Broadening open science: data affordances for assessing transdisciplinary research

GraspOS Pilot Report of Utrecht University, Copernicus Institute of Sustainable Development at the Faculty of Geosciences

Anestis Amanatidis
Copernicus Institute of Sustainable Development

Jarno Hoekman
Copernicus Institute of Sustainable Development

Arne Hefting
Copernicus Institute of Sustainable Development

Carolina Castaldi - Human Geography and Planning

Jeroen Bosman – Utrecht University Library

11-07-2023 v1.0

This pilot was part of the GraspOS project, which 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.

Table of Contents

List of Table

List of Figures

Division of responsibilities and contributions

This taxonomy is based on the Contributor Roles Taxonomy[1]. We are aware that beyond the five main contributors and authors to this report, a plethora of people contributed to its creation. This includes critical and productive meetings with other pilot representatives from GraspOS, especially under coordination of Laura Himanen and Iiris Liinamaa from CSC in Finland, the thought-provoking consortium meetings of GraspOS, the colleagues at OpenAIRE as well as research managers at Utrecht University, hallway conversations with colleagues at Utrecht University, the administration and the secretariats from both CWTS in Leiden and Copernicus Institute and more.

	Anestis Amanatidis	Arne Hefting	Jarno Hoekman	Carolina Castaldi	Jeroen Bosman
Conceptualization	X		X	X	X
Data curation		X			
Formal analysis	X	X			
Funding acquisition			X	X	
Investigation	X	X			
Methodology	X		X		X
Project administration	X		X		
Software		X			
Resources					
Supervision	X		X	X	
Validation	X	X	X		X
Visualization	X				
Writing – original draft	X	X			
Writing – review & editing	X		X	X	X

[1] See https://credit.niso.org/contact/

1. Introduction

This document serves as a final deliverable concerning the Utrecht University pilot of the GraspOS project conducted at the Copernicus Institute of Sustainable Development (hereafter: Copernicus Institute) at Utrecht University. Primarily, it examines the use of current research information systems for assessing transdisciplinary research at an institute level and with two audiences in mind: research information providers on the one hand, and evaluators who consider their use on the other. Practically, it documents the GraspOS pilot project at the Copernicus institute using the inorm's SCOPE approach[2] and presents the emergent learnings we had throughout our use of this approach.

General concerns with research assessment are well-documented and discussed in a variety of places, including the Leiden Manifesto (Hicks et al. 2015), the San Francisco Declaration on Research Assessment (DORA), the Agreement on Reforming Research, Assessment (CoARA), the Position paper of the national programme for Recognition & Rewards, the Hong Kong principles for assessing researchers (Moher et al. 2020), and the Dutch Strategy Evaluation Protocol (SEP) 2021-2027. These concerns rest on the broad consensus that the way that research(ers) are traditionally assessed do not support healthy research cultures and are not in line with contemporary research practices and attached expectations, for example in view of demands that spring from socioenvironmental pressures. These concerns include the limitations of primarily metrics-based assessments and the unintended consequences they produce (Wilsdon et al. 2015); or the narrow enactments of what counts as research output, for example scholarly papers published in international peer-reviewed journals, which become designed into systems usually deployed for evaluating research - like institutional repositories (Aubert Bonn and Bouter 2023). They also relate to the proprietary nature of information, data and knowledge held by research information companies which makes dominant research assessment practices that rely on these databases for informing their evaluations not transparent (Barcelona Declaration on Open Research Information et al. 2024).

The emerging concerns already led to actions to address these issues through various primary framings. For instance, the Coalition for Reforming Research Assessment is primarily concerned with 'research quality' and 'impact'[3]. The position paper that springs from a collaboration of the Dutch national research funders and knowledge institutions[4] addresses a wider variety of concerns, including the diversification of career paths, recognition of teams as units of evaluation, an emphasis of quality of work rather than maintaining an emphasis on quantities of research outputs, as well as leadership and open science. These framings increasingly highlight the interconnectedness between research assessment and open science and also stress that policy initiatives to incentivise research assessment and open science risk developing in isolation from one another (UNESCO 2021; European Commission 2016).

The GraspOS project can be understood as one initiative that mobilises various action to address concerns regarding dominant forms of research assessments and connect them to open science initiatives. Funded by the European Science Cloud[5], GraspOS is a collaborative project across 18 organisations that aimed at understanding open science and responsible research assessment practices to "develop fit-for-purpose tools and services with the necessary characteristics to accelerate adoption [of Open Science and RRA practices]"[6]. Amongst others, it aims to tinker and

---

[2] See (International Network Of Research Management Societies-Research Evaluation Group 2023)

[3] See CoARA. The agreement, 2022. URL https://coara.eu/agreement/ the-agreement-full-text/. Accessed 02.06.2025.

[4] VSNU, NFU, KNAW, NOW and ZonMw. (2019). Room For Everyone's Talent. Towards a new balance in the recognition and rewards of academics. Accessed 02.06.2025 via https://www.nwo.nl/sites/nwo/files/media-files/2019-Recognition-Rewards-Position-Paper_EN.pdf

[5] GraspOS is funded under HORIZON-INFRA-2022-EOSC-01-01 call.

[6] GraspOS Application Form; p.2

test (largely pre-existing) tools and services for responsible research assessment in the contexts of nine pilots spanning different organisations: from national research funders to university departments and national (research) information portals. These pilots are the sites where evaluative events are to be studied, tested and learned from for the development of tools and services that support the transition to an open science-aware assessment system, whilst adhering to principles of the Coalition for Reforming Research Assessment[7].

