

# The analysis of the available software infrastructures for supporting research assessment reform

Authors:

Janne Polonen (TSV, Finland, [janne.polonen@tsv.fi](mailto:janne.polonen@tsv.fi))

Anna-Kaisa Hyrkkänen (TSV, Finland, [anna-kaisa.hyrkkanen@tsv.fi](mailto:anna-kaisa.hyrkkanen@tsv.fi))

Marita Kari (TSV, Finland, [marita.kari@tsv.fi](mailto:marita.kari@tsv.fi))

Dragan Ivanovic (University of Novi Sad, Serbia, [dragan.ivanovic@uns.ac.rs](mailto:dragan.ivanovic@uns.ac.rs))

## Abstract

The CoARA agreement was published in July 2022. There are more than 600 signatories at the moment which agree on the need to reform research assessment practices to recognise the diverse outputs, practices and activities in order to maximize the quality and impact of research. We investigated whether available software infrastructure in the research domain can be used to capture record assets needed for assessment recommended by the CoARA agreement. There is discrepancy between elements needed for the purpose of responsible research assessment and available information in research domain infrastructures. The analysis was conducted as part of the GraspOS project.

## The CoARA agreement

The agreement was published in July 2022. The goal of this document is to give recommendations for reforming research assessment to be more responsible and include elements of the Open Science paradigm. The agreement recognises the diversity of research activities and practices, with a diversity of outputs. Moreover, the agreement recommends rewarding early sharing and open collaboration. It means, the research assessment should include the full range of research outputs, such as scientific publications, data, software, models, methods, theories, algorithms, protocols, workflows, exhibitions, strategies, and policy contributions. Moreover, it should include different activities such as peer review, training, mentoring and supervision of Ph.D candidates, leadership roles, science communication and interaction with society, entrepreneurship, knowledge valorisation, and industry-academia cooperation. Furthermore, early knowledge and data sharing, as well as open collaboration including societal engagement where appropriate, should be incentivised and recognized by research assessment.

# The GraspOS project

The GraspOS project aims to revolutionize the rewards and recognition system in science by introducing innovative metrics and indicators, fostering a culture of quality, transparency, and trust in research. It seeks to establish a community-driven qualitative information system and a revamped recognition system. GraspOS is creating an open, trusted federated dataspace for next-gen research metrics, offering data and tools to support research assessment reforms at various levels, led by the Athena Research Center in Greece and a diverse consortium of partners.

## Methodology

The analysis was based on results of the surveys and resource review.

Two surveys were conducted during the spring 2023 by the Federation of Finnish Learned Societies (TSV):

- GraspOS landscape survey for pilots (March 2023)
- GraspOS landscape survey on Reforming Research Assessment (May 2023)

The purpose of the surveys was to gain overview of the state-of-the-art research assessment practices at the research performing and funding organisations, and other organisations involved with research assessment. The first and the second survey included 28 and 49 questions, respectively, concerning the respondents' background, assessment frameworks and policies, quantitative and qualitative criteria and practices, as well as supporting software infrastructure, i.e. practices and sources of information used to support assessments. Both surveys were conducted online by using the LimeSurvey tool. The GraspOS landscape survey for pilots was launched on the 2nd of March 2023 and disseminated to the GraspOS project pilots' representatives (10 responses collected). The GraspOS landscape survey on Reforming Research Assessment was launched on the 11th of May 2023 on the GrapsOS website. A total of 54 full submitted responses were received.

Resource review was conducted in the period January-May 2023. The methodology for finding available resources includes:

- Previous knowledge of the project participants
- Searching of publications/citation databases using keywords of the project
- Social networks
- Relevant resources chaining
  - citation analysis to/from other resources
  - searching of relevant project/network websites

All collected resources were characterized by certain fields. The result of this characterization can be found at [this link](#) in the tab Infrastructures.

