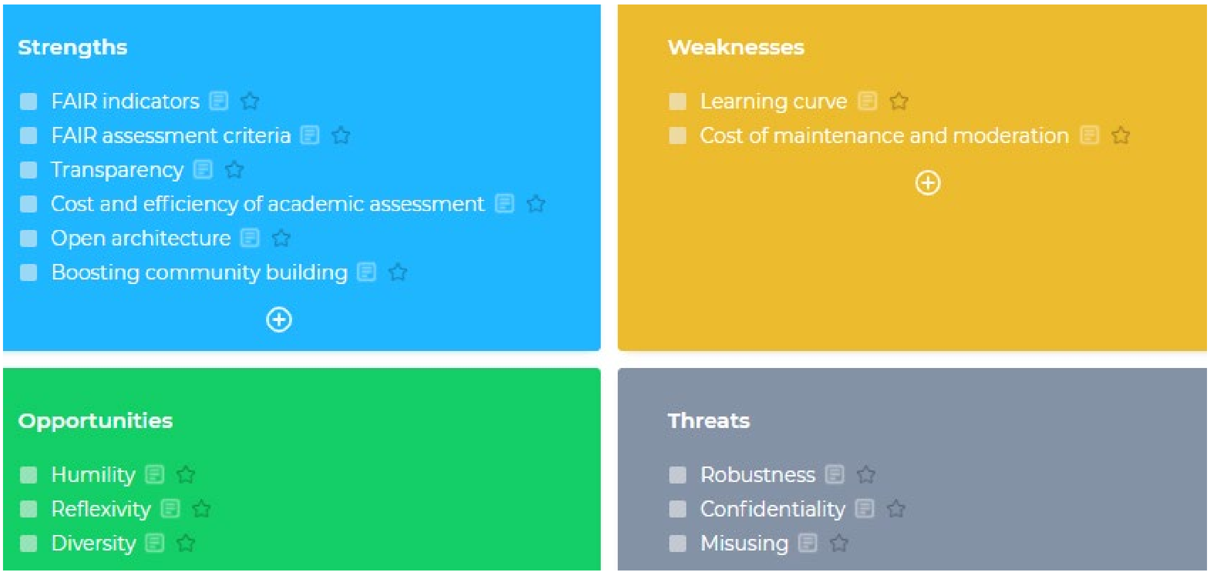


Figure 5. SWOT analysis of the proposed eInfrastructure.

STRENGTHS

1. FAIR indicators

   Bibliometric scholars develop more and more indicators and metrics. Some newly developed indicators may be better than the indicators we have, but we would never know about it because they are not discoverable and well adopted by the available citation tools [5]. The "Register of academic assessment indicators" enables adding information about new indicators, finding and reusing available indicators, as well as raising discussions about pros and cons of indicators and its application in the assessment process. This is in accordance with Next-generation metrics recommendation: "The EC should encourage the development of new indicators, and assess the suitability of existing ones, to measure and support the development of open science" [1].

2. FAIR assessment criteria

   "Repository of academic assessment policies" enables definition of machine actionable criteria and description of those criteria with metadata. Defined criteria are discoverable and reusable for some new academic assessment policies.

3. Transparency

   The proposed eInfrastructure improves transparency of an academic assessment process through definition of FAIR assessment criteria. Although it is optional due to privacy issues (see Threats below), it is also possible to store CVs/Portfolios in "Repository of CV/Portfolio templates" and results of assessment in the "Repository of academic assessment policies" and make academic assessment process completely transparent. The transparency could lead to Impartiality - fair treatment for all applicants.

4. Cost and efficiency of academic assessment

   Applicants' and evaluators' effort are decreased. The application process is more efficient by using CV/Portfolio templates represented by using machine executable instructions for collecting and formatting data for the assessment. Moreover, the assessment process is also more efficient in the aspect of automatic calculation of meeting criteria defined by using quantitative indicators. However, development of local academic assessment platforms requests additional effort made by research software engineers. Those local platforms communicate with numerous data sources through one proxy - the “Repository of CV/Portfolio templates” platform. This simplified implementation of collecting data needed for assessment. On the other side, development and maintenance of “Repository of CV/Portfolio templates” and “Repository of Academic assessment policies” request additional effort made by research software engineers, and moderating its contents requests additional human efforts as well, and it is listed as a weak point of the proposed eInfrastructure (see Weaknesses below).

5. Open architecture

   The proposed eInfrastructure is easily extensible with new indicators, data providers, academic assessment criteria, etc. Indicators can reach its full potential only if they are underpinned with an open and interoperable data infrastructure.