Utrecht University assumed the role of one of nine pilot projects. The Utrecht University pilot concentrates on one of four departments of the Faculty of Geosciences, namely the Copernicus Institute of Sustainable Development. The mission of the Copernicus Institute is to develop high quality and relevant knowledge related to sustainable development and to have a significant impact on the transition to a sustainable society. The research is marked by interdisciplinary collaboration and transdisciplinary engagement. It is structured along five multidisciplinary sections, namely Energy and Resources, Environmental Sciences, Environmental Governance, Innovation Studies, and the Urban Futures Studio. The first two have a predominantly science and engineering perspective, while the latter three mostly draw on social sciences. Below we document the findings of this GraspOS pilot, which operated across multiple organisational levels with different evaluative needs. The pilot followed the SCOPE approach that underlie the working structure of all GraspOS pilots, and the report is structured accordingly.

The next chapters will offer methodological reflections on the pilot process, speaking from the point of view of the GraspOS project team at Utrecht University. We first introduce three different organisational contexts in which our inquiry took place. With this contextualisation, we offer qualitative insights of departmental open science (evaluation) by presenting a case study on the departmental evaluation that took place in 2021 covering the period 2014-2020. We also expand on a departmental initiative formed to discuss and promote societal impact of research of the department. As elaborated upon below this then provided an overview of evaluation and monitoring values and demands concerning societal impact creation and transdisciplinarity as a particular matter of concern on the intersection of Open Science and Responsible Research Assessment. In line with this we subsequently present a bifold analytical comparison of research information infrastructures including the UU local CRIS system the CRIS and of OPENAIRE to test the extent to which they can be appropriated for assessing transdisciplinarity. This is complemented by a section that describes hermeneutic additions of researchers' interpretations of themselves and their research group based on the two data sources mentioned before. Finally, we draw conclusions and formulate recommendations.

2. Goals of the pilot

The initial goals of the UU pilot were the following:

(1) Document how Open Science is understood, operationalized, and evaluated in context at Utrecht University

(2) Co-develop Open Science assessment protocols at three respective levels.

(3) 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.

(4) Inform the VSNU Knowledge Base on practices and integration feasibility.

[7] See CoARA. https://coara.eu/

Throughout the piloting period, these goals were refined based on new knowledge that we elaborated in the roadmap document internal to the GraspOS project in September 2024[8]. This resulted in a reformulation of the goals as follows, with slight changes:

(1) Document how Open Science is understood, operationalized, and evaluated in context at Utrecht University with specific attention to OS evaluation on a Departmental level

(2) Co-develop open science assessment recommendations for the university department

(3) Test the viability of OpenAIRE data and the local UU repository (the CRIS), particularly how they can inform OS evaluation on the level of the department

(4) Inform the VSNU Knowledge Base on practices and integration feasibility

These changes were done in consideration of the GraspOS project goals and an assessment of what was being valued within the pilot in line with the first phases of the SCOPE approach (see Chapter 4-6). In particular, based on an analysis of open science assessment values and needs at the Departmental level, it became apparent that an important part of how Open Science is understood within the department related to activities focused on societal impact creation and transdisciplinary ways of working. In line with this and during pilot conduct, the department also put in place a working group for societal impact, with complementary goals to the pilot. We aligned with the working group by asking how societal impact can be evaluated on a departmental level and focused our efforts on (1) providing input for the next departmental evaluation, (2) scrutinising the (publication) data that is normally used for evaluating (open) science. This resulted in lessons for evaluating transdisciplinary research at the departmental level and feedback for the data providers (here: the CRIS and OpenAIRE) as to how they afford the evaluation of more holistic open science, such as transdisciplinary research, and how they can improve their services and tools in line with this.

3. SCOPE methodology

3.1. Working structure

The project team had a working structure where biweekly meetings were used to discuss the progress of the ongoing work and recalibrate between each other. Anestis Amanatidis focused mostly on the execution of the work, where Jarno Hoekman and Carolina Castaldi took up an advisory and supervisory role. Jeroen Bosman joined the team meetings in order to shape the process as an open science expert from the Utrecht University library. During the comparative work, Arne Hefting joined the team to assist the data extraction and analysis. Writing was primarily done by Anestis Amanatidis with inputs and feedback from the team.

3.2. SCOPE and subsequent learnings

The SCOPE methodology (International Network Of Research Management Societies-Research Evaluation Group 2023) was used for structuring the pilot activities. SCOPE stands for (S)tart with what you value, (C)ontext considerations, (O)ptions for evaluation, (P)robing and (E)valuating one's evaluation. Using the SCOPE method as a structuring device for the entirety of the process proved useful in that it allowed us to keep learning and adjusting course as new learnings emerged. While we used SCOPE as intended in the first three stages, we diverted from the SCOPE guideline for Probing and Evaluation. Where according to SCOPE probing is a matter to identify harmful impacts, we are thinking of probing as 'testing' of and 'experimenting with' the options for

[8] See Appendix (10.6)

evaluations that were identified previously with an emphasis on reflexive learning throughout the process. In line with this, we use the 'evaluation' stage to draw conclusions from the entirety of the SCOPE process of 'executing' the pilot, where our evaluation (of the research information systems tested) is considered part of the probing stage. Table 1 broadly summarises the guiding questions per stage and methods and approaches used. Each stage will be described in more detail from the next section onwards.

S	Guiding research question: how does open science emerge as a matter of evaluation at Utrecht University, and the Copernicus Institute in particular? Methods: exploratory interviews, document analysis, informal conversations
C	Guiding research question: how is open science enacted in departmental evaluations at Copernicus Institute? Methods: interviews, document analysis
O	Guiding research question: what GraspOS tools and services afford the evaluation needs of the department and how? Methods: service provider workshop in Athens 30.01. and 31.01.2024 and assessment and justification in the project team based on S and C.
P	Guiding research question: how does OpenAIRE and the CRIS compare for use in departmental evaluations considering transdisciplinary research? Methods: bibliographic comparison, 'data snapshot' interviews
E	Guiding research question: what did we learn and what recommendations can we formulate for the departmental evaluation of Copernicus Institute? Methods: Discussions in the project team and informal conversations

Table 1: SCOPE overview, research questions and methods

4. Start with what you value

This section documents how Open Science is understood, operationalized, and evaluated in context at Utrecht University. It introduces aspects surrounding open science emergence on a university-wide level, how open science is conceived of and assessed on a departmental level, and also introduces the societal impact working group, which is a considerable initiative within the department. We focus specifically on the evaluative concerns each of these aspects raise for research(ers) in the pilot university department.

