10.5281/zenodo.13629147

Pilot findings and progress report

Deliverable 5.2

Due Date of Deliverable	31 August 2024
Actual Submission Date	30 August 2024
Work Package	5
Type	R = Report
Approval Status	Submitted
Version	1.0
Number of Pages	80

Abstract

This interim report on the status of the pilots and main achievements describes the work the pilots have done in order to move forward to the next phase of the project, to testing and evaluating the tools, services and/or infrastructures developed in GraspOS. To support this work and to test and evaluate the Open Science Assessment Framework (OSAF) being developed by Work Package 2, the pilots have been following the SCOPE Framework, which is the basis on top of which the OSAF is being built. One of the central results from this work is the realisation that each evaluation setting (i.e. pilot) is different, there are no "one-size-fits-all" solutions for evaluations, therefore, within the project, we need to be sensitive to differences.

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

Revision history

Version	Date	Reason	Revised by
0.0	3.6.2024	First Draft	Clifford Tatum Andrea Mannocci Carol Delmazo Odile Hologne Laura Himanen Iiris Liinamaa
0.1	25.6.2024	Peer review	Elina Koivisto Jarno Hoekman
0.2	23.8.2024	Peer review comments addressed	Iiris Liinamaa Joonas Nikkanen
1.0	30.8.2024	Final Version after proofreading	Iiris Liinamaa

Author List

Organisation	Name	
CSC	Laura Himanen	laura.himanen@csc.fi
CSC	Iiris Liinamaa	iiris.liinamaa@csc.fi

Contributor List

Organisation	Name	Contact Information
CNR	Andrea Mannocci	andrea.mannocci@isti.cnr.it
CSC	Joonas Nikkanen	joonas.nikkanen@csc.fi
CWTS	Clifford Tatum	c.c.tatum@cwts.leidenuniv.nl
OPERAS	Fotis Mystakopoulos Carol Delmazo	fotis.mystakopoulos@operas-eu.org carol.delmazo@operas-eu.org
UU	Anestis Amanatidis	a.amanatidis@cwts.leidenuniv.nl
UNIBE	Ana Djordjevic	anadj@chem.bg.ac.rs
UEF	Katri Rintamäki	katri.rintamaki@uef.fi
UEFISCDI UEFISCDI UEFISCDI	Alina Irimia Ioana Trif Ioana Spanache	alina.irimia@uefiscdi.ro ioana.trif@uefiscdi.ro ioana.spanache@uefiscdi.ro

Table of Contents

List of Figures

1. Executive Summary

This deliverable describes the process leading up to the next phase of the GraspOS project for the pilots: the testing and validating of tools, services and infrastructure being developed in Work Packages (WP) 3 and 4. For this process, the pilots have tested and evaluated aspects of the Open Science Assessment Framework (OSAF) being developed in WP2. More specifically, the pilots have taken part in several workshops, online and face-to-face, where different stages of the SCOPE Framework have been followed in order to design and plan more responsible evaluation settings. The SCOPE Framework is embedded in the OSAF, which provides both a common approach to Responsible Research Assessment (RRA) and a scaffolding for implementing assessment-specific infrastructure.

The results of this process are summarised as follows:

- Each evaluation (i.e. pilot) setting is different, there are no “one-size-fits-all” solutions for evaluations, therefore, within the project, we need to be sensitive to differences.
- There is a great variety between the pilots in terms of their role in the evaluation settings, especially in terms of having a mandate to make the necessary interventions to reform the evaluation. This needs to be taken into consideration in regard to piloting activities.
- Respecting context and diversity is imperative for realisationing more responsible research assessments, not only in targets of evaluation (cf. Reform on Research Assessment[1]), but also in the actual evaluation process.
- Designing an evaluation requires several aspects to be taken into account. These include, defining what is wished to be evaluated, who or what is the target of the evaluation, and why the evaluation is needed, i.e. defining the values and considering the context, are the most important, but also challenging steps.
- Data, tools, or services available for evaluation should be made to adapt to the needs and requirements guided by the values and context.

[1] https://coara.eu/agreement/the-agreement-full-text/

2. Introduction

In the first half of the operation for GraspOS project, the pilots have concentrated on testing and evaluating aspects of the Open Science Assessment Framework[2] (OSAF) being developed in WP2. This form of codevelopment aims to provide reciprocal input into both the OSAF and the ongoing pilot developments. More specifically, the pilots have taken part in several workshops, online and face-to-face, where different stages of the SCOPE Framework[3] have been followed in order to design and plan more responsible evaluation settings. The SCOPE Framework is embedded in the OSAF, which provides both a common approach to Responsible Research Assessment (RRA) and a scaffolding for implementing assessment-specific infrastructure, such as the Assessment Portfolio and Assessment Registry. One central result of this work is the realisation that each evaluation (i.e. pilot) setting is different, and thus there are no "one-size-fits-all" solutions for evaluations. Therefore, within the project context, we need to be sensitive to differences. The original pilot categorisation is still somewhat relevant, being based on the level of evaluation, but mainly only for reporting purposes.

It needs to be emphasised that forced cooperation between pilots that differ in terms of their level of evaluation, their target of evaluation, their purpose of evaluation, their stakeholders, and their own role in the evaluation setting is not only difficult but potentially detrimental. During the first half of the project, it has become clear that creating unifying guidelines and forcing pilots (evaluations) into similar moulds in terms of methodology and tools (for example) is not an option. Instead, respecting their context and diversity not only in targets of evaluation (cf. Reform on Research Assessment[4]), but also in the actual evaluation process is imperative to achieve more responsible research assessments. This is an important outcome in line with the use of the SCOPE framework feeding back into our understanding of what the use of the SCOPE framework is likely to deliver.

In this report, we briefly present the OSAF method, focusing on the SCOPE Framework, and then explain how the different stages of SCOPE have been followed in a series of dedicated workshops organised for the pilots:

- S - Stakeholder mapping workshop for pilots in October 2023 (see chapter 4.1)
- C - Pilot workshop in November 2024 (see chapter 4.2)

[2] https://zenodo.org/doi/10.5281/zenodo.10475459
[3] https://doi.org/10.26188/21919527.v1
[4] https://coara.eu/agreement/the-agreement-full-text/

- O and P - Pilot workshops in January 2024 (see the chapter 6.1) and May 2024 (see the chapter 6.2)

The goal of the workshops was to work towards a roadmap on how the pilots will test and evaluate the tools, services and/or infrastructure it has chosen as options for evaluation.

3. Methodology

The OSAF is being developed to "facilitate the use of research assessment specific infrastructure informed by sensibilities of both the Open Science and Responsible Research Assessment movements." (Tatum et al., 2023, p.15) The OSAF has three components, (1) SCOPE+i method, (2) Assessment portfolio, and (3) Assessment Registry. The content of the SCOPE+i method includes development of resources (e.g. templates, guidelines, and checklists) that aim to provide practical support in formulating new assessment practices in local contexts. The OSAF is guided by the principles presented in the CoARA Agreement (European University Association et al., 2022) as well as in the SCOPE Framework. In addition to being guided by the SCOPE Framework principles, on an operational level, the SCOPE framework itself has been adopted as the common approach for piloting Open Science aware responsible research assessments, and all nine pilots follow the SCOPE process. The OSAF approach to research assessment begins with the premise that context, purpose and values inform the development of an assessment protocol, and each context is different. This is realised by following the first four stages of the SCOPE Framework which is a framework for research evaluation in form of a step-by-step process:

- S - Start with what you value
- C - Context consideration
- O - Options for evaluation
- P - Probe deeply.

The connections to SCOPE stages during activities and workshops are presented in more detail in the following sections. The final stage, E - Evaluate your evaluation, takes place after conducting the evaluation designed according to the first four stages, so at this point in the project timeline it is not relevant for the pilots.

4. Start with what you value

The first stage of the SCOPE Framework is about finding out what is valued about the entity under evaluation. This may seem like a redundant reminder, but it is very often the case that evaluators start with the data sources available to them (e.g. bibliometric data) or with the values of third parties, such as government, funders, and university ranking agencies. In the OSAF "start with what you value" is a part of assessment readiness which is the first assessment event phase.

For this phase, the first step is to understand who are the stakeholders, who should be involved in the discussions about value. For this purpose, an online workshop was organised for the pilots in October 2023 (presented in more detail in section 4.1.). The second step is to understand what it actually is, what you value about the entity you are seeking to evaluate, and here looking a little deeper into what a “value” means and the different layers of values can be helpful. Values were discussed in a face-to-face (F2F) workshop organised in November 2023 (presented in more detail in section 4.2.).

4.1. Stakeholder mapping workshop

To be able to “start with what you value”, pilots needed first to explore who is the “you”. To facilitate this exploration, as a pre-assignment to the workshop, the pilots conducted a stakeholder mapping following a dedicated template (Annex 1). The template requires identifying all relevant stakeholders, considering their role in the evaluation, and when they have a role, as well as their relationship to the outcome of the evaluation. The important questions were: 1) who determines what is valued, and 2) who defines the purpose of the evaluation. The stakeholder mapping template is part of the OSAF method, which is built on the template created for the pilot workshop, and developed based on pilot feedback and experiences.

In the online workshop, the pilots first presented their stakeholder mappings to give participants a better understanding of the environments in which each of the pilots operate. In moving forward with the project, it is important to clearly understand what the possibilities and limitations of each of the pilots are in terms of being able to implement GraspOS tools, services and infrastructures.

The stakeholder mapping showed that there is a great variety between the pilots in terms of their own role in the evaluation setting. Some pilots have a mandate to make the necessary

interventions to reform the evaluation, some pilots have no authority, and some pilots are in between these two opposites. When planning ahead, these differences need to be taken into consideration in regard to piloting activities, as well as possibly modifying the Key Performance Indicators (KPIs) of some of the pilots.

In terms of stakeholders, there was less diversity. The role of decision-making on conducting an evaluation, evaluation criteria and the utilisation of evaluation outcome depends on the level the evaluation takes place (national, institutional, thematic). This group of stakeholders included governmental actors, senior management, scholarly societies and research funding organisations. For the role of planning evaluations, most pilots reported on committees, steering groups and scientific councils, as well as support services, such as bibliometric teams and research services. Researchers were reported to have two roles, evaluators and targets of evaluation.

The public, media and businesses are considered as utilisers of evaluation outcome as audiences. Naturally, the stakeholders reported as decision-makers in terms of the evaluation setting (criteria and utilisation of outcome), were also reported as the utilisers of the evaluation outcome.

After the presentations, the pilots were divided into three groups according to the original categories (national level, institutional level, and thematic level) to further discuss

A. the role/meaning of stakeholders in evaluation design
    - specific questions addressed:
        - Who is in control of the research process?
        - Who creates knowledge strategy for the organisation?
B. the role/meaning of the evaluation itself
    - specific questions addressed:
        - How will this evaluation design influence knowledge creation and sharing?
        - More specifically, will it incentivise certain types of activity?
        - Will it have a negative or positive influence?
        - etc.

One of the central challenges in discussing these issues was the diversity of pilots, which caused the discussions to be conducted more as small reports of each individual pilot's stand in regard to the questions as opposed to mutual conversation or exchange of ideas.

4.2. Values in F2F workshop

The second step in starting with what you value is to understand what you value about the entity you are seeking to evaluate. To facilitate this exploration, as pre-assignment to the workshop, the pilots conducted a "start with what you value" statement using the SCOPE format of super-values, values, and (if applicable) sub-values. Value statement template is part of the OSAF method, which is built on the format of the statements done by the pilots for the workshop, and developed based on pilot feedback.