6. Boosting community building

   Part of “Register of academic assessment indicators'' is a forum, which will enhance building a community of bibliometricians and other interested parties in development of indicators. The forum can be used for discussing benefits of using some indicators, strong and weak points, as well as for discussing and reusing best practices for definition of new indicators. Moreover, part of “Register of CV/Portfolio templates'' is a forum, which will enhance building a community of interested parties in development of CVs/Portfolios. At the end, part of “Register of Academic assessment policies'' is a forum, which will enhance building a community of policy makers, evaluators and other interested parties in academic assessment criteria.

WEAKNESSES

1. Learning curve

   Languages for representing assessment criteria and for representing CV/Portfolio templates should be learned, as well as APIs of those two platforms. It requests time, some basic technical skills and previous knowledge.

2. Cost of maintenance and moderation

   There is a cost of development and maintenance of “Repository of CV/Portfolio templates” and “Repository of Academic assessment policies”. Moreover, those systems can’t work properly without moderation of contents stored in these platforms. This moderation requests human efforts, meaning it costs.

OPPORTUNITIES

1. Humility

   The proposed eInfrastructure supports a mix approach - quantitative and qualitative measures, usage of narrative evaluations and indicators.

2. Reflexivity

   The proposed eInfrastructure can be used for experimenting with novel approaches, evaluating assessment processes, or even for cyclical and iterative assessment.

3. Diversity

   The architecture supports including plural characteristics and opening up the range of contributions in the assessment process for the purpose of objectivity and integrity, e.g. selecting best project for funding and best candidate for hiring. The academic assessment performed using proposed eInfrastructure can include various research objects (datasets, publications, software, etc.) and activities (teaching, training, supervising, peer reviewing, engagement with industry, public engagement, securing funding, etc.). Moreover, it can be used by numerous stakeholders - funders, publishers, institutions, researchers. At the end various criteria (appropriateness and inclusiveness) can be defined for different disciplines (SSH, engineering, etc.), project types (applied, fundamental, etc.), career paths (data stewards, research software engineer, researchers, career breaks vs no career breaks, path in industries, etc.).

THREATS

1. Robustness

   Assessment still depends on comprehensives and accuracy of data. Although the eInfrastructure can automatically exclude retracted manuscripts and citations of retracted manuscripts from the assessment process (e.g. by using openretraction api - http://openretractions.com/api/doi/10.1002/job.1787/data.json), some indicators might be gamed.

2. Confidentiality

   Due to GDPR and privacy issues, some policy makers might decide to not publish CVs/Portfolios and academic assessment results. Even worse, some policy makers might decide to publish it without asking for permission, and therefore violate privacy.

3. Misusing

   Due to lack of human resources (reviewers) on one side, and automatic calculation of quantitative indicators offered by the proposed eInfrastructure, some policy makers might choose easier path for assessment, meaning might create assessment policies which use quantitative indicators for quality things (Quantitative measures are for quantifiable things [6]), or even define academic assessment based only on quantitative indicators. In this way, without qualitative assessment performed by experts, some innovative ideas might be unrecognized. Moreover, the eInfrastructure supports criteria based on proxies' indicators (publication channels' indicators) which usage is not recommended by DORA and some other initiatives.

4. CONCLUSION

European Commission should develop a strategy and enable funding for sustainability of the academic assessment process and eInfrastructures. Development and maintenance of the eInfrastructure requests equipment and human resources. Moreover, moderation of the infrastructure content also requests additional human resources, and therefore it costs. However, benefits of objective and transparent assessment are much stronger than introduced costs. eInfrastructure could help evaluators to make the right decision and to select the best projects for funding or best candidates for hiring, and therefore enhance further development of science and society in general.

5. REFERENCES

[1] Wilsdon, J. R., Judit Bar-Ilan, Robert Frodeman, Elisabeth Lex, Isabella Peters, and Paul Wouters. "Next-generation metrics: Responsible metrics and evaluation for open science." (2017).

[2] Wouters, Paul, Mike Thelwall, Kayvan Kousha, Ludo Waltman, Sarah de Rijcke, Alex Rushforth, and Thomas Franssen. "The metric tide." literature review, Supplementary Report I to the Independent Review of the Role of Metrics in Research Assessment and Management, HEFCE, London (2015).

[3] O'Carroll, Connor, Bernard Rentier, Cecilia Cabello Valdès, Fulvio Esposito, Eva Kaunismaa, Katrien Maas, Janet Metcalfe et al. "Evaluation of Research Careers fully acknowledging Open Science Practices-Rewards, incentives and/or recognition for researchers practicing Open Science." Publication Office of the European Union (2017).

[4] Federation of Finnish Learned Societies. "Good Practice in Researcher Evaluation: Recommendation for Responsible Evaluation of a Researcher in Finland." (2020), https://avointiede.fi/sites/default/files/2020-03/responsible-evalution.pdf

[5] Gadd, Elizabeth. "Influencing the changing world of research evaluation." *Insights* 32, no. 1 (2019)

[6] "INORMS Research Evaluation Working Group." (2016), https://arma.ac.uk/rethinking-the-rankings/