4.1. Three levels

We encountered three contextual aspects and attached evaluative processes at different levels:

- the *Open Science Programme*, who are concerned with *monitoring and evaluating open science on a university-wide level*
- the *department*, which is subject to the strategy evaluation protocol, a *periodic evaluation*

- the Impact Working Group, a bottom-up initiative that is concerned with evaluating societal impact

These three aspects of open science were primary focus of our initial pilot analysis. We discuss how open science comes to bear on these levels.

On a strategic level of the university, open science is organised in a university-wide *open science programme*. The scope of the programme is broadly conceived of as five pillars that represent the strategic priorities for implementation of open science, including issues surrounding Open Access, FAIR Data and Software, Public Engagement, Open Education, and lastly Recognition and Rewards[9]. Activities are supported by the Open Science office and an Open Science platform which offers guidance and advice regarding Open Science implementation. Within the five strategic priorities, one can find different aspects surrounding openness and formulations of open science ambitions. This includes direct endorsements to national strategies and goals[10] concerning publishing open access; the promotion of reuse and verifiability of data, data management and consideration of FAIR principles when handling data within research; engagement of citizens and other stakeholders whose lives and activities may be affected by knowledge produced at the university; responsible research assessment practices; as well as open education, meaning re-use and findability of teaching materials, but also accessibility to education concerning language or financial barriers (Miedema 2022).

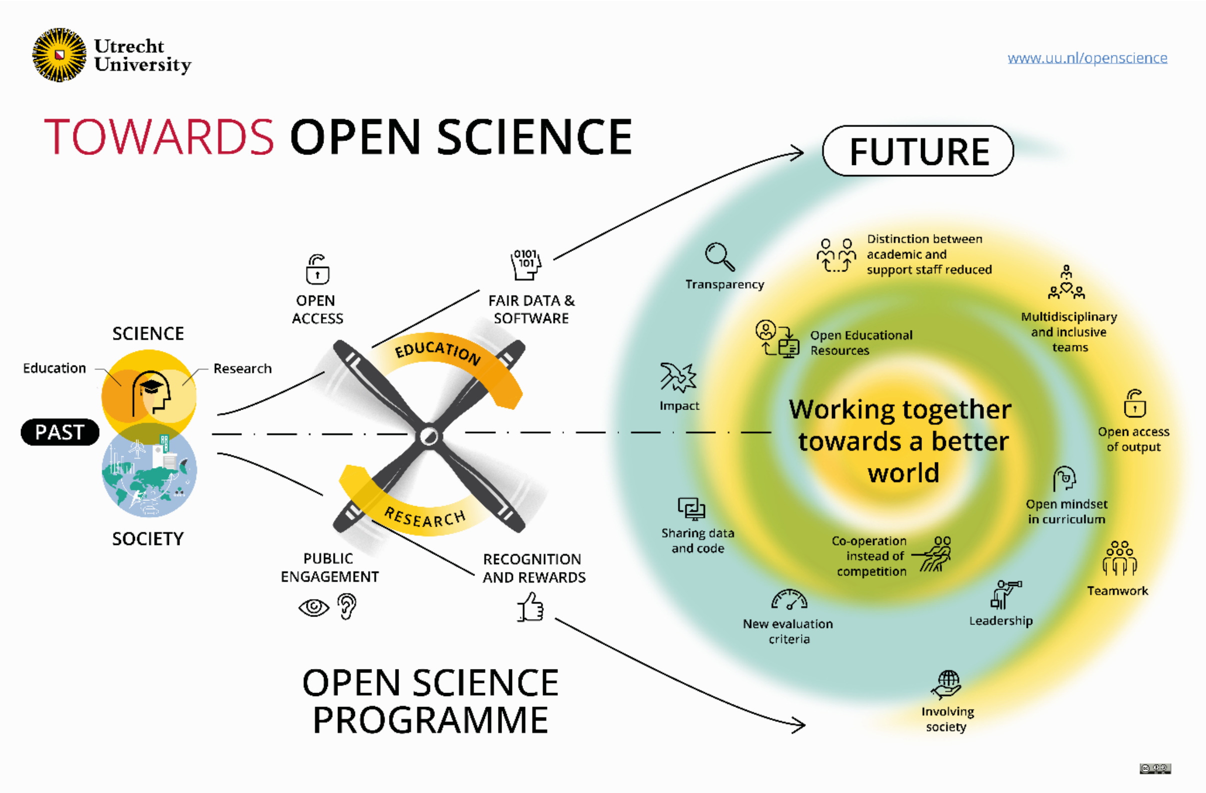

Figure 1: Illustration by Utrecht University on their vision on Open Science

On a *departmental level*, meaning the Copernicus Institute of Sustainable Development, we have analysed how the departmental evaluation of research according to the periodic strategy evaluation