According to the SCOPE full guide, a value is "a judgement made about what is important" (International Network Of Research Management Societies-Research Evaluation Group 2023, p. 8). However, the guide recognises three different layers of values, the consideration of which supports getting to the heart of what is valued about a particular entity, and should therefore be the target of evaluation:

1. **super-values** are at the highest level, are often stated as single words (e.g., openness, diversity, inclusivity), and can be useful in steering an evaluation, but lack the level of detail to be used in the design of an evaluation
2. **values** are at the next level down, and they can be understood by asking how the super-value manifests itself: these are the things you want to evaluate
3. **sub-values** are at the lowest level of granularity, these are what your values look and feel like

As the pilots are diverse, they were given the freedom to adapt the approach as needed. It was also allowed to make a statement that was preliminary or aspirational, for example, depending on the local circumstances. In the value statements, the pilots were quite unanimous in considering diversity as a super-value underpinning their evaluation setting. How it translated to values, i.e. how it manifests itself, and what it is more specifically that the pilots wish to evaluate varied somewhat, but there was a strong consensus on wanting to enable the consideration of a diversity of contribution, in terms of outcomes as well as researchers' roles and careers. In addition, considering a broad range of activities, including research work ranging from theory to application and multiple forums for disseminating research work was emphasised. Impact, understood in many ways, was another generally identified super-value. It was considered to manifest itself as outreach, expertise for the benefit of society and communication. Openness was considered as a super-value, but also as a value connected to super-values like collaboration, transparency and impact. Other super-values mentioned were experimentation, reflexivity, responsibility, knowledge,

excellence, equal opportunities, professionalism and efficiency. The last two relating to the services or systems being developed as part of a pilot setting.

Sub-values could also be identified, written out mainly as practical ideas on how to evaluate the given values.

Going through the value statements strengthened the understanding of pilots (and evaluations) being unique, as the same super-values could be understood in different ways in terms of what should be evaluated.

In the workshop, the pilots were divided into three groups according to the original categories (national level, institutional level, and thematic level) to further discuss values in terms of whose values count and, more specifically, who should be included in formulating the evaluation criteria, or the actual meaning of a given value and what it means to evaluate that value, how can it be evaluated. Also, the role of each of the pilots in formulating what is valued was discussed.

4.3. Pilot summary on start with what you value

To support the work on the OSAF the pilots were asked to report on what was their perception of how the workshops and pre-assignments have supported them in moving forward in the project. The pilots were provided with a template (Annex 2.) to facilitate summative reporting in this interim report.

In regard to the first stage, start with what you value, the pilots reported on their learnings for both workshops described above in chapter 4.2. The feedback provided by pilots will be useful in the development of the tools and services by the project as a whole. The plans are described in the unique roadmaps per each pilot that can be found in the Annex 3.

4.3.1. Stakeholder mapping

Stakeholder mapping was considered to be very important for all the pilots. Simply paying attention to the sheer amount of stakeholders that are somehow affected by an evaluation, as well as their diversity, supported seeing the different contexts and uses involved in one evaluation setting. Through the stakeholder mapping, the pilots were able to assign importance, or weight, to stakeholders based on their role in evaluation as well as their relationship to outcomes, which, for example, allows focusing resources to cater to the needs of the most relevant stakeholders. This was also considered to support identifying relevant

people or groups to take part in specific pilot activities based on their relationship with the evaluation setting. Stakeholder-centred approach was seen to ensure addressing real-world use cases and ultimately maximising value for intended audiences.

The mapping was also considered revealing the potential implications of evaluations to not only their targets, but also their operational environment.

Another key contribution of the stakeholder mapping was its ability to clarify the pilots' roles in regard to their evaluation setting. As mentioned earlier, in moving forward in GraspOS it is essential to consider the pilots' authorities to make interventions to reform evaluations, and this implies recognising the pilots' challenges and limitations.

4.3.2. Value statement

Completing the value statement helped the pilots to think more clearly about the starting point of evaluation. It also highlighted the need to keep updating "what is valued" and supported identifying what needs to be changed in order to keep the focus of evaluation relevant.

At least two types of starting points could be detected, technical and ideological. Technical in the sense that the value was attached to the way an evaluation is operationalised or how an evaluative agency works (or should work). And ideological in the sense that the value was considered as something that guides what is evaluated (much as SCOPE originally intended). Where the starting point for evaluation is, depends on the function of the given pilot, as well as who takes part in the discussions about what is valued. There were some considerations on who ultimately is in charge of interpreting the values guiding the evaluation, which implies a need to educate evaluators on what it means to "start with what you value".

Through the value statement exercise, it became evident for the thematic pilots that defining values is very challenging in their context. They considered the exercise to focus on an organisational level, which makes sense, as evaluations are mostly conducted by organisations and, to some extent, also nations, not thematic communities, and the values identifying the targets of evaluation should therefore be assigned at the level of organisations or nations.

5. Context considerations

The second stage of the SCOPE Framework is context consideration. Evaluations need to be context-specific across two dimensions: who or what is being evaluated and why. It is impossible to figure out the proper methods and tools for evaluation without knowing the context. The OSAF gives a special emphasis on contextual factors, so in addition to the SCOPE Framework, it also draws on an expert report, Indicator Frameworks for Fostering Open Knowledge, commissioned by the European Commission[5]. In the OSAF, "context considerations" along with "purpose", is a part of assessment readiness which is the first assessment event phase.

When defining the entity being evaluated, its size and discipline need to be considered. Size, because when it comes to indicators, the smaller the entity, the less reliable indicators are in terms of their ability to tell something meaningful about that entity. And discipline, because it makes a difference in terms of, for example, funding opportunities, methods used and publication practices. As a result, an evaluation approach completely suitable for one discipline might be inappropriate for another. The pilots explored the who or what, as well as the why in the face-to-face workshop conducted in November 2023 (presented in more detail in section 5.1.)

[5] Indicator frameworks for fostering open knowledge practices in science and scholarship:

https://data.europa.eu/doi/10.2777/445286

5.1. Workshopping context

In the workshop, the pilots were presented with the evaluation impact matrix (Figure 1.)

		Country	HEI	Group	Individual
Analysis	To understand				
Advocacy	To show off				
Accountability	To monitor				
Acclaim	To benchmark				
Adaptation	To incentivise				
Allocation	To reward				

Low impact
Medium impact
High impact

Figure 1. Evaluation impact matrix. Adopted from the SCOPE Full guide. (International Network Of Research Management Societies-Research Evaluation Group, 2023, p.12.)

The evaluation impact matrix plots six key reasons for evaluation against four different entity sizes. Each segment is RAG (red-amber-green) rated to indicate high, medium or low impact for evaluations in a specific context. The purpose of the matrix is to highlight that evaluations in some settings have more impact on the entity being evaluated and are, therefore, a higher risk. For the majority of the pilots, the reasons for evaluation are to understand (analysis) and/or to show off (advocacy). However, all reasons included in the matrix are represented in some of the pilots. And well in line with the different levels of evaluation represented in the pilots, evaluations take place for all four different entity sizes. So here again, we witness the diversity between the pilots.

Taking into consideration that the pilots are each located in a particular context, or rather in layers of contexts, i.e. national, organisational, infrastructural, epistemic, and career stage, in the workshop an exercise on identifying contributions in relation to context and purpose was

conducted around the same value, advancing open science. All pilots were to use 'advancing open science' as their starting point for evaluation, and identify unconventional and overlooked contributions that could be used as evidence of it (conventional referring to open access publications and open data). Surprisingly, for the most part, the three groups did not come up with the same contributions, in fact only open software, citizen science and engaging with open science were identified in two groups. The rest of the contributions were open methodology, open infrastructure, popularised publications, outreach and collaboration. The pilots were very aware that for some of these contributions finding data is very difficult, if not impossible, but the idea was to brainstorm without the restrictive requirement for available data.

5.2. Pilot summary on context considerations

In regard to context consideration, the pilots were asked to report on how taking into account the context of their pilot (i.e. the entity size, discipline, etc.) supported their work.

Generally speaking, context consideration was executed by the pilots as ticking the relevant segments in the matrix (Figure 1). However, when given more thought, the level of evaluation, i.e. the size of the entity being evaluated, was mainly considered through the potential risks involved. Focusing on the risks encouraged widening the scope of thinking, and understanding that even though a pilot setting is not meant for a certain function (e.g., evaluating individuals, or in fact evaluating at all), there is sometimes a possibility to use the setting for unintended purposes. Considering the context raised the pilots' awareness of possible risks involved in evaluation, as well as what their role was in instigating and mitigating these risks. Discipline was considered by the pilots mainly in terms of possible methods to be used in evaluation with the emphasis on how they could be developed to mitigate unfair advantages or disadvantages to do with disciplinary differences, as well as properly reflect thematic characteristics.

Some of the pilots used context considerations to tailor for specific entity sizes and disciplines to achieve better implementation, which naturally requires cooperation with the targets of evaluation. But in general, the emphasis in context consideration, for all of the pilots, was in defining the purpose of evaluation.

6. Options for evaluation

In line with the SCOPE Framework, in the workshops presented in sections 4 and 5, the pilots considered what they value about the thing they wish to evaluate, who or what is the target of evaluation, and why they wish to evaluate it in the first place. Figuring out your values as well as the context supports choosing appropriate methods and tools for the evaluation, which is the third stage of the SCOPE Framework: options for evaluation. The purpose of this stage is to explore all the options - both qualitative and quantitative - for evaluating the identified values in the chosen context. In the OSAF, "options for evaluation" is a part of assessment design, which is the second assessment event phase.

Considering that each evaluation setting is unique, which means that there are many and varied values and contexts, it is impossible to provide a comprehensive list of options for evaluation. However, the OSAF aims to support decision-making by providing recommendations and guidelines on the options.

For the pilots in the context of this project, options for evaluation translate to choosing appropriate tools, services and/or infrastructures developed within GraspOS. The first workshop on options was organised in January 2024 (presented in more detail in section 6.1.), and the second in May 2024 (presented in more detail in section 6.2.)

6.1. Workshopping options, part 1

In the first workshop on options for evaluation, the aim was to facilitate interaction between the services and pilot representatives. Discussions were held on the services' functionalities, on pilot needs in terms of tools, services and infrastructures, as well as on the current tools, services and infrastructures used by the pilots.

The tools, services and infrastructures developed in GraspOS can be divided into three categories: 1) enrichment services, 2) monitoring services and 3) data services. In the workshop, the pilots (divided according to the original categories of national, institutional, and thematic level) would attend three sessions, one for each service category. It was expected that if the pilots are categorised based on the level of evaluation, they would have similar needs in terms of tools, services and infrastructures. However, as already emphasised several times, it has become apparent within the course of this project that evaluation settings are unique in terms of their values and contexts, and therefore in terms of their choice of options for evaluation as well. For this reason it was possible for the pilots to get an initial idea on

which tools, services and/or infrastructures might be useful for their evaluation settings, but an in-depth understanding of how this relates to practice, i.e. how the pilot should move forward in the project towards testing and evaluating the tools, services and infrastructures was not reached.