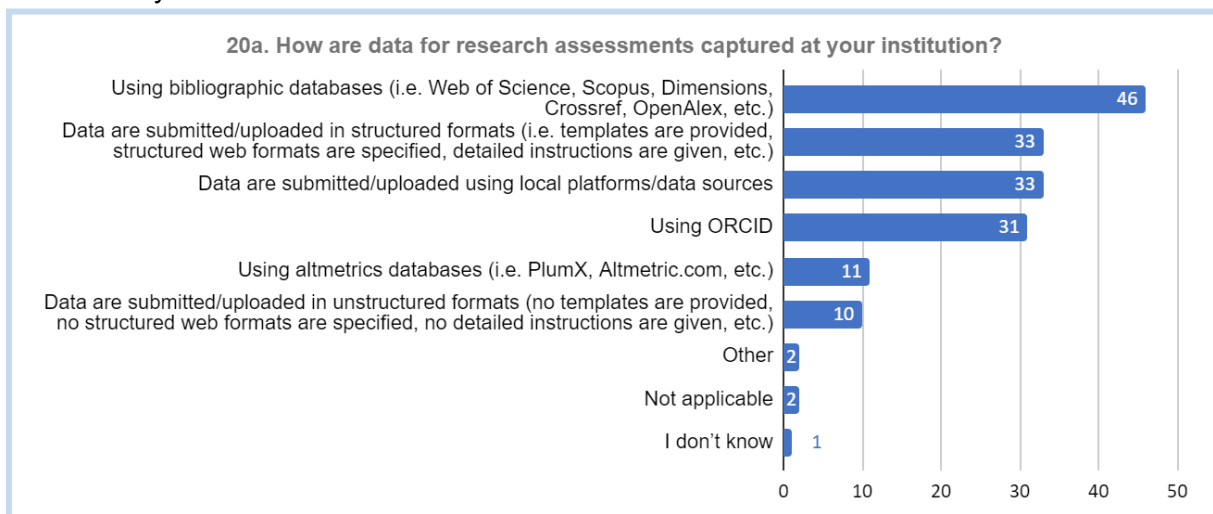
## Results and Discussion

Figure 1 shows the survey response on a question “How are data for research assessments captured at your institution?”. The large majority of respondent organizations (85.2%) capture data supporting research assessment by using bibliographic databases (i.e. Web of Science, Scopus, Dimensions, Crossref, OpenAlex, etc.). In around 60% of the respondent organizations

data are submitted/uploaded in structured formats, for example using templates, web formats and instructions, often using local platforms/data sources, and using ORCID.

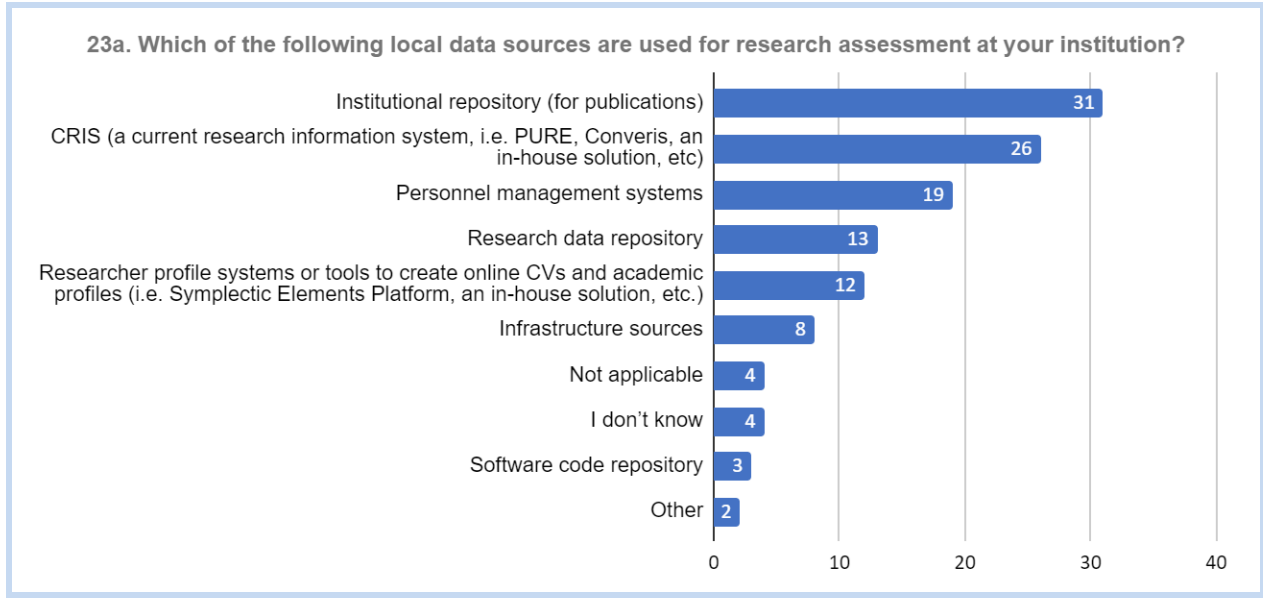
Some respondents described their responses further:

1. "Local platforms" seems to overlook national CRIS systems that include quality assurance by a governmental organization. Such platforms exist, they offer advantages over commercial databases (being publicly funded and in the public domain), as well as advantages over non-verified data sources (like ORCID) being quality-controlled (that is, verifying all entries through mandatory evidence) by a public/governmental actor
2. "The Current research information system of the University has a central role in research evaluations and smaller scale analysis. Data from international citation databases is used as well.
3. In the larger scale evaluations the data used has been captured from the Current research information system of the University as well as using bibliographic database Web of Science. University Library's publication data and publication metrics specialists deliver the data in the agreed format to all the stakeholders of research assessment. Units that have been assessed have had a possibility to check the data used in the analysis.



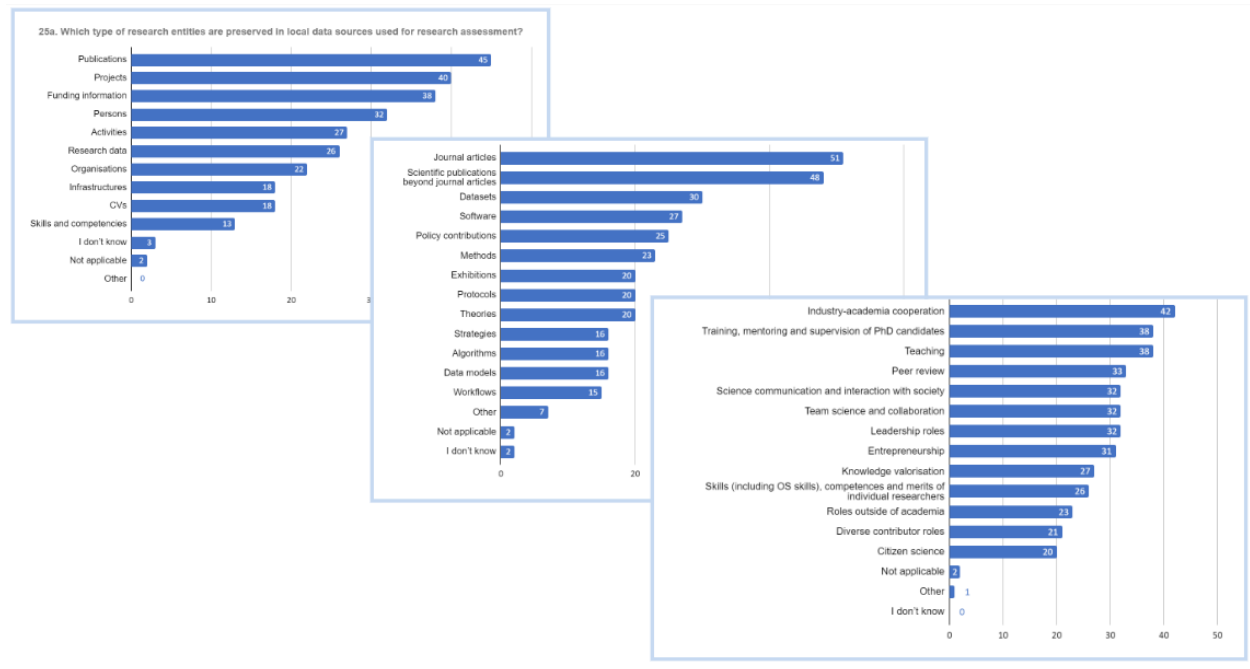
**Figure 1.** The survey question about capturing data for the research assessment process

Figure 2 shows the survey response on a question "Which of the following local data sources are used for research assessment at your institution?". Majority of the organizations use for research assessment either institutional repository (57.4% of the respondents) or a CRIS - Current research Information System (48.1%). 41 organizations (75.9% of all respondents) have a local publication repository and/or CRIS, and 25 of them have one or the other kind of local system. Other local systems, such as personnel management or profiling systems or data or software repositories, are more rarely used and almost always in conjunction with local CRIS or publication repositories.