[9] Detailed descriptions about the emergence of the open science programme and descriptions about what these priorities entail can be found on p.16 of this report and the corresponding section overall.
[10] See National Programme on Open Science Strategy 2030: https://www.openscience.nl/sites/open_science/files/media-files/final_npos2030_ambition_document_and_rolling_agenda.pdf

protocol plays out vis-á-vis open science. In particular, we focused on the writing of the self-assessment report covering the period 2014-2020 that forms the basis of the Departmental evaluation in line with the so-called Strategic Evaluation Protocol. In this protocol, open science comes as assessment criterion: ”(a) the extent to which the research unit involves stakeholders, (b) the extent to which the research unit opens up its work to other researchers and societal stakeholders in the context of its strategy and policy. Finally (c), it also considers the extent to which the research unit reuses data where possible, how it stores data according to FAIR principles, how it makes its research data, methods and materials available, as well as when publication are available through open access.” (VSNU, KNAW, and NWO 2020, p.9). It is clear from this description that open science is broadly conceived of and, next to publishing, data and software concerns, includes a focus on engaging societal stakeholders. This explicit reference to societal stakeholders persists throughout and is also formulated as explicit criterion in the protocol, which emphasises under the heading of open science that ‘research units’ are encouraged to evaluate to which extent the department opens up its work to non-academic stakeholders and how it does so exactly. Importantly, this criterion encourages to formulate how such interactions with societal stakeholders will take shape in the strategy of the department. Of course, this emphasis under this heading is interesting in that it diversifies concerns of open research publication that oftentimes accompany notions of open science, especially so in (inter)national strategic documents (VSNU et al. 2019; UNESCO 2021).

Also on the departmental level, but as a less formalised bottom-up initiative, we followed and contributed to the *impact working group*. This group was put together to assess what it means to have societal impact and formulate recommendations to the board of the Copernicus Institute of Sustainable Development concerning societal impact. Whilst the group didn’t form in explicit reference to open science, the group discussed research and teaching in ways that, in other contexts, would be labelled as open science. Moreover, within the context of the aforementioned UU-wide open science programme and Strategic Evaluation Protocol, public and stakeholder engagement are considered a key pillar of open science activities. For us, observing the ’impact working group’ entailed staying sensitive to matters of openness in science that may travel in different forms. Although they were not called in a certain way, we noticed that departmental-level research and teaching practices, values and strategies did align and resonate with open science and its operationalisation in specific strategic ambitions at the university level and within national science system contexts. For the ’impact working group’, some exemplary practices include collaboration with non-academic partners by way of facilitating open spaces for constructive discussion and engagement with wider public, or elimination of output orientation in research assessment and reconsidering hiring practices and adjusting reward and recognition for individual researchers[11]. All of which are, arguably, considered as important aspects of open science (evaluation) elsewhere, particularly when considered in a holistic way. These observations in the impact working group were supported by informal conversations: at the printer, over lunch, at the coffee machine or during after-work gatherings. There, it became clear that the notion of impact is a much more prominent topic of conversation that more ‘traditional’ concerns of open science.

In the end, through iterative loops with the different sections, the group formulated ‘impact pathways’ that address different social groups that the group knew researchers and teachers in the department interact with: policymakers, civic communities, and private sector actors. Figure 2 presents the model that was drawn up and agreed upon depicts a temporality and directionality toward what seems to be a universally-accepted concern across the department in line with its mission: that research and teaching contribute to resolving sustainability challenges around food, land, water and the circular economy. Whirling arrows that represent researchers from different disciplinary backgrounds cut through different social groups towards this goal, indicating collaborations. In general, the model is drawn to depict a relatively high complexity of the

[11] *These “impact practices” used here as examples came from a department-wide workshop aiming at identifying what research and teaching practices are understood as ‘societally impactful’ in December 2023.*

environment in which research and teaching happens. This model also became the working model for the formulation of the societal impact strategy.

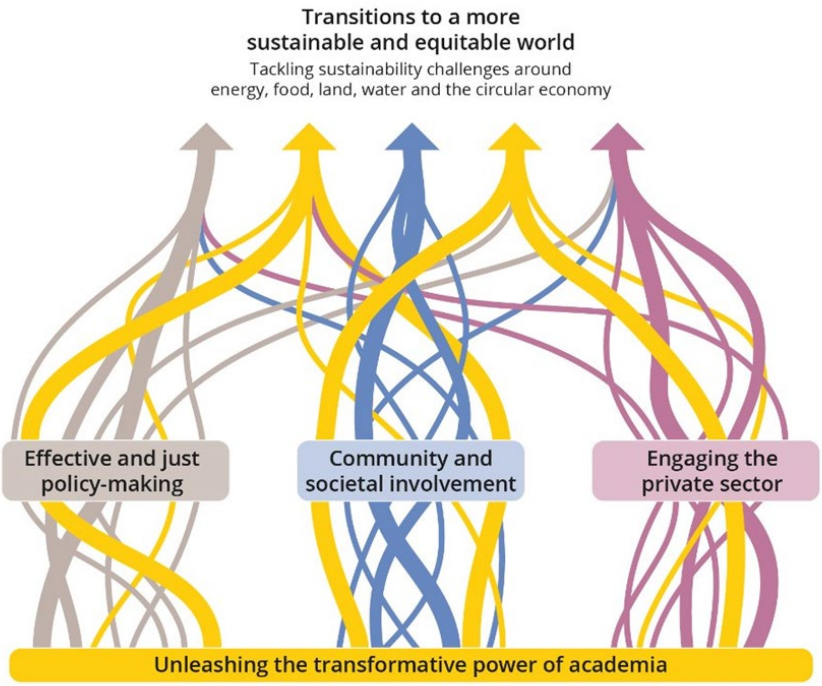

Figure 2: Illustration of societal impact strategy of Copernicus Institute

What is important to note is that these multiple paths represent varying research repertoires that become concerned with extra-academic domains, such as policymaking, public participation, or engagements with private sector actors; on sustainability issues regarding energy, food, land, water or circular economy. As such, from the get-go, the very problematisation of societal impact becomes a stakeholder-inclusive affair which is in line with what is put central as an assessment criterion for Open Science in the Strategic Evaluation Protocol. We also note that the interactions between scientists and societal actors go by different terms in different communities (e.g., public participation and engagement, co-creation, stakeholder engagement, transdisciplinary research). Each of these terms come with their own histories, connotations and differences (Schrögel and Kolleck 2019; Fritz and Binder 2020). In the context of the pilot we decided to speak of 'transdisciplinarity' (i.e. transdisciplinary research). We do so because we noted in informal conversations and the impact working group that academics in the department frequently use *this* term to refer to both existing and emerging practices regarding stakeholder interactions as well as broader aspirations with regard to creating societal impact. Moreover, we also use this term because the literature surrounding transdisciplinarity explicitly engages with questions surrounding reflexive learning and experimenting with expansion of existing research repertoires.