6.2. Workshopping options, part 2

In the second workshop on options for evaluation, the starting point was different in comparison to the first one: pilots presented their evaluation cases to the GraspOS service representatives, and there were no categorisations of pilots or services - each pilot was considered as a unique evaluation setting. As pre-assignment for the workshop, the pilots were asked to report on the tools, services and/or infrastructure they had identified as relevant for their evaluation setting. In addition, for the purpose of getting ideas flowing ahead of the workshop, they were asked to describe what was needed to make their evaluation setting work. And here so-called blue-sky thinking was welcomed. These reports were made available in advance to the service providers, so all participants, the pilot representatives as well as the service providers, would be prepared.

The pilot presentations were in the form of use cases, which consisted of three elements:

1. User stories, i.e. simple statements that focused on what the pilot wishes to achieve in GraspOS,
2. the pilot's insights on how the tools, services and/or infrastructure developed in GraspOS can support realising the user stories, and
3. the pilot's insights on what is preventing the realisation of the user stories (in the context of the project), i.e. what is conceived to be missing from the tools, services and/or infrastructures developed in GraspOS.

The aim of the workshop was to draft a roadmap for each of the pilots specifying the tools, services and/or infrastructures to be piloted, identifying the relevant partners, i.e. service providers, and the expected outcomes.

The decision not to group the pilots or the services according to the categories in the second workshop proved to be the right one. When the pilots were considered as unique evaluation settings, they as well as the service providers were able to concentrate on the relevant pilot-specific questions that needed to be solved before moving forward to the next stage of the project, the testing and evaluating of tools, services and/or infrastructures.

6.3. Pilot summary on options for evaluation

For this stage, the pilots were asked to describe how the first two stages, 'starting with what you value' and 'considering your context' supported identifying relevant options for evaluating.

In the first place, defining values and context made it easier to recognise the requirements for tools, services and infrastructure to be used in evaluation, as well as prioritising the requirements. Obviously in choosing options for evaluation, having a mandate to implement new tools, services and infrastructure makes it easier, so discussing options was very different for pilots with no such mandate. However, even without a mandate it was considered easier to at least suggest the implementation of new tools, services and/or infrastructure when the choices are based on an idea that they could support commonly accepted values and take the given context into consideration.

So being able to practically discuss different options for evaluation depended a lot on the maturity of the pilot. However, the very premise of the GraspOS project, i.e. the pilots' specific role in testing and evaluating tools, services and/or infrastructure developed within the project, has had an effect on the pilots' maturity in terms of being able to implement tools, services and/or infrastructure: it is likely that the pilots have been quite set on the options they wish to choose right from the beginning. This kind of starting point, from the tools, services and infrastructure available, is opposite to what SCOPE Framework suggests. And it was very interesting to witness how going through the two first stages of the SCOPE Framework slowly turned the tables. The pilots defined what they wish to evaluate, who they wish to evaluate, and why, and based on that process, some pilots even came to the conclusion that none of the tools, services or infrastructure can support their evaluations.

As explained earlier, when planning the second workshop a different point of view was chosen, based on the learnings of the first one. And it turned out, that when the starting point is not the available tools, services and infrastructure and how the pilots attempt to utilise them, but instead the pilots' needs and requirements (based on their individual values and contexts) and how the tools, services and infrastructure try to adapt to them, all pilots were eventually able to identify relevant tools, services and/or infrastructure being developed within GraspOS for them to test and evaluate.

7. Probe deeply

According to the SCOPE Framework, many of the problematic approaches to research evaluation can be avoided if they are 'probed' for harmful impacts and possible unintended consequences in advance. The SCOPE full guide suggests the following four key questions that should be asked of any options for evaluating

1. Who might this discriminate against?
2. How might this be gamed?
3. What might the unintended consequences be?
4. What is the cost-benefit?

In the OSAF 'probe deeply' is part of assessment design, which is the second phase. For examining evaluation options beforehand, the OSAF will provide guides, for example, on RRA obstacles and on equity, diversity and inclusion, as well as a checklist for responsible assessment.

As part of the interim report, pilots were asked to consider the tools, services and/or infrastructure they are planning on piloting from the point of view of unintended consequences or other challenges involved. The four questions presented above were provided to the pilots, but as they might not be relevant for all pilot cases, the pilots were free to probe starting from their own premises. It was also emphasised that as none of the pilots have had any experience from piloting any of the tools, services or infrastructure, the stage should be considered as more of a thought experiment.

7.1. Pilot summary on probe deeply

Most of the pilots used only the four questions SCOPE suggests to probe for harmful impacts and possible unintended consequences. However, there was mention of questions around relevance as well as technical aspects, which are obviously something that need to be probed for in advance for choosing options for evaluation. The majority of pilots reported probing for unintended consequences which included enabling evaluations where it is not intended to happen, misuse of developed indicators, and an unbalanced consideration of merits when a diversity of merits is offered as evaluation material. In addition, there were concerns over losing coverage when implementing a new service, creating a false perception of having to

"tick all the boxes" when expanding profiles with new types of merits and outputs, as well as increasing administrative burden.

The question of discrimination raised nearly as many concerns, mostly on behalf of individuals, but also of disciplines, publishing cultures and different types of outputs and activities. Gaming or cost-benefit were not considered as issues for most of the pilot settings.

8. Conclusions and results

During the past 12 months, the pilots have gone through the four stages of the SCOPE Framework: start with what you value, context considerations, options for evaluation and probe deeply in a series of workshops, both online and face-to-face. The following is a collection of concluding remarks that the WP5 pilots can offer for the further development of OSAF based on their experiences.

- The stakeholder mapping proved to be of pivotal importance for all pilots. It makes sense to keep it as wide as possible to first get the full picture of all stakeholders involved, but maybe the OSAF template could enable drilling down to specific stakeholder groups identified as more important for the design and realisation of a given evaluation.
- The value statement template should acknowledge different starting points, at least the two identified in this report: technical and ideological. While the ideological starting point, i.e. what is it that you wish to evaluate based on what you value, is the original intention of the SCOPE Framework, it is evident that for organisations facilitating and organising evaluations, the values that guide the operationalisation of an evaluation are also important. In addition, the OSAF should see how it could better support the value discussions within thematic communities, specifically addressing the gap between the super-values common to the thematic community and the operationalisation of those values into evaluation designs.
- In context consideration, the pilots mainly focused on defining the purpose of their evaluation. This is already reflected in the OSAF resources, as a designated template for defining purpose will be available there. However, it is important to emphasise that considering the risks involved in evaluation depending on the entity size as well as discipline should not be handled as merely a tick the box in the matrix -exercise and maybe specifically instructing that this risk-mining would certainly benefit from engaging the targets of evaluation.
- When it comes to choosing options for evaluation, it is important to highlight the role of values and context. Especially in the case of a mature evaluation setting, where the

options have already been chosen at some point or another. One solution could be to add something on the premise of the evaluation setting in context consideration, maybe an assessment of maturity in terms of already existing options, and the perceived willingness to let go of them.

- The probing stage should be done consequently with choosing options for evaluation, based either on earlier experience of using the given option, or as a thought experiment. However, it is necessary that the same probing questions are asked also after the evaluation has been conducted, so as a phase of OSAF it could be added also to the final one, assessment evaluation & dissemination.

The pilot specific result of the process described in this deliverable is a practical roadmap describing how the pilot will test and evaluate the tools, services and/or infrastructure it has chosen as options for evaluation. The nine unique roadmaps for testing and evaluating GraspOS tools, services and/or infrastructure, one for each pilot, are included as Annex 3.

9. References

European University Association, Science Europe, European Commission, Stroobants K. (2022). Agreement on reforming research assessment.

International Network Of Research Management Societies-Research Evaluation Group (2023). The SCOPE Framework: A five-stage process for evaluating research responsibly. https://doi.org/10.26188/21919527.v1

Tatum, C., Anli, Z., Waltman, L., Hyrkkänen, A.-K., Pölönen, J., & Nordling, J. (2023). GraspOS Deliverable 2.2 "OSAF". Zenodo https://doi.org/10.5281/zenodo.11091512

10. Annexes

- Annex 1. Template for GraspOS Pilot Assessment Stakeholder Mapping
- Annex 2. Template for Pilot reporting
- Annex 3. Pilot Roadmaps

Annex 1. Template for GraspOS Pilot Assessment Stakeholder Mapping

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

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

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

Determining the purpose of the evaluation:

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

Step 1. Table for collecting stakeholder information

stakeholder name/title	affiliation(s)	role(s) in the assessment	relationship to the outcome	stage of evaluation: planning/conducting /utilizing
		E.g. evaluand(s) assessment design evaluator self evaluation	E.g. decision-making, consulting, following, evaluand(s),	

		coordinator…	evaluator, reporting…	
(example) *Name /* Professor of…	CWTS, Leiden U QSS editor	member of evaluation organizing committee conducted self assessment evaluand	decision-making evaluand reporting	

Step 2. Illustrating stakeholder relationships

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

Annex 2. Template for Pilot reporting

The structure of the interim report is based on the INORMS REG SCOPE Framework (https://doi.org/10.26188/21919527.v1). A brief introduction of each stage will be added to the final report.

The idea of this interim report is to summarize the work the pilots have done in order to be able to start testing and evaluating GraspOS tools, services and infrastructures. The questions are aimed at finding out how the activities, i.e. workshops and their pre-assignments, have supported the pilots in moving forward in the project.

1. Start with what you value

    a. Learnings from the Stakeholder mapping (online workshop 2.10.2023)

    For the workshop, pilots were asked to do a stakeholder mapping. To facilitate doing the mapping, the following questions were presented:

    Start with what you value *about the thing you are evaluating/monitoring*:
    - Who determined what is valued?
    - Was it a collaborative effort?
    - Who wasn't talked to (but might have been relevant)? Why?

    Determining the purpose of the evaluation:
    - Who defines the purpose?
    - What does that imply for the assessment?

    Based on these questions, as well as your stakeholder mapping, please describe how identifying stakeholders, and the different roles they may have, supported your pilot?

    NB! The stakeholder mappings will be used to summarize the type of stakeholders pilots can have, as well as the roles identified, so there's no need to attach them to your report.

    b. Learnings from the Value statement (F2F workshop 22.11.2023)

    For the workshop, pilots were asked to provide a value statement using the SCOPE format of supervalues, values, and (if applicable) sub-values.

Please describe how identifying the pilot value(s) in the form of a value statement supported your pilot?

NB! The value statements will be used in the report, so there's no need to attach them to your report.

2. Context considerations

The meaning of context and purpose in regard to evaluation/monitoring was discussed at the F2F workshop in Espoo on 22.11.2023.

It is important to identify what is the purpose of the evaluation and who (or what) is the target of evaluation prior to considering how to undertake the evaluation as evaluations in some settings have more impact on the entity being evaluated and are therefore a higher risk in terms of consequences. In the matrix below, six different evaluation purposes have been plotted against four different entity sizes to provide a sense of where there may be greater impacts and risks.

Please position your pilot on the matrix below considering your evaluation purpose and the size of the entity being evaluated:

		Country	HEI	Group	Individual
Analysis	To understand				
Advocacy	To show off				
Accountability	To monitor				
Acclaim	To benchmark				
Adaptation	To incentivise				
Allocation	To reward				

Low impact
Medium impact
High impact

Please describe how considering the context of your pilot (entity size, discipline, etc.) supported your pilot?

3. Options for evaluation

   The options for evaluation, which in the context of GraspOS means the tools, services and infrastructure being developed within the project, were presented in more detail to the pilots in the F2F workshop in Leiden on 31.1.2024.