**Figure 2.** The survey question about local platforms used for the research assessment process

We also investigated the research assessment elements recommended by the CoARA agreement which can be captured from those local data sources (Figure 3).



**Figure 3.** Research assessment elements which can be captured from local data sources

Moreover, we investigate the same question for global data sources OpenAlex, OpenAIRE, Crossref, ORCID record registry, and BIP! Services. Also, we analyzed the capacity of the CERIF interchange model and VIVO ontology to support interchanging of the information



journal articles							
Datasets	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Software	No	Yes	Yes	No	No	Yes	Yes
Policy contributions	No	Yes	Yes	No	No	Yes	Partially <sup>1</sup>
Methods	No	No	Yes	No	No	Yes	Yes
Protocols	No	No	No	No	No	Yes	Yes
Exhibitions	No	No	Yes	No	No	Yes	Yes
Theories	No	No	No	No	No	No	No
Strategies	No	No	Yes	No	No	No	No
Algorithms	No	No	Yes	No	No	No	No
Data models	Yes	Yes	No	No	No	Yes	No
Workflows	No	No	No	No	No	No	No

**Table 3.** Support for representation of other research activities and roles in global platforms and models

Activities and roles	OpenAlex	OpenAIRE	ORCID profiles	Crossref	BIP! Services	CERIF	VIVO
Industry-academia cooperation	No	No	Yes	No	No	Yes	Yes
Training, mentoring and supervision of PhD candidates	No	Yes	Yes	No	Yes	Yes	Yes

<sup>1</sup> legislations, and standards can be represented in the VIVO ontology

Teaching	No	Partially <sup>2</sup>	Partially <sup>3</sup>	No	No	Partially <sup>4</sup>	Yes
Peer review	No	No	Yes	Yes	Yes	Yes	Yes
Leadership roles	Partially <sup>5</sup>	Yes	Yes	No	No	Yes	Yes
Entrepreneurship	No	No	Yes <sup>6</sup>	No	No	Yes	Yes
Science communication and interaction with society	No	No	No	Yes <sup>7</sup>	No	Partially <sup>8</sup>	Partially <sup>9</sup>
Team science and collaboration	No	No	No	No	No	No	No
Skills, competence and merits	No	No	Yes	No	No	Yes	Yes
Knowledge valorization	No	No	Yes	No	No	Yes	Yes
Roles outside of academia	No	No	Yes	No	No	No	Yes

<sup>2</sup> Learning material (such as lecture) can be catalogued and linked with the author of the material

<sup>3</sup> Learning material (such as lecture, speech) can be catalogued and linked with the author of the material

<sup>4</sup> Learning material (such as course, presentation) can be catalogued and linked with the author of the material

<sup>5</sup> There is a role of the publication corresponding author

<sup>6</sup> Spin-off companies, inventions, and patents can be listed in a ORCID profile

<sup>7</sup> Event data can be used for assessment of "science communication and interaction with society" (twitter post' views, blog views, etc.)

<sup>8</sup> Newsclipping can be represented in the CERIF data model

<sup>9</sup> Blogs, interviews, newsletters can be represented in the VIVO ontology

Diverse contributor roles (Data steward, software engineer, and data scientists)	No	No	No	No	Partially <sup>10</sup>	Yes	Yes
Citizens science	No	No	No	No	No	No	No

## Conclusion

There is discrepancy between elements needed for the purpose of research assessment and available information in research domain infrastructures.

It is not likely that some research infrastructure can offer a reliable source of all recommended information for research assessment prescribed by the CoARA agreement. Therefore, the federated software infrastructure for supporting Open Science-aware research assessment should be built. There are following challenges in building this federated infrastructure:

- Interoperability
  - standardization of cataloging formats and practices,
  - wide adoption of persistent identifiers (DOI, ORCID, ROR ID, RAiD), and
  - definition of protocols and application interfaces for exchanging information
- Sustainability of the components (funding)
- Openness of data (data should be the academic-community property)

---

<sup>10</sup> CRediT taxonomy for contribution roles is supported which includes Data curation and Software