Table 2 summarises the key differences and similarities between the three levels:

	OS Monitoring (Open Science Programme)	SEP Evaluation	Societal impact working group
Key Evaluation Carrier	Questionnaire Knowledge from representatives of the Open Science Platform	Self-assessment report	Series of meetings + department-wide workshop
Level, Purpose	University-wide, to understand and monitor	Departmental, To monitor (both learning and accountability)	Departmental and faculty, to improve
Collaborative science-society encounters as	Stakeholder engagement	Societal impact	Transdisciplinarity
Ideas of open science	Clear idea of openness, narrow (e.g. focused on digital infrastructures for research, stimulation and uptake through policies and support)	Mixed ideas between open science and other forms of openness (e.g. societal impact). Often occurs as 'open access'.	Issues of openness, without speaking of them as issues of open science (societal engagement, translation of research outputs)
Role of publication data	Used to evaluate the share of open access articles	Used to exemplify departmental outputs, including share of OA	None
Period	Was biyearly, now discontinued as focus groups are now used	Six-yearly	Single action
Use of research information infrastructure	CRIS, DOAJ, Unpaywall, SAP/BI for OA Monitor	Bibliometric analysis, CRIS (The CRIS)	No

Table 2: Evaluative events that the UU-GraspOS team has identified and inquired

4.2. Conclusions for GraspOS Pilot

The parallel inquiry into the three levels during the first year of GraspOS showed that formalised strategic activities at the open science programme feature 'open science' prominently and treat it largely as a set of new research practices to become implemented in everyday routines of researchers through a combined approach between top-down policies and bottom-focused support services, such as recognition and reward structures[12] as well as the exchange of learnings across university structures. These include implementation measures and learnings on both publishing practices (open access) and broader research practices (stakeholder engagement).

Within the Department, there is little explicit reference to 'open science' as a term with stabilised meanings and practices. However, its broader ambitions which emphasise engagement of societal actors in knowledge production feature prominently in the strategy evaluation protocol and the concerns of the Department focused on societal impact and transdisciplinarity.

Based on these observations we present three key learnings:

Firstly, the analysis of what is being valued showed that an implicit ambition of open science as enacted at the department and university more broadly indeed concerns the production of knowledge with societal actors. In particular, 'transdisciplinary research' features as a prominent process-oriented approach for stakeholder engagement and societal impact creation.

[12] See Open Science Monitor from Utrecht University, which has surveyed the state of open science twice so far (2020, 2022).

Secondly, and in effect, this concerns both bottom-up institutionalisation and top-down implementations of 'open science' which we consider as equally important empirically for what concerns 'open science at UU' as both carry normative implications on how research should be done, recognised and rewarded - albeit being enacted differently. Such symmetrical treatment allows us to recognise policies, strategies, and other formalised initiatives emerging from executive level of the university as equally meaningful for GraspOS as everyday, mundane research practices that we observe in the department.

Finally, our situatedness within the department called for conducting this analysis in the context of decision-making and decision-makers at the departmental level. As a consequence, we refined the pilot aims based on this step as to making the mentioned points regarding broadly defining OS and acknowledging top-down and bottom-up relevant in the context of the departmental evaluation.

In terms of evaluation, we also understand our observations as signifying a knowledge gap in evaluating stakeholder engagement as open science practices proliferate, although high-level reflections are already being articulated for open science (Rafols, Meijer, and Molas-Gallart 2024).

This observation and the nascent institutionalisation of transdisciplinary research also surfaced new evaluative concerns: How to make transdisciplinary research practices visible for evaluations? And how to do so on a *departmental* level specifically?

5. Context

The preceding analysis emphasised a knowledge gap that we wanted to address. Namely to understand how open science as a matter of transdisciplinarity is evaluated departmentally. For informing the context in which this evaluation takes place, we analysed more closely how the departmental periodic evaluation attends to such reading of open science.

The departmental evaluation is underpinned by the national strategy evaluation protocol of the Netherlands. The assessment process, as well as the criteria, follow the protocol that is set by VSNU[13] – a coalition across 14 Dutch universities with an own legal organisational structure and 45 staff who represent the Dutch universities. This 'Strategy Evaluation Protocol' is renegotiated periodically and explicitly complies with and refers to DORA[14], which indicates the interconnectedness to global movements in research evaluation.

Figure 3: Cover of the SEP evaluation protocol 2021-2027

The protocol outlines three broad assessment criteria: *research quality, viability,* as well as *societal relevance.* Next to these broad assessment criteria, the SEP protocol specifies four aspects of research culture that the evaluation should address, *open science, PhD Policy and Training, Academic Culture,* and *Human Resources Policy.*

[13] 'Universities of the Netherlands' (UNL) in English. See https://www.universiteitenvannederland.nl

[14] See San Francisco Declaration on Reforming Research Assessment. https://sfdora.org/. Accessed 06.07.2023.

As such, in the last SEP evaluation, open science was on the agenda as (at least) one of four criteria for conducting Copernicus Institute's evaluation. A stronger strategic focus on open science had also been recommended by the evaluation committee in the prior evaluation which prompted the institute to put in place vehicles to promote open science practices among the institute's researchers.