   Were you able to identify tools, services and/or infrastructure relevant for you pilot setting?
   - If yes, which tools, services and/or infrastructure are you planning on piloting?
   - If not, please explain why.

   Please describe how the first two stages, starting with what you value and considering your context supported identifying relevant options for evaluation or monitoring. Not applicable for pilots that did not identify tools, services or infrastructure.
   [*NB! depending on the schedule, the coming workshops could feed into this section of the report, so maybe start with the first two sections until we have more confirmed plans*.]

4. Probe deeply

   For pilots that have identified the tools, services and infrastructures for piloting:
   Please consider the tools, services and/or infrastructure you are planning on piloting from the point of view of unintended consequences or other challenges involved. Keep in mind that this can be done from whatever point of view is relevant for your pilot, so it can be technical, or ideological, or both.

   You can also use the questions suggested in the SCOPE Framework as support:
   1. Who might your evaluation approach discriminate against?
   2. How might your evaluation approach be gamed?
   3. What might the unintended consequences be?
   4. Consider the cost-benefit of the evaluation.

   Please note, that these questions may not be relevant for all pilot cases, and also refer more to the evaluation setting, than just the options for evaluation, so feel free to probe starting from your own premises.

This stage should be considered as more of a thought experiment, as none of the pilots have any actual experience from piloting any of the tools, services or infrastructure.

For pilots that have not identified tools, services or infrastructure for piloting: please describe the challenges preventing you from identifying relevant tools, services or infrastructure for piloting.

Please note, that you can use this space to discuss pilot-internal challenges (e.g., needing a deeper understanding of pilot focus or purpose) or service-related challenges (e.g. not considering any of the services etc. offered within GraspOS relevant), or both.

Annex 3. Pilot Roadmaps

The Annex 3 includes a collection of roadmaps produced by each pilot.

1. Pilots supporting OS-aware RA for funders and national stakeholders

1.1 National funding monitoring platforms

GraspOS pilot – UEFISCDI pilot roadmap

Authors: Ioana Spanache, Ioana Trif, Alina Irimia

Type of pilot: Research Funding Organization, institutional & technological pilot

Short description:

The focus of UEFISCDI's pilot within the GraspOS project is to analyze, develop and test the implementation of the openness researcher profile within our national platforms (e.g. BrainMap), by also aligning with CoARA's Agreement on Reforming Research Assessment, open science practices, and incorporating elements specific to narrative CVs and Openness Profile.

For this we wish to build a test version (a mock-up) of the openness researcher profile, based on the current BrainMap researcher profile to be tested in terms of content, as well as user experience, with representatives of research communities in Romania (researchers, representatives of RPOs, and other).

The **BrainMap** platform is an in-house solution developed by UEFISCDI with over 60000 registered researchers, innovators, technicians and entrepreneurs from all over the world, that provides information regarding over 10 000 R&I projects funded through national and international calls, and contains modules such as Research Outputs Registry, Explore by Skills (based on data extracted from individual profiles), Explore by Map (identifies users according to their countries of provenience), and an Organizations' Registry (in progress). When it comes to the evaluation process, Brainmap is also used as a pool for selecting (international and national) experts to act as reviewers for projects submitted under different funding calls.

The newly designed Researcher profile proposes new information that can be collected in the platform such as new types of research results and activities in accordance with CoARA, related to both open science practices and other types of practices which are related to responsible research assessment, as well as elements specific to narrative CVs and Openness Profile. By expanding the types of research outputs to be included in the researcher profile we will obtain a more comprehensive picture of the contributions a researcher has brought to science and society in general and in this regard we expect to be able to recognize a greater variety of researchers' activities.

In a nutshell - contributions we aim for:

- Researchers are recognized and rewarded for all their contributions to science
- A bigger diversity of types of contributions to and activities of research that are recognized and rewarded in evaluation processes

Goals & links to GraspOS

- test the viability of the tools and services (indicator toolboxes) in accordance with the commitments of the organization's policies, especially related to indicators that cover more types of research products and research activities;
- analyze and test the implementation of the openness researcher profile within national platforms including user experience;
- align with CoARA's Agreement on Reforming Research Assessment;
- experiment and do a feasibility analysis on integration with the OpenAIRE Research Graph.

KPIs

- 1 assessment protocol in OSAR
- 1 Openness profile template
- 100 individual researchers involved in testing the Openness profile template
- 1 national funder
- 2 workshops

Contact persons

- Ioana Spanache, PhD - Policy & Evaluation Specialist UEFISCDI - ioana.spanache@uefiscdi.ro
- Alina Irimia, PhD - Open Science Knowledge Hub Coordinator UEFISCDI - alina.irimia@uefiscdi.ro

Plan for implementation

No.	Activity	Tasks	Expected outputs	Estimated Timeline	Comments
	Analyze and test the implementation of the Openness Researcher Profile within national platforms				
1	Analysis of research	- analyze assessment processes, criteria and	pilot analysis - intended to	April - July	

	assessment processes, criteria and indicators used at national and institutional levels	examples of indicators used at institutional level for certain funding instruments - analyze criteria and indicators used at national level - indicators/ research outputs being used for career advancement of researchers (to principal investigator positions or Associate Professor/ full Professor in academia)	document the proposed Openness profile	2023 (completed)	
2	Review existing literature - relevant documents Openness profile, Open Science, types of research outputs and activities that need to be recognized	- review relevant literature related to concepts such as the Openness profile, responsible research assessment, new types of RA practices and research contributions to be taken into consideration in RA processes, Open Science practices, examples of narrative CVs and other	input for the newly designed Researcher profile template	April - October 2023 (completed)	
3	Analyze institutional platforms and corresponding databases regarding type of data collected and monitored indicators	- analyze institutional databases and types of data collected, including relations between data - analyze current BrainMap research profile specificities and design - analyze functionalities and types of data collected through EvoC - the institutional platform used for contracting, monitoring and reporting of research projects at national level	input for the new Researcher profile template - section dedicated - types of research results with information collected	January - June 2024 (completed)	
4	Elaborate specifications for the Researcher profile template - including	- elaborate a first version of the redesigned Researcher profile (including a dedicated Openness profile section, a comprehensive	1st version of the Researcher profile template	April - June 2024 (completed)	

	Openness profile, a broader diversity of research outputs and activities, and elements of narrative CV	list of research outputs and activities to be included, elements of narrative CV in the introduction) based on the types of data that are already being collected through institutional platforms and are already showcased on the current researcher profile - incorporate a broader diversity of indicators inspired by CoARA commitments and by the OPUS project Research Assessment Framework - identify initial potential sources for data corresponding to indicators included			
5	Receive feedback internally and from the GraspOS team	- receive feedback from GraspOS project partners including in light of the project's services and tools - discuss the template internally - with other departments and management representatives - potential several iterations - adjust the template accordingly	- revised version of the Researcher profile template	June - December 2024	
6	Design a mock up version for the Researcher profile	- design a visual representation of the Researcher profile, including a Openness profile dedicated section	- visual template of the Researcher profile	January - May 2025	
7	Test the template with researchers and research communities at national level,	- elaborate the concept and the format for 2 dedicated workshops - create & send invitations - create event landing page	- 2 organized workshops - 100 consulted researchers - final version	May - November 2025	

	including user experience	and registration page - implement the workshops and consult 100 researchers about the mock up - analyze received input from participants - incorporate received feedback from participants into the template and adjust the researcher openness profile accordingly	of the Researcher profile template		
8	Publish and disseminate the final version of the redesigned Researcher profile template	- make the template available on UEFISCDI webpage as well as on the Open Science Knowledge Hub webpage - disseminate information about it on social media channels and through newsletter to reach larger audiences - publish the template on Zenodo to reach larger audiences	- published version of the Researcher profile template - 1 online article published on UEFISCDI webpage - 1 dedicated message published and disseminated on social media channels - 1 dedicated message disseminated through newsletter - 1 dedicated message disseminated internally	November - December 2025	
	Test GraspOS tools & services				
1	Test the viability of the GraspOS	- Exploring and analyzing GraspOS tools and services		Completed January - July	The opportunities and challenges

	tools and services in accordance with the commitments of the organization's policies, especially related to indicators that cover more types of research products and research activities	that could be relevant for the openness profile (e.g. RAiD, ORCID, Open Citation, BIP! Ranker, BIP! Scholar) to see to what extent it will be possible to implement or integrate some of them for the purpose of the pilot. - API extraction for analyzed services. - Organizing meetings with service providers in order to better understand specific services and tools (e.g. RAiD, Research Graph). - Explore the possibility of using them, especially in relation to the proposed indicators in the new template for the Researcher profile - Test the indicator toolboxes	1 dedicated report that explores potential uses of GraspOS services & tools	2024 September 2024 - November 2025	associated with the possibility of integrating GraspOS tools and services with UEFISCDI platforms are related to both relevance and technical aspects.
2	Experiment and do a feasibility analysis on integration with the OpenAIRE Research Graph	- Analyze the OpenAIRE Research Graph, in terms of types of data covered and functionalities; - attend dedicated community calls or review them online, - attend dedicated meetings & discussions - Elaboration of a feasibility analysis on integration with	1 Feasibility analysis report	February - October 2025	

		the OpenAIRE Research Graph			

1.2 National CRIS (Research.fi)

GraspOS pilot - Research.fi pilot roadmap

Aim: enriching research.fi Open Access publications with citation information				
Task	**Preliminary schedule**	**Contributors**	**Actions**	**Comments**
Preliminary evaluation and planning: • availability of identifiers (in research.fi) • type of identifiers (in research.fi) • querying OpenCitations via API or data dump? • preliminary plan on needed changes to datamodel	09-10/24	CSC OC	• Telecom in September with OC • Internal planning and refining roadmap	Where will the citation data "live"? Will it be a clickable OC logo, or a number (of citations) and then you are directed to OC pages for more information? Or integral part of Research.fi portal? Ways to handle data updates from OC - API or data dump Datamodel updates needed for Research.fi preliminary discussed with OC

PoC on feasibility	10-12/24	CSC	• Testing out how much of Research.fi publication is covered by OC • PoC on technical solution (OC's API or data dump)	What is a "considerable" number? How to assess coverage?
Implementation considerations for research.fi from technical point of view	11/24-02/25	CSC OC for consulting on technical implementation	• Locating the citation information in the research.fi portal, providing design for the implementation within portal	Depending where the data is: if we have it, to the publication information, if not, on the right side (links to other objects, maybe?) Updating the datamodel to enable bringing in new information
Implementation considerations for research.fi from national point of view	11/24-01/25	CSC National Research.fi Steering Group	• Discussion and decision on which publications to include • Discussion on how citations should be handled in research.fi as part of the national publications	4 types of OA → delayed OA not included in the funding model, self-archiving is challenging for referencing, ... Do we want to steer, or celebrate ALL OA, regardless? Which considerations has to be made when providing citation data within research.fi,

				where this kind of information is not yet available - national perspective?
Development of research.fi data model to accommodate citations	12/24-02/25	CSC	• Expansion of current research.fi data model and taking into consideration citations as first-class data entities (https://opencitations.hypotheses.org/816)	Citations as "only" as links between publication or handling them the way OC prefers as first-class data entities https://opencitations.hypotheses.org/816 ?
Technical implementation for importing data from OpenCitations	02-05/25	CSC OC	• Implementation of citation information within the research.fi portal with updates to data model • Implementing automatic update mechanisms for citation information i.e. API or data dump automation	

Aim: enriching researcher profiles with open science activities, merits and outputs				
Task	**Preliminary schedule**	**Contributor**	**Actions**	**Comments**
Preliminary evaluation and decision which plan to choose: • Plan A: WP2 OP as a source of OS activities information • Plan B: OpenAIRE as a source of OS activities information • Plan C: National data as a source of OS activities information • (Plan D: some combination of the above)	09-12/24	CSC CWTS OpenAIRE UEF	• Finding out the feasibility of WP2 OP being able to produce information on OS activities ○ If not, querying OpenAIRE for information filtered for Finland and Finnish organisations ○ If not, using UEF's results on OS indicators found from local and national CRISs • Preliminary plans for needed updates for research.fi datamodel (if open activities need to be explicitly included in data model)	In order to create an openness profile (= a separate new information entity listing OS acitvities, etc.) in research.fi researcher profiles new type of information is needed and possible research.fi data model work as well. The ideal is, that an external service (OP/OpenAIRE) harvests potential information, and research.fi is then integrated to this service. Also, within the context of GraspOS, OP is supposed to define which activities, merits or outputs are considered to be connected to OS. We will follow the development of WP2

				OP on this (DL to be assigned) Keeping an open mind on maybe doing a combination, i.e. several integrations, if it is not too resource intensive in terms of PMs

Implementation considerations for research.fi from technical point of view and preparing a beta site for OS activities within research.fi	12/24-02/25	CSC
		
		CWTS/OpenAIRE for consulting on technical implementation
		
		UEF for consulting on local and national CRIS systems information
			• Preparing a beta site for OS activities within research.fi
			  ○ Decision on information content
			  ○ Decision on design for OS activities within research.fi researchers profile pages
			  ○ Technical plan on data exchange with OP/OpenAIRE
				If not possible, a MVP is a mock site to see what it would look like, and what type of information could be included for researchers to choose from. (vrt. UEFISCDI)
				
				Taking into consideration UEF results on local and national databases - what can be offered as OS that's already in research.fi?

Implementation considerations for research.fi from national point of view	12/24-02/25	CSC	• Collecting feedback from researchers and HEIs on PoC • Workshop for national funders in Finland on OS activities and utilization for this kind of data	How to best collect feedback from individuals and HEIs? Workshop also for funders on openness profile (or the usability of OS related information)
Technical implementation for research.fi and handling data exchange	02-05/25	CSC	• Work on and publication of beta site for OS activities within research.fi • Implementing automatic update mechanisms for OS activities	Decision on moving forward from beta site to OS activities being integral part of research.fi needs to be discussed based on the feedback for PoC and beta site

2. Pilots supporting OS-aware RA at research organisations

2.1 University of Utrecht

23.08.2024

Anestis Amanatidis, Jarno Hoekman, Carolina Castaldi

GraspOS Roadmap

Utrecht University, Copernicus Institute of Sustainable Development

DESCRIPTION OF PILOT CONTEXT

Open science is of strategic importance to Utrecht University. Policies and initiatives by different units (e.g. open science programme) invest substantially in open science. There are open science practices observable, especially with regard to open access publishing, open data and software. Public and stakeholder engagement is promoted as part of open science, too[6].

The engagement of stakeholders in research is a core practice in sustainability research. In the Copernicus Institute of Sustainable Development, it plays an important role, as the collaborations identified during the last SEP evaluation[7] show. At the department, such engagement with stakeholders dominantly goes under the term of transdisciplinary research. For the purpose of this research and the clear parallels to open science policy at the UU, we understand transdisciplinary research as a key open science practice at the department.

In terms of responsible research assessment practices, the Faculty of Geosciences, in which the Copernicus Institute is embedded, is embracing RRA-related policies in their hiring and promotion guidelines and protocols[8]. For instance, Utrecht University expanded the

[6] See Open Science Monitor 2022:
https://www.uu.nl/sites/default/files/230203%20UU%20general%20%20.pdf
[7] See for instance the opinion piece by Ismael Rafols and Louise Bezuidenhout (2024):
https://www.researchprofessionalnews.com/rr-news-europe-views-of-europe-2024-6-judge-open-science-by-its-outcomes-not-its-outputs/ or the UNESCO recommendations on open science:
https://unesdoc.unesco.org/ark:/48223/pf0000379949
[8] See Vision on Recognition and Rewards:
https://www.uu.nl/sites/default/files/UU%20Vision%20Recognition%20and%20Rewards_2023.pdf

MERIT-model with their own TRIPLE model[9], which we had discussed and described in detail in the initial pilot analysis. The transition to more responsible assessment happens in view of adapting tenure and hiring practices to facilitate more diverse and inclusive career paths of researchers in line with CoARA. However, whilst incentives to promote transdisciplinary research exists, and especially so at the Copernicus Institute, there are open questions about the *evaluation* of transdisciplinary research, which is the focus of the GraspOS UU pilot.

Also, between February 2023 and January 2024, the Copernicus Institute has appointed an impact working group with representatives from all its different sections[10]. The goal of this group was to articulate a strategy for 'impact'. The recommendations of this impact strategy currently lead to the institutionalisation of 'impact' in the department through explicit positions ('impact officer'). Crucially for the GraspOS team, 'impact' as posited by the impact working group closely relates to transdisciplinary research and open science, as it is considered a key 'ingredient' for strengthening relationships with societal stakeholders and creating impact.

Currently, the institutionalisation of 'impact' at the department surfaces new evaluative questions: how to evaluate transdisciplinary research practices well? How to keep track of collaborations between scientists and non-scientists? This presents the GraspOS team with an opportunity for intervention, testing new methods for evaluation of transdisciplinary research in a real-life setting.

Following from the SCOPE approach taken and a literature review, there are two important observations for the UU pilot for GraspOS:

- *there is limited knowledge on **how to collect evaluative data** for evaluations that concern transdisciplinary research practices*
- *there is limited knowledge on **what kind of data** fits evaluations of transdisciplinary research, which inherently collapses the separation between excellence 'during' and impact 'after' research.*

By way of following the SCOPE approach, we also found that we need to approach transdisciplinary research and its evaluation by following a process-oriented approach for evaluation in order to move away from a sole focus on research outputs. This was evident for at least two reasons: first, so that it aligns with the needs of those in the department that are concerned with and want to put 'impact' into practice. Secondly, because our focus on the departmental periodic evaluation. That is, because institutionally, this evaluation serves an

[9] See https://www.uu.nl/en/news/from-merit-to-triple

[10] See p.24 in: Laura Himanen. (2023). GraspOS Deliverable 5.1 "Report on pilot setup, current practices and initial requirements". Zenodo. https://doi.org/10.5281/zenodo.11091718

evaluation of the current departmental strategy and as a moment of intervention to re-assess, and possibly re-align it with the values and culture of the department. We think that a process-oriented approach to evaluation thus allows us to cater to these two points.

Another reason for the SCOPE approach is that evaluation methods actively shape what is considered relevant and good research (de Rijcke et al. 2016), which complicates the assessment of transdisciplinary research practices in an already interdisciplinary department and bears responsibility on the GraspOS team to conduct this pilot in the most careful way possible. Especially so if the goal is to contribute to future departmental evaluations.

For us, the commitment to reflexivity and process-orientation makes not only the *research,* but also *the institutional context* a matter of transdisciplinary evaluation. To help us think about process-oriented evaluation of transdisciplinary research, we follow Brenninkmeijer (2022) and distinguish between values, interactions and strategies as units of evaluation. Roughly speaking, **values** describe how researchers and staff express the kinds of 'end goals' they wish to achieve through research through *doing* everyday things; **interactions** describe what new relations are made between different actors[11] during knowledge production processes; and **strategies** describe how values are collectively held and pursued institutionally through, e.g. strategies and guidelines, but also initiatives, or other 'devices for organising order'.

OPERATIONALISING THE CONCEPTUALISATION FOR GRASPOS

In order to test the evaluation of the departmental *interactions, strategies* and *values* for transdisciplinary research at Copernicus Institute, we operationalise these as follows:

Interactions: GraspOS tools and services, in particular the data that underlies the OpenAire Graph, may be part of approaching *interactions*. In particular, we aim at using the data from OpenAire for two major aims that we have with regards to the periodic departmental evaluation of Copernicus Institute for Sustainable Development:

(1) First, to find ways to **identify** – for departmental evaluations – **the collaborations of a particular customised group of individual researchers (teams) and the department as a whole**.

[11] Actors are thought of as a diverse set of people, organisations, projects, but also policies, material things and more that may be affected by new relations (e.g. research that is dependent on limited material resources may, in turn, shape research practices and thus their valuations (e.g. 'telescope time' with astronomers). In the context of transdisciplinary research, however, we focus on collaborations only.

(2) Secondly, to **find ways to identify the type of collaborations** that happen in research teams.

Ideally, these two aims should help the GraspOS team identify a way to assess transdisciplinary research for the next periodic departmental evaluation (2025-2026).

Strategies: a qualitative assessment of the guidelines, strategies, initiatives, existing evaluation protocols and other devices that order and organise the institutional context in which Copernicus Institute operates. This analytical focus of this assessment is transdisciplinary research in view of existing institutional affordances for research(ers). This will be done via document analysis, reliance on previous observations (the working group on impact), and, potentially, further interviews.

Values: a qualitative assessment of the 'valuing' activities (Boenink and Kudina 2020) of researchers at Copernicus Institute. In terms of methods, this draws from already-conducted observations since the start of the GraspOS project and includes observations and interviews of everyday practices and routines of researchers and a description of the culture of the department. Further data gathering may happen in the future depending on the results.

PILOT GOALS

This roadmap relates in following ways to the GraspOS project and the goals set by the GraspOS grant agreement:

- Document how open science is practiced in the context of sustainability research
- A departmental-level analysis and evaluation of open science with a focus on transdisciplinary research
- Test and evaluation of experimental, mixed methods approach to evaluating the stakeholder engagement and transdisciplinary research-aspect of open science
- Develop a template for the evaluation of transdisciplinary research at the level of an department
- Collaborating with the library for potential translation of this departmental evaluation method into other departments within the faculty and other faculties of the university. The background is that the library often provides bibliometric services to inform departmental evaluations and is thus implicated with a wider range of 'evaluation events'.
- Assess how to translate this evaluation to the level of research teams and groups of researchers contributing to similar impact pathways and goals

- Develop general guidelines and learnings, drawing from the specificities of the pilot at Copernicus Institute
- Potential to contributing to GraspOS WP2 Open Science Assessment Framework Indicator Toolbox.

Furthermore, in relation to the UU pilot KPIs, this roadmap lays out previously set performance indicators for the Utrecht University pilot. The following bullet points elaborate on KPIs and changes, if there are any:

- 'Co-develop Open Science assessment protocols at three respective levels.': Assessment protocols will be developed at the level of the department. We will then discuss whether and how these protocols can be translated to other departments within the UU, with a focus on other departments within the Faculty of Geosciences. Moreover we will consider the conditions for adapting assessment protocols to evaluation at lower levels of research teams.

- 'Test the viability of the indicators, tools, and services, particularly how they can inform OS monitoring and narrative CV writing activities as well as their societal impact.': The plan indeed tests indicators, tools and services in view of OS monitoring and evaluation with a focus on societal impact. This may also focus on narrative writing as has been the dominant way of evaluating transdisciplinary research in a previous departmental evaluation, but not in the context of CVs as the lowest level of assessment is the research team and not the individual.