This resulted in several initiatives, including the development of research hubs with the purpose of connecting university researchers with stakeholders and data stewards to support researchers with open—source data sharing and open access publishing, interlinking research assessment with open science on a departmental level.

In our analysis, we understood the SEP evaluation as an event where values surrounding openness, but also the Copernicus Institute's values come to bear. The last SEP evaluation was concluded in 2021 (for the period 2014-2020). To illuminate the process around which the SEP evaluation was conducted, we focused on the categories used to describe concerns of open science and issues that emerged in the writing 'SEP evaluation self-assessment report', which underlies the evaluation and is then complemented by the visit and external evaluation of a committee. This allowed us to develop the GraspOS value proposition at Utrecht University as one that attends to a specific evaluative event. Hence, we could reflect on how evaluative knowledge was gathered to account well for open science. This is particularly interesting as the last SEP evaluation was the first one to use the new evaluation protocol, which made the job for the evaluation committee and the self-assessment writing team particularly challenging as there was no inspiration to be drawn from by any preceding example.

Open Science[3]:
The assessment committee considers the extent to which the research unit involves stakeholders, if possible and relevant, in the preparation and execution of the aims and strategy. It also considers to which extent the research unit opens up its work to other researchers and societal stakeholders in the context of its strategy and policy. Furthermore, the committee considers whether the research unit reuses data where possible; how it stores the research data according to the FAIR[4] principles; how it makes its research data, methods and materials available; and when publications are available through open access. Even if Open Science was not yet considered by the research unit for the past period, the assessment committee evaluates the unit's considerations and plans for the future with regard to Open Science.

In the self-evaluation, the research unit reflects on how it involves stakeholders, to which extent the research unit opens up its work to other researchers and societal stakeholders, how it pays attention to other aspects of open science and what its future plans are in this respect.

Figure 4: Open Science criterion description from SEP evaluation

In general, the self-evaluation is expected to be a coherent narrative argument reflecting on the aims and the strategy of the department. The idea is that the narrative argument is "supported by factual evidence" (p.19) and, "where appropriate", quantitative indicators (ibid.). The protocol stresses that the choice of indicators should be based on the argument which the self-assessment report wants to develop, rather than enrolling such indicators irreflexively. Then, the protocol also suggests benchmarking against peer research units and presenting case studies to highlight distinctive and societally relevant accomplishments. This should then be followed by strategic implications for the future of the department and through the use of an analysis of the strengths, weaknesses, opportunities and threats that exist in the institutional context of the department.

As such, the self-assessment report can be understood as the assembly of (evaluative) choices made by both the Board of Directors of Copernicus Institute, the team who supported the preparation and the conduct of the process, but also the technical affordances offered by both the research information infrastructures that were available to the composition of the self-assessment report and the technical ability and advice of the Utrecht University Library who supported the process. In view of open science specifically, the protocol suggests the evaluation to reflect on three key themes:

(a) the extent to which the research unit involves stakeholders,

(b) the extent to which the research unit opens up its work to other researchers and societal stakeholders in the context of its strategy and policy.

(c), it also considers the extent to which the research unit reuses data where possible, how it stores data according to FAIR principles, how it makes its research data, methods and materials available, as well as when publication are available through open access.

By examining the self-evaluation document and through interviews with those involved in compiling the report we created an overview of the evaluative aspects which were employed and that indicate associations to open science as presented either explicitly or implicitly. Table 3 outlines categories that were used, and indicators that fall under these categories, and what data source and/or form these evidencing practices took. A detailed table can be found in appendix 10.1.

Research products for societal target groups		*the CRIS*
	Published policy reports (national, European, global)	
	Visits on Copernicus Institute website	
	Examples of (research) interaction with stakeholders	*Form; short text*
	Example of research output for different audiences	*Form; short text*
Use of research products by societal target groups		
	News mentions related to publications	*Altmetrics*
	Newspaper items (national, international)	
	Radio appearances	
	Television appearances	
	Other news items	
	Mention of publications in Wikipedia articles	
	Mentions of publications in blogs	
	Social media mentions (Facebook and Twitter)	
	Mentions in policy documents	
	Publications co-financed by ministries in NL and abroad	
	Publications co-financed by the European Commission	
	Examples of co-creation, knowledge networks and platforms	*Narrative*
	Examples of use of research products by societal target groups (international, national)	*Narrative*
Marks of recognition from societal target groups		*Form*
Examples of marks of recognition by societal target groups (films, presentations, committee membership, etc.)		*Form*

Table 3: Indicators used in association to open science in the self-assessment report

Whilst this is only a selection of quantitative and qualitative indicators that were employed, there are a couple of observations that deserve explicit mentioning.

One is that an explicit effort was made to contextualise the knowledge that was presented in the form of research data. For example, the self-assessment report outlines a column next to every indicator that elaborates in short descriptions what actually is evidenced by indicator. This speaks to the limits, or wrong interpretations of quantitative and qualitative indicators.

Another observation is that primarily, evidence of both academic and non-academic achievements usually come in the form of *objects* of research or collaborations. Some examples include presentations, papers, or policy documents. Some indicators, such as *examples of use*, or *examples of interaction* usually take the form of narrative descriptions. These are exceptions to the dominant use of objects for evidencing.

From our interviews with a member of the writing team of the self-assessment report, the local university repository was only of little help in evidencing as it provided only little data that they felt could be appropriated for the self-assessment report. This prompted them to manually collect case descriptions of projects or other interactions by directly communicating with the department's staff and asking them to send in descriptions. Other data sources included the Human Resources Management system for current staff data and also research information providers that were used by the Utrecht University Library to conduct a bibliometric analysis.