- 'Inform the VSNU Knowledge Base on practices and integration feasibility.': When results are available of this pilot, they will be communicated accordingly.

- 'KPIs 3-5 assessment protocols in OSAR; 20 researchers; 1 research group; 1-2 openness profile templates': It is still unsure to us what assessment protocols in OSAR mean for us. We are engaging more than 20 researchers in a department that spans 5 research groups (sections). The openness profile does not seem to be relevant and applicable to the context of our research.

PLANNING

Generally, we foresee the following steps: First, we will articulate the research and data requirements for OpenAire. This includes what type of datasets are needed and for which period, as well as further extra requirements that relate to customisability of the sample to be represented. This will form the basis for conversations with OpenAire partners about the

assessment and analysis of using OpenAire data to approximate the evaluation of transdisciplinary research practices in this novel way. The focus in this step will be to experiment with different methods to visualise (or list) collaborations between actors that Copernicus Institute is involved in; as well as experiment with methods to determine the type of those collaborations identified.

Secondly, while the first step is ongoing, we will manually curate a dataset with the same requirements. To do this, we will use a variety of methods and combine the data so as to have a representative dataset on which basis we can compare and thus assess the use case we are presenting. To do so, we will first collect all research projects that have been active during the specified periods. Then, we will trace the collaborations that have happened for each project.

The third step will be to analyse the results by comparing the two datasets (step one and step two, respectively). This comparative analysis, we expect, will give us insight into the conditions of use, but also the potential strengths, pitfalls and uses for the OpenAire dataset(s). Comparing the datasets to each other thus will allow us to identify strengths and weaknesses for potential future evaluations. We expect that the datasets will show us where collaborations have happened or are happening, but much less so what kinds of collaborations they are.

This is why, as a fourth step, we want to experiment with the available datasets in view of the question of evaluating *quality* in transdisciplinary research practices. This emphasis on quality comes from the state of the art of the literature on transdisciplinary research evaluation. This serves as a qualitative enrichment of the dataset that was collected and experiments with quali-quantitative approach to transdisciplinary research evaluation. Depending on the results, multiple, related steps could be taken, depend on our assessment of the experimentation:

- Contribute to the WP2 indicator toolbox with particular focus on transdisciplinary research evaluation.
- Put forward a tested evaluation method to capture transdisciplinary research at Copernicus Institute of Sustainable Development for the upcoming departmental evaluation (2025-2026).
- Support the Utrecht University library as a service provider who becomes involved in the evaluation processes of different departments and groups with a focus on translating findings to other departments within the Faculty of Geosciences.
- Strengthen evaluation of transdisciplinary research with a focus on the level of research teams and the four core thematic ‘impact pathways’ by which the Copernicus Institute of Sustainable Development is increasingly becoming structured.

Below, in a table overview, we show a breakdown of the different steps:

Goal	***Timing***	***Description***
1. **Interactions:** Identifying which stakeholders play a role in research projects of Copernicus Institute of Sustainable Development or a custom, inter-faculty team (Fall 2024)		Discuss with OpenAire the identification of interactions with stakeholders. Based on earlier discussions during the Athens workshop, some data options could include: - OpenAire funding data - Author affiliation data - Full-text analyses - Project funding mentions in scientific publications
2. **Interactions:** Triangulating the data from OpenAire with 'internal' data using a bibliographic approach / manual curation (Fall 2024)		This manual collection of data will probably be a mix of different data sources. These include: • Author affiliation data (WoS or Dimensions) • Internal data from research administration and project control • Project funding mentions in papers • Potential links to existing workflows and initiatives (e.g. communication)
3. **Interactions:** Analysis of results (Winter 2024)		*Comparative assessment of the related datasets as to their differences and similarities to determine feasibility of evaluating transdisciplinary research with existing data. Some analytical questions include:* • *What are the use conditions for the data? Or what type is, could be, and is not usable?* • *What are differences between the automated OpenAire-results and the customised dataset?* • *What are the weaknesses of the automated OpenAire-results and the customised dataset?* • *What are the strengths of the automated OpenAire-results and the customised dataset?*
4. **Strategies, values:** Experiment with new approaches to TDR evaluation (Winter		*The literature review revealed that evaluations tend to demarcate the distinction between research quality (as something that happens during research) and impact ('after research'). We want to experiment with assessment methods that dissolve this boundary, which is criticised in current literature on TDR evaluation (see Franssen (2022))*

2024 and Spring 2025)	*with a focus on evaluating* **...the quality of collaborations by:** **Intensity**: 'depth' of involvement (e.g. assessment of stakeholder involvement in the *development* of problematisations) **Openness**: diversity of stakeholders (also other than human ones) mobilised during research process **Quality**: continuity of stakeholders' voiced during the research process **...the content of collaboration by:** **Openness:** diversity of stakeholders represented in the problematisation. **Quality:** role, abilities, concerns assigned to each stakeholder in problematisation **Reflexivity:** awareness of politics of problematisations (they *always* benefit some actors over others).
5. **Experiment with the available datasets** in view of the question of evaluating quality in transdisciplinary research practices. (Spring 2025)	*Using computational methods and locally available tools and services, we want to experiment with new and local approaches to evaluating transdisciplinary research on the departmental level. This will be done in collaboration with the Utrecht University library. These tests include:* - *Customisability of tools and services* - *Using CoreTexT to extract key terms* - *'Issue Mapping' (Marres 2015)* - *And more, depending on the datasets*
6. **Writing up** (Early summer 2025)	*Writing of updated D5.2 Progress report + dissemination*

Table 1: Road map of UU pilot on (tools and services for) evaluating transdisciplinary research

REFERENCES

Siri B. Borlaug, Silje Marie Svartefoss, Henrik Karlstrøm & Thomas Franssen (unpublished). *Evaluating transdisciplinary research quality – a review of the literature.* Private communication.

Boenink, M., & Kudina, O. (2020). Values in responsible research and innovation: from entities to practices. Journal of responsible innovation, 7(3), 450-470. doi.org/10.1080/23299460.2020.1806451

Brenninkmeijer, J. (2022). Achieving societal and academic impacts of research: A comparison of networks, values, and strategies. In Science and Public Policy (Vol. 49, Issue 5, pp. 728–738). Oxford University Press (OUP). https://doi.org/10.1093/scipol/scac022

De Rijcke S., Wouters, P. F., Rushforth, A. D., Franssen, T. P., and Hammarfelt, Bet al. . (2016) 'Evaluation Practices and Effects of Indicator Use—A Literature Review', Research Evaluation, 25: 161–9.

Franssen, T. (2022). Enriching research quality: A proposition for stakeholder heterogeneity. Research Evaluation, 31(3), 311–320. https://doi.org/10.1093/reseval/rvac012

Jorrit P Smit, Laurens K Hessels, The production of scientific and societal value in research evaluation: a review of societal impact assessment methods, *Research Evaluation*, Volume 30, Issue 3, July 2021, Pages 323–335, https://doi.org/10.1093/reseval/rvab002

Langfeldt, L., Nedeva, M., Sörlin, S., & Thomas, D. A. (2020). Co-existing notions of research quality: A framework to study context-specific understandings of good research. Minerva, 58(1), 115-137.

Marres, N. (2015) Why Map Issues? On Controversy Analysis as a Digital Method. Science, Technology & Human Values, 0162243915574602.

Ochsner, M., Hug, S. E., & Daniel, H.-D. (2013). Four types of research in the humanities: Setting the stage for research quality criteria in the humanities. Research Evaluation, 22(2), 79–92. doi:10.1093/reseval/rvs039.

2.2 University of Eastern Finland

GraspOS pilot (affiliated entity to CSC): University of Eastern Finland (UEF) 'Novel methods for responsible research assessment and Open Science evaluation' Roadmap

Short description

Type of pilot: research-performing institutions

RRA maturity: UEF has signed The Agreement on Reforming Research Assessment in 2023. UEF's CoARA Action plan is written and will be published in 2024.

SCOPE approach:

UEF Strategy 2030:

- Value: 'We are courageous, open and responsible.'
- Strategic programme 'Research-based understanding': 'Open science improves the quality and impact of research. The use of research findings by society is intensified and science reaches an increasing number of people. We strive for the utilisation of research-based knowledge from different disciplines in society, as well as for new innovations.'
- Target: 'Strengthening open science, broad utilisation research and science communication'

- Actions: 'Open access publishing; Science communication and social impact communication; Strengthening support for international conferences and science events'

Main data sources:

- UEF CRIS, the UEF research information system, includes information about the publications and other research activities of the UEF research units and researchers.
- Research.fi, a service provided by the Ministry of Education and Culture, collects and disseminates information on research conducted in Finland. CSC - IT Center for Science is responsible for the implementation of the service on behalf of the Ministry of Education and Culture in collaboration with Finnish higher education institutions, research institutes and other research organizations as well as research funders. The service contains e.g. metadata of publications by Finnish organizations. In general, the submission of information is voluntary for organizations. The submission of publication information to the Ministry of Education and Culture has been mandatory for universities and universities of applied sciences since 2012. VIRTA Publication Information Service is an advanced data warehouse solution to integrate institutional data at the national level. VIRTA collates bibliographic information of all scientific publications from institutional Current Research Information Systems (CRIS) and publication repositories. The publication metadata in VIRTA is publicly available in research.fi -service.

UEF open science policy and implementation:

- UEF Open Science and Research Policy
- eRepo, the UEF open institutional repository, is a key component in the UEF open access ideology. eRepo records, publishes and stores permanently the UEF theses,

serials, self-archived articles, and metadata of research data made available by UEF's researchers.

Pilot E.: University of Eastern Finland (UEF)

- Utilizing new indicators and metrics in responsible research assessment and in monitoring open science

Goals & links to GraspOS

- Make a university-level analysis and evaluation of indicators and metrics of publishing activities of the Finnish national service VIRTA / Research.fi
- Test and evaluate less used indicators and metrics of publishing activities
- Monitor UEF's publishing activities and compare them to those of other Finnish universities
- Explore the possibilities to utilize less used indicators and metrics in knowledge management, in impact assessment, in responsible research assessment and in monitoring open science
- Identify the needs of the university (UEF) management regarding the use of less used indicators and metrics of publishing activities in knowledge management, in impact assessment, in responsible research assessment, and in monitoring open science
- Promote and enable open science by exploring and assessing possibilities to utilize less used indicators and metrics of publishing activities e.g., in recognizing merits in open science

KPIs

- 5–10 university leaders and leading experts interviewed/involved
- University and scientific community informed of project outputs

Implementation approach

Steps	Outputs	Tasks	Comments
1	UEF-level: Indicator A: UEF CRIS data	Choosing appropriate units and filters, preparing a new dataset, summative calculations and standardized indicators.	2024 Q2
2	UEF-level: Indicator B: VIRTA data	Choosing appropriate units and filters and preparing a new dataset, summative calculations and standardized indicators.	2024 Q2
3	UEF-level: Hybrid indicator (A, B)	Developing a hybrid indicator model that combines indicators A and B.	2024 Q2
4	National level: Indicator C: OpenAIRE Graph/Explore data	Choosing appropriate units and filters and preparing new dataset based on allowed max. 2000 rows datasets, exploring possibilities for summative calculations and standardized indicators. Exploring and comparing data available from OpenAIRE Explore and from OpenAIRE Graph (API documentation, see https://graph.openaire.eu/docs/data-model/entities/research-product). Observations on OpenAIRE Explore data content and consistency regarding hybrid indicator calculation.	2024 Q2–Q3
5	UEF-level: Hybrid indicator (A, B, A_r)	Testing scalability by adding indicator A_r (proportional activity per recorded researcher) into hybrid indicator model that was developed in Output 3.	2024 Q2–Q3
6	Three Universities level:	Choosing appropriate units and filters and preparing a new dataset, evaluating summative calculations and standardized	2024 Q3

	Indicator B: VIRTA data (Indicator C: OpenAIRE)	indicators. Testing if the institution is suitable for unit/common identifier that receives a hybrid indicator value.	
7	VIRTA data OpenAIRE Explore	Comparing UEF data with selected Finnish universities VIRTA and OpenAIRE Explore data regarding hybrid indicator calculation.	2024 Q3
8	Appraisal of OpenAIRE metadata fields for the assessment of open science	Studying OpenAIRE (and Research.fi) metadata fields for the appraisal of open science at the European level.	2024 Q3 – 2025 Q1
9	Interviews: Evaluation of the hybrid indicator model	UEF leaders and leading experts to be interviewed in order to identify the needs of the university (UEF) management regarding the use of hybrid indicator model and less used indicators and metrics of publishing activities in knowledge management, in impact assessment, in responsible research assessment, and in monitoring open science. Reflection on validity and relevance of the hybrid indicator.	2024 Q4 – 2025 Q1
10	Presentations	Conference presentations on hybrid indicator model (oral presentations and/or poster presentations)	2024 Q4 – 2025 Q4
11	Communication	1. Scientific article on the hybrid indicator model offered to be published in an international academic journal (in English) 2. Article on the hybrid indicator model aimed for the academic community offered to be published in a trade journal (in Finnish)	2024 Q1 – 2025 Q4

		3. Blog post(s) on the GraspOS project, the UEF pilot, and the hybrid indicator model to be published in the UEF library's blog (in Finnish and in English)	
12	Spin-offs	1. Basic requirements for analyzed data and units 2. Notes on how analyzed data could be made more consistent	2024 Q1 - 2025 Q2

Contributors: All tasks are contributed by the UEF pilot team.

2.3 University of Belgrade

GraspOS Chemistry Science UNIBE pilot Roadmap

Short description

Type of pilot: institutional and thematic

RRA maturity: UNIBE is fully compliant with national 'Open Science Platform' and institutional 'Rulebook on Open Science at the University of Belgrade - Faculty of Chemistry' principles.

SCOPE approach: We value Green Open Access practices, participation in OS training, collaboration, quality (research excellence), dissemination, ethical scholarly production and leadership.

GraspOS UNIBE collection in institutional repository: https://cherry.chem.bg.ac.rs/handle/123456789/5989

Existing main data sources and OS policy monitoring:

- https://cherry.chem.bg.ac.rs/ - for publications and other research outputs
- https://cherry.chem.bg.ac.rs/ - for research data
- https://cherry.chem.bg.ac.rs/APP/ - a type of Openness Profile as an internal bibliographic database to browse, search and export information about authors, publications and funding (projects)

- https://open.ac.rs/images/doc/Open-Science-Policy-Serbia.pdf - national Open Science Platform
- https://cherry.chem.bg.ac.rs/handle/123456789/5991 - Rulebook on Open Science at the University of Belgrade - Faculty of Chemistry (UNIBE)

Goals & links to GraspOS

- To develop an assessment protocol for in-departmental career assessment:
  - Establishing badges for researchers
  - Establishing badges for groups of researchers (departments)
  - Establishing badges for participation in Open Science seminars conducted by a librarian
- To enrich a researcher Openness Profile as an add-on to the local repository.
- To integrate information from OpenAIRE Graph, BIP! Scholar and OpenCitations to the local researcher dashboard.
- To evaluate indicators and metrics on how they can improve researcher career assessment.

KPIs

- 1 assessment protocol in OSAF - upgraded Rulebook on Open Science at the UNIBE
- 1 reward system prototype
- 3 badge models for the reward system
- 10 individual researchers involved
- 1 department enrolled
- 1 community of practice
- 2 lectures/workshops

Objectives

- Better assessed OS engagement (Green Open Access publications and attendance at local Open Science seminars) of UNIBE Chemistry Science researchers

Implementation approach

Steps	Outputs	Tasks	Contributors	Comments
UNIBE reward system				
1	Badges for researchers	- Ideation on the logic of badge implementation - Ideation on badge design - Choosing the most appropriate logic and design for researcher badges - Small scale implementation and testing - Small scale validation - Large scale implementation	UNIBE with University of Belgrade Computer Center (RCUB)	Deciding on the logic and appearance of researcher badges and implementation
2	Badges for departments	- Ideation on the logic of badge implementation - Ideation on badge design - Choosing the most appropriate logic and	UNIBE, RCUB	Deciding on the logic and appearance of department badges and implementation

		design for department badges - Small scale implementation and testing - Small scale validation - Large scale implementation		
3	Badges for participation in OS seminars	- Ideation on the logic of badge implementation - Ideation on badge design - Choosing the most appropriate logic and design of badges for seminar attendance - Small scale implementation and testing - Small scale validation - Large scale implementation	UNIBE, RCUB	Deciding on the logic and appearance of badges for seminar attendance and implementation
4	Upgraded Rulebook on OS at UNIBE	- Defining badges for researchers - Defining badges for groups of researchers (departments)	UNIBE, RCUB	Redefinition of the UNIBE institutional Rulebook on OS with the goal of establishing a reward system, contributing to a

		- Defining badges for participation in Open Science seminars conducted by a librarian		later evaluation framework
5	Preliminary statistical report after reward system implementation	- Data collection and visualization - Report draft - Final open report of preliminary data	UNIBE, RCUB	Starting point for long-term sustainability and monitoring
6	Dissemination workshop and community of practice	- Organizing dissemination workshops or lectures and community of practice - Presenting OSAF and metrics to be used for researcher Openness evaluation - Depositing all materials in the institutional repository	UNIBE, RCUB	
Open Aire/GraspOS indicator services embedded in our repository and services				
7	Integration of OpenAire Research Graph indicators services in our repository	- Integrating the Search API for collecting different citation metrics	UNIBE, OpenAire	

8	Integration of BIP!Ranker indicators in our repository	- Integrating the public API for linking records to BIP!Ranker profiles	UNIBE, OpenAire, BIP!Ranker	Linking repository records to BIP!Ranker while data about citation is collected from OpenAIRE Graph due to difference in synchronizations of citation count
9	Integration of OpenCitations indicators in our repository	- Integrating the public API for collecting OpenCitions citation metrics	UNIBE, OpenCitations	

2.4 National Research Council Italy (CNR)

GraspOS Pilot CNR

Short description

In 2022, the National Research Council of Italy (CNR) approved the "Relaunch Plan", which includes a reform of the research assessment system. In November 2022, CNR signed the CoARA agreement; as per its declared commitment, the whole apparatus and processes for assessing the career progressions of CNR researchers and technologists are deemed to change. Consequently, the announced assessment campaign opened in 2023 embodied (to some extent) some core principles and commitments promoted by the Agreement.

We will analyse both the criteria applied in the last competitive call (2020) and the ones instantiated in the new assessment campaign and study how the latter takes a step away from the legacy one. Furthermore, we will run two parallel surveys to probe for the reception of the change in both the evaluators and evaluands, as well as the general sentiment towards the implementation of the Reform at CNR. Finally, the pilot will focus on intersecting the innovations introduced in the latest CNR assessment campaign (narrative CVs, diversity of contributions) with GraspOS-federated data and services to understand how these could facilitate the application and evaluation phases.

Goals & links to GraspOS

- Collect and perform a comparative analysis of the assessment criteria of the two selections for career progressions at CNR.
- Collect feedback from both evaluators and evaluands to understand how the change is perceived and, more broadly, what is the general sentiment about the Reform and its implementation at CNR.
- Experiment with GraspOS federated services and Assessment Portfolios to understand how these can support evaluands and evaluators in performing their tasks during the assessment campaign
- Report back to CNR offices responsible for assessment design.

KPIs

- For the analysis of the criteria, please refer to the table below

Data from the 2020 call: • 7 departments • 27 research areas x 2 possible profiles (senior researcher, director researcher) • 4 research areas for technologists x 2 possible profiles (senior technologist, director technologist) • 61 individual calls (54 + 7) • Positions: 280 senior researchers, 150 director researchers, 70 senior technologists, 20 director technologists	Data from the 2023 call: • 7 departments • 35 research areas x 2 possible profiles (senior researcher, director researcher) • 10 research areas for technologists x 2 possible profiles (senior technologist, director technologist) • 90 individual calls (20 + 20) • Positions: 1010 senior researchers, 180 director researchers, 210 senior technologists, 35 director technologists

Tentative roadmap

Steps	Activity	Tasks	Output(s)	Comments	Timeline
Comparative analysis of assessment criteria					
1	Analysis of 2020 assessment criteria	- Collection of 2020 calls and criteria - Charting general criteria and committee-specific criteria	First working spreadsheet with criteria charted (to be published once the full analysis is over)		Apr 2023 – Sep 2023
2	Analysis of 2023 assessment criteria	- Collection of 2023 calls and criteria - Charting general criteria and committee-specific criteria	Finalised worksheet produced and released. A publication is planned.		Feb 2024 – Oct 2024
3	Comparative analysis	Criteria of the 2023 and 2020 assessment campaigns are compared and confronted with CoARA goals and commitments	We plan to publish a paper with the results and insights from the analysis.		Nov 2024 – Jan 2025
Evaluator/evaluands feedback					
1	Survey preparation	Prepare two surveys, one addressing evaluators and the other assessing evaluands	The two surveys will be produced and released		Sep 2024 – Dec 2024

2	Run survey	Disseminate and collect survey replies		The dissemination of the survey, tentatively scheduled in January, is subject to possible delays in the selection process. All the committees must have finished their work.	Jan 2025– Mar 2025
3	Analysis of the results	Analyse the replies to the survey	The results and the analysis will be released		Apr 2025 – May 2025
Experimentation with Assessment Portfolios, narratives, and GraspOS federated services					
1	Contribute to BIP! Researcher narrative template	Translate the CV template of the latest CNR selection into a digital-twin template in BIP! Scholar that can be used to showcase an assisted/automated compilation of the CV and blend research product portfolio with narratives	A new template reflecting CNR narrative CV in BIP! Researcher		Feb 2024 – Jun 2025 (tentative)
2	Contribute to OpenAIRE Researcher Profile	Contribute to the available visualisations to be shown in	Feedback is included in OpenAIRE Researcher		Jun 2024 – Jun 2025

		OpenAIRE Researcher Profile by taking the inspiration from the CV template of the latest CNR selection	dashboard and visualisations		(tentative)
Report back to CNR offices					
1	Report	Prepare an extensive report about the selection to be handed to CNR staff in charge of the assessment design	Technical report for internal usage		Summer 2025

3. Pilots supporting OS-aware RA for thematic disciplines

3.1 Social Sciences and Humanities Domain

GraspOS SSH Pilot

Short description

SCOPE approach:

Goals & links to GraspOS

- Develop an OS assessment protocol for SSH, engaging with OPERAS members.
- Develop researcher Openness profiles (Assessment Portfolio) for different flavours of SSH and in different functions in an organisation (researcher, publisher, editor)
- Integrate metrics with OPERAS services (metrics and portal/ PRISM).
- Consider including OS certifications by linking to the Skills4EOSC project.