Importantly, the case descriptions of colleagues resulted in the development of a number of case-studies showcasing the activities of colleagues focused on various sustainability challenges. These case studies are written as narratives of approximately 1.5 page and including key references to research outputs. As mentioned in the self-evaluation report they are meant to showcase "how rigorous research has proven to be highly relevant for societal partners" and illustrate how impact has been achieved on various topics such as "global and European policies related to sustainable development" and "day-to-day policy processes related to circular economy". It is also explicitly mentioned that these forms of impact are hard to grasp: "most of these high impact processes are not highlighted in the media and therefore less visible. However, it is these types of impacts that make us most proud".

In sum, what is striking is that 'societal impact' is reiterated extensively in the text of the self-assessment report, which in itself indicates the relative strategic importance of boundary-spanning work between science and society for the department. Coupled with the observation that collecting evidence systematically for evaluation of such work turned out to be more difficult than anticipated, the question that may result from the SEP evaluation for GraspOS is one of sourcing evaluative data: What GraspOS tools and services afford the evaluation needs of the department and how?

6. **Options for evaluation**

The preceding two analyses showed us that assessing how GraspOS can support the evaluation of transdisciplinary research on a departmental level would be beneficial as evaluative knowledge on such interactions was difficult to get a hold on and demanded improvised methods for collecting evidence by the writing team. This provides the GraspOS with an opportunity for intervention. Importantly, the service providers internal to GraspOS was the extent of tools and services that could be considered, which on the one hand allowed for collaboration on our issue at hand; but also limited us in the sense that other, 'external' services and tools could not be considered. We therefore took the repertoire of tools and services that GraspOS providers offered as options for evaluation assessed and justified what tool or service provider fits our pilot purpose based on the assessment of what is valued (S) as well as the context (O). The table below summarises the process of assessment and justification in the project team:

Value proposition	**GraspOS Tools and Services**	**Justification**
Enrichment of research outputs with missing attributes	OpenAIRE Broker, OpenAIRE text mining modules, OpenAIRE Metadata Validator, SCRE Pipeline	The focus on transdisciplinary research focused on *interactions* and there were no sufficient existing outputs that indicated transdisciplinary research
Enrichment of research output links	Semantic Citation Classifier, BIP! Citation classifier	Focus on academic citations only (Semantic Citation Classifier), focused on computer science (BIP!)
Enrichment of research outputs with novel metrics	EC KIP OS indicators, BIP! Services (toolbox)	Potentially interesting to link outputs to non-academic entities. However, the services focused either on academic indicators (BIP!), or focused on research products (EC KIP OS Indicators)
Enrichment of research outputs with usage data	EOSC accounting for services, UsageCounts, OPERAS Metrics	Impact metrics signifying impact in academic field (EOSC accounting for services), focus on scientific data sources (UsageCounts), focused on monographs and books (OPERAS)
Dashboarding	OS Institutional Dashboard (OpenAIRE), EOSC OS Researcher Dashboard (BIP!Scholar)	Potentially interesting, but focused on research output of researchers (both)
Scholarly resources	OpenAIRE Graph, OpenCitations, Scholexplorer, BIP! DB, OpenAIRE Usage Counts, OPERAS Metrics, GoTriple Platform, FAIRCORE4EOSC RAiD	Diverse applications. OpenAIRE Graph seemed like a promising service to analyse in order to understand the ‘base’ data underlying OpenAIRE services. Other services were field specific (e.g. OPERAS), focused on citations (e.g. OpenCitations), etc.

Table 4: GraspOS tools and services in relation to UU pilot

Based on this exercise, working with OpenAIRE Graph data seemed to be the most viable and useful for our purposes, as it allowed us to assess how the basic data that OpenAIRE Graph can be appropriated for evaluations of transdisciplinary research. Also, as the context considerations showed that a lack of evaluation data was underlying the departmental evaluation previously, we decided to focus on a collaboration with OpenAIRE in order to analyse to what extent OpenAIRE can enrich the already-existing data in the local repository of Utrecht University.

We found that particularly interesting as the strategic and scientific importance of transdisciplinary research further blurs the distinction between science and society. For us, this also means that infrastructures that cater to research evaluation increasingly need to cater to the evaluation of research practices that, oftentimes, remain invisible in traditional research information as noted in the self-evaluation report of the Copernicus Institute as well.

Comparison between the CRIS and OpenAIRE data

In order to assess OpenAIRE data in that regard, we chose to conduct a comparison between OpenAIRE data and data from the CRIS. The CRIS usually provides a baseline of research information that Utrecht University has available. As such, the comparison compares externally generated and organised data with internally generated and coordinated data. The comparison is based on two criteria: the *ideal extent* to which both data sources can inform transdisciplinary research evaluation, and the *real extent* respectively. This difference emerged after receiving the data corpuses and realising that relatively many columns (i.e. monitoring functions) were insufficiently populated with data of Copernicus Institute or not at all. This produced a distinction between *what could be assessed if data were existent (potential data)* and *what can be assessed given the data that exists (actual data).*

CRIS and OpenAIRE data affordances for evaluating transdisciplinary research

To evaluate how the real data can potentially afford evaluations of transdisciplinary research, we drew up a 'table of interactions' that was also sent to OpenAIRE colleagues in order to prompt a first brainstorm about the kind of data that the Utrecht University pilot needed and to prompt reflections on potential data limitations. For this end, we had asked the OpenAIRE colleagues to fill in a column with notes which reflect on the limitations for monitoring and capturing data or respective interactions (rows). Unfortunately, these reflections and notes had not been filled in when received the document again.

The left column describes *Interactions* that we observe mattered in transdisciplinary research. The second column elaborated on the 'entities' that are involved in every particular interaction, such as persons, outputs, datasets, or other potentially computationally capturable entity. Finally, the right column describes potential publication-based data that we knew exists in the realm of GraspOS.

includes academic events (conferences, symposia..), peer-reviewed publications, pre-prints, reviews...
*includes non-academic events (civic, artistic, governmental), policy briefs, tv-appearances, newspaper articles...

iteractions	entities	data
cad. co-authorship*	authors COP, authors other UU, output	publication data, with author information and author affiliations, ORCID, Authors PIDs
cad. co-authorship*	authors COP, authors non-UU acad affil., output	publication data, with author information and author affiliations
cad. co-authorship*	**authors COP, authors non-acad affil, output**	**publication data, with author information and author affiliations**
hare non-acad. co-authorship**	authors COP, authors other UU, output	publication data, with author information and author affiliations
hare non-acad. co-authorship**	authors COP, authors non-UU acad affil., output	publication data, with author information and author affiliations
hare non-acad. co-authorship**	**authors COP, authors non-acad affil, output**	**publication data, with author information and author affiliations**
hare research project (w. grant)	persons COP, persons other UU, project grant, grant agreement	project/grant data, with consortium details and/or participant data, acknowledgements publications
hare research project (w. grant)	authors COP, persons non-UU acad affil., project grant, grant agreement	project/grant data, with consortium details and/or participant data, acknowledgements publications
hare research project (w. grant)	**persons COP, persons non-acad affil., project grant, grant agreement**	**project/grant data, with consortium details and/or participant data, acknowledgements publications**
iteracted in research project (as in: not a formal project	person COP, persons non-UU academic affiliation	final project reports, mentioned in publications, evaluation reports, media coverage
iteracted in research project (as in: not a formal project	**persons COP, persons non-acad affil**	**final project reports, mentioned in publications, evaluation reports, media coverage**
hared (creating and giving a) presentation	persons COP, authors other UU, presentation, event	data on e.g. conferences (abstract submissions, proceedings) / media coverage of shared presentations/events
hared (creating and giving a) presentation	persons COP, authors non-UU, presentation, event	data on e.g. conferences (abstract submissions, proceedings) / media coverage of shared presentations/events
hared (creating and giving a) presentation	**persons COP, authors non-academic affil., presentation, event**	**data on e.g. conferences (abstract submissions, proceedings) / media coverage of shared presentations/events**
ive presentation to non-academic receivers	**person COP, organising entity, presentation, event**	**data on e.g. conferences (abstract submissions, proceedings) / media coverage of shared presentations/events**
iteracting expertise	persons COP, persons other UU, problem statement(s), advisory board, governmenta	grant agreements, other funding descriptions (CORDIS)
iteracting expertise	persons COP, persons non-UU acad affil, problem statement(s), advisory board, gov	grant agreements, other funding descriptions (CORDIS)
iteracting expertise	**persons COP, persons non-acad affil., problem statement(s), advisory board, govern**	**grant agreements, other funding descriptions (CORDIS)**
aming a problem	persons COP, persons other UU, problem statement(s)	grant agreements, other funding descriptions (CORDIS)
aming a problem	persons COP, persons non-UU acad affil, problem statement(s)	grant agreements, other funding descriptions (CORDIS)
aming a problem	**persons COP, persons non-acad affil., problem statement(s)**	**grant agreements, other funding descriptions (CORDIS)**
rovide information to stakeholders	**persons COP, persons non-UU**	**(internal) data on transfer of IP, data or insights**
takeholders provide information	**persons COP, persons non-UU**	**(internal) data on transfer of IP, data or insights**
ieaningfully mentioned together in (full text) of output	COP (org), academic organisation, output	co-word analysis of full-text academic and non-academic publications
ieaningfully mentioned together in (full text) of output	**COP (org), non-academic organisation or person, output**	**co-word analysis of full-text academic and non-academic publications**
ieaningfully cited together in academic output	person non-UU, persons COP, output	academic publication data, with author information and author affiliations
ieaningfully cited together in non-academic output	**person non-UU, persons COP, output**	**non-academic publications, with author information and author affiliations**
iention (full-text) in acad. output	authors COP, authors other UU and v.v.	full-text publication data, ideally contextualised
iention (full-text) in acad. output	authors COP, authors non-UU acad affil and v.v.	full-text publication data, ideally contextualised
iention (full-text) in acad. output	**authors COP, authors non-acad affil and v.v.**	**full-text publication data, ideally contextualised**
iention (full-text) in non-academic output**	authors COP, authors other UU and v.v.	full text publication data, ideally with data on mentions of (output/outcomes) research of COP or its members
iention (full-text) in non-academic output**	authors COP, authors non-UU acad affil and v.v.	full text publication data, ideally with data on mentions of (output/outcomes) research of COP or its members
iention (full-text) in non-academic output**	**authors COP, authors non-acad affil and v.v.**	**full text publication data, ideally with data on mentions of (output/outcomes) research of COP or its members**
eceived funding	COP, org, public research funder, project	project funding data (e.g. CORDIS, OpenAire, National Research Funders)
eceived funding	persons (COP), public research funder, project	project funding data with participant information (e.g. grant agreements)
eceived funding	COP (org), private research funder, project	funding acknoweldgements, private funder websites
eceived funding	persons (COP), private research funder, project	funding acknoweldgements, private funder websites, researcher profiles

Table 5: List of interactions. Note: boldly marked rows correspond to transdisciplinary processes

Hermeneutic additions to data snapshots

Next to this bibliographic comparison, we collaborated with the newly instituted impact team of the department and elaborate on a questionnaire that we sent collaboratively to the department. This questionnaire asked researchers to reflect on how they believe their work makes a difference. After the responses came in, we invited a selection of these researchers into a conversation about their research data that we sourced from the CRIS and OpenAIRE and invited them to reflect on their previously formulated 'impact statements' by extending the printed sheet with a network visualisation of their research data with what they deem important. The assumption that we rested this exercise on is that we thought that such addition can provide insights to the kinds of entrenched categories by which research information is conceived and that this would be important feedback to the GraspOS partners, especially in view of emergent research practices. The triangle in Figure 5 visualizes and summarises our operationalisation.