KPIs

- 3-5 Assessment protocols in OSAR
- 1 Generic Researcher Openness/Assessment Portfolio profile
- 10 national representatives
- 2 workshops (1 Stakeholder Exercise – 2 – Review outputs).

Objectives

- The pilot will provide general assessment criteria for Social Sciences and Humanities, considering the specificity of the domain: monographs, OA books, diamond OA journals and infrastructure developed by OPERAS at EU level.

Contact points:

Carol Delmazo (carol.delmazo@operas-eu.org)

Fotis Mystakopoulos (fotis.mystakopoulos@operas-eu.org)

Implementation approach

Steps	Activity	Tasks	Output(s)	Comments	Timeline
Community feedback					
1	Community of Practice	- Discussion on role of Open Science in relation to Research Assessment for the SSH Thematic Area	https://zenodo.org/records/10562683	Co-hosted as part of T6.2 with CWTS	December 2023 - January 2024
2	Consultation Workshop	- Preparation: dissemination, gathering of possible participants - Organisation of three Consultation Workshops - Transcription of the full content **Pending** - Anonymisation - Analysis	https://zenodo.org/records/11071150 **Pending** Publication of transcripts Publication of related articles		February - September 2024

Infrastructure/Services					
3	OpenAIRE Monitor	- Initial discussion - Alpha version - Review internally - Decide on next steps		Monitor for internal use is active. Under evaluation: • Internal review raised some concerns about data quality and data provenance issues. • Considering the option to use the monitor as a Minimum Viable Product (which product will be decided later)	May - October 2024
4	BIP! Scholar (Narrative CVs)	- Prepare the questionnaire - Test BIP! Scholar using volunteer Researchers from SSH fields	- N/A	Feedback from the consultation workshops raised concerns about data-centric profiles. Next steps: Identify volunteers (national nodes from	September - November 2024

				OPERAS) to specifically evaluate BIP! Scholar for SSH	
5	Openness Profiles	- Discuss SSH potential usage	- Potential SSH recommendation/Guidelines	Review potential for SSH Need to propose a meeting with Clifford (need to find a date).	September - December 2024
Open Science Assessment Framework (OSAF)					
6	- Templates	- Narrative Template - Contextual Factors template		Propose meetings - Janne for Contextual Factors	October - December 2024
7	- Guides	- Guidance on the diversity of OS Contributions, roles and activities - Guidelines for evaluands and evaluators - Guidance on translating values, purpose and context into an assessment protocol		Propose meetings	October - December 2024

		- Guidance/Template on what to include, how to document an assessment protocol - Checklist for Responsible Research Assessment			
8	- Infrastructure	- Open Research Information Sources		Here it would be good for OPERAS to create an assessment of the resources and highlight strengths and weaknesses in relation to the activities/outputs/process of SSH Research	October - December 2024

3.2 Agricultural and Veterinary Sciences Domain

GraspOS Agri-Vet Science INRAE pilot Roadmap

Short description

Type of pilot: institutional and thematic

RRA maturity: INRAE is fully compliant with COARA principles

SCOPE approach: we value openness, collaboration, quality (research excellence), interdisciplinarity, societal and economical impact through diverse research activities : scholarly production, training, leadership, dissemination

Denis Tagu, Francoise Boudet-Bône, Camille Brard, Edith Legouy, Frédéric Gaymard. A qualitative and multicriteria assessment of scientists: the case study of INRAE, France. 2022. https://hal.inrae.fr/hal-03890041v1

Existing main data sources and OS policy monitoring:

- https://hal.inrae.fr for publications and other research outputs except data,
- Internal bibliometric database with curated affiliations
- https://data.inrae.fr for research data
- OS Monitoring https://science-ouverte.inrae.fr/fr/la-science-ouverte/le-barometre-de-la-science-ouverte-inrae : OA rate for publications, data and code within publications

Goals & links to GraspOS

- Incorporate new types of indicators to measure OS engagement and uptake.
- Developing an 'Openness Profile' for researchers so as to follow them along their career assessment

KPIs

- 1 assessment protocol in OSAF
- 2-3 Openness profile templates
- 20 individual researchers enrolled
- 2 workshops

Objectives

- Better assessing OS engagement (public, data, code, citizen science) of our researchers in Agricultural Sciences and Veterinary Science
- Analyze the impact in terms of interdisciplinarity, reproducibility of research results and on Society (does openness support innovation or contribute to interdisciplinarity

Implementation approach

To implement this pilot we will use OpenAire services to test them and contribute to their enrichment – A Gantt we'll be done when this table is validated

Steps	Outputs	Tasks	Contributors	Comments
INRAE OpenAire Monitor				
1	An INRAE openaire Gateway with validated data https://inrae.openaire.eu	- Data sources declaration - Data validation by mapping our internal sources with OpenAire graph - Data enrichment if needed ?	INRAE with OpenAire Support	The process to enrich of correct OpenAire Graph should be clarified
2	Thematic classification of INRAE data according Frascati classification and SDG	- Data classification - Classification validation	OpenAire INRAE	Methodology for classification to be discussed
3	INRAE OS monitor on openaire https://monitor.openaire.eu/	- Monitor setting - OS indicators validation by comparison with INRAE open science dashboard	INRAE with OpenAire support	At this stage we will better understand what is feasible or not with openaire then, it will be possible to go further
4	New indicators exploration and implementation	- Analysis of the feasibility of interdisciplinarity indicator - Feasibility of an indicator "contribution to policy" and	CWTS, INRAE, OpenAire	Enrichment of the gateway with indicators at the level of each publication ? Enrichment of the monitor ?

		"contribution to innovation" "citizen engagement" for publications - Data Fairness - reproducibility - Data citations - Implementation and tests		
5	Pilot Report on feasibility and relevance from the RRA point of view	- Report with 3 issues : o technical feasibility o Assessment of Agri-Vet sciences by comparison to other INRAE thematic o Feasibility to extend to thematic monitor on european Agri-Vet Sciences	INRAE	
6	Restitution workshop	- Organization of a workshop to share the lessons learnt	INRAE	Which audience ? internal, external
Open Aire/ GraspOS indicator services embedded in our repositories and services				
7	Integration of OpenAire indicators services in our repositories	- HAL INRAE : metrics on publications - DATA INRAE : metrics on data	INRAE	
8	Openess profile template	- Openness profile design	CWTS, INRAE, OpenAire	

9	Implementation of openness profile with 20 voluntary researchers	- In HAL INRAE i.e. https://cv.hal.science/olivier-le-gall - or in an openaire dedicated service	INRAE, OpenAire	Not clear how to provide this openness profile and help researchers to write their narrative CV
10	Restitution workshop	- workshop organization	INRAE	Which audience ? internal, external

3.3 Computer Science Domain

Computer Science Pilot Roadmap

Short description

Type of pilot: thematic (Computer Science)

SCOPE approach: in the Computer Science thematic pilot, we value Open Science and the fundamental principles of responsible research assessment, while we stress the importance of taking into account researchers' contributions in conferences and workshops and their activities in producing research software. We also point out the intrinsic interdisciplinarity nature of the Computer Science field, which is expected to affect research assessment processes.

More details about the values of the pilot can be found in the following report:

Angelo Di Iorio, Kumar Guha, Silvio Peroni, Laurent Romary, Thanasis Vergoulis. Pilot analysis - Computer Science. 2023. ⟨hal-04362464⟩.

Goals & links to GraspOS

- Develop comprehensive and inclusive templates for CS researcher profiles that will offer recognition for the full spectrum of CS research activities taking into consideration the efforts related to Open Science and the specificities of the field.
- Suggest new types of indicators to assist the evaluation of Open Science engagement and uptake in CS related research by interested stakeholders (e.g., research funding organizations, governmental bodies) .
- Contribute to the design, tuning, and evaluation of GraspOS tools, services, and datasets that will help with the aforementioned goals. More specifically, the initial plan is to work with:
    - Software-Sync and Software-Viz: the tools will be tested to extract software citations and deposits
    - OpenAIRE Researcher Profile & BIP! Scholar: the services will be used to test templates for CS researcher profiles including related evidence, indicators, and narratives
    - OpenAIRE Connect: the tool will be used to provide OS monitoring and analysis capabilities
    - OpenAIRE Graph, OpenCitations Dataset, and BIP! NDR Dataset: the datasets will be used to provide data inputs for researcher profiles and for

statistical analyses. The OpenAIRE Graph brings a variety of useful scholarly metadata and incorporates DBLP which is very important information for the CS domain. OpenCitations Dataset also brings a variety of scholarly metadata and citations. Finally, BIP! NDR Dataset includes citations from conference or workshop papers from the CS domain that do not have a DOI (and are not present in other major citation corpuses).

KPIs

- 1 CS researcher profile template
- 100 CS researcher profiles created
- 1 report on Open Science engagement and uptake in CS

Objectives

- Offer insights and test ideas towards addressing well-known issues in current evaluation processes for CS researchers that usually arise because discipline-agnostic profiles and analyses tend to overlook critical aspects of CS-related research activities and contributions.
- Offer valuable insights on Open Science engagement and uptake in the CS domain.

Tentative Roadmap

Activity	Tasks	Output(s)	Timeline
Supporting activities			
Software mentions	- Testing of the tool and presentation of the results to the research teams' scientific advisors for evaluation of the tool.	Report on the tool's possible usage for evaluation..	Q3 2024
Evaluate the value of citations from CS conferences and workshops	- Investigate the coverage and effect of the citations provided by the BIP! NDR dataset using samples of the CS researchers profiles created.	Report on the coverage and effect of the BIP! NDR dataset to provide data inputs for researcher profiles and for statistical analyses	Q3 2025
Open-Science-aware CS researcher profiles			
Investigation of evidence, indicators,	- Determine a list of appropriate components	- "Gold-standard" list of components, that defines a	Q4 2024

and narratives that can be incorporated in CS researcher profiles with a focus on Open Science.	that a CS researcher profile should include. - Investigate implementation feasibility in the OpenAIRE Researcher Profile and BIP! Scholar leveraging also other tools, services, and datasets from GraspOS.	CS researcher profile template (publish it on the GraspOS catalogs) - List of components that can be implemented in the OpenAIRE Researcher Profile - List of components that can be implemented in BIP! Scholar.	
Support the implementation of CS researcher profile in BIP! Scholar	- Test early demos of the CS researcher profile template in BIP! Scholar providing feedback.	- A subset of the CS researcher profile template is supported in BIP! Scholar.	Q1 2025
Support the implementation of CS researcher profile in the OpenAIRE Researcher Profile	- Test early demos of the CS researcher profile template in the OpenAIRE Researcher Profile providing feedback.	- A subset of the CS researcher profile template is supported in the OpenAIRE Researcher Profile.	Q2 2025
Creation of profiles in both platforms	- Coordinate the creation of profiles from various CS researchers.	- An array of CS researchers profiles ready to be used for evaluation purposes.	Q2 2025
Final evaluation of the profile templates and tools	- Experiment with both profile platforms and identify strengths and weaknesses. - Report the results (to be included in D5.2)	- Results of the evaluation are published in D5.2	Q3 2025
Tracking Open Science (OS) engagement and uptake in CS			
OpenAIRE connect	Study of the use of monitoring indicator for INRIA use-case.	Define the way OpenAIRE connect coud be used for indicator production	Q1 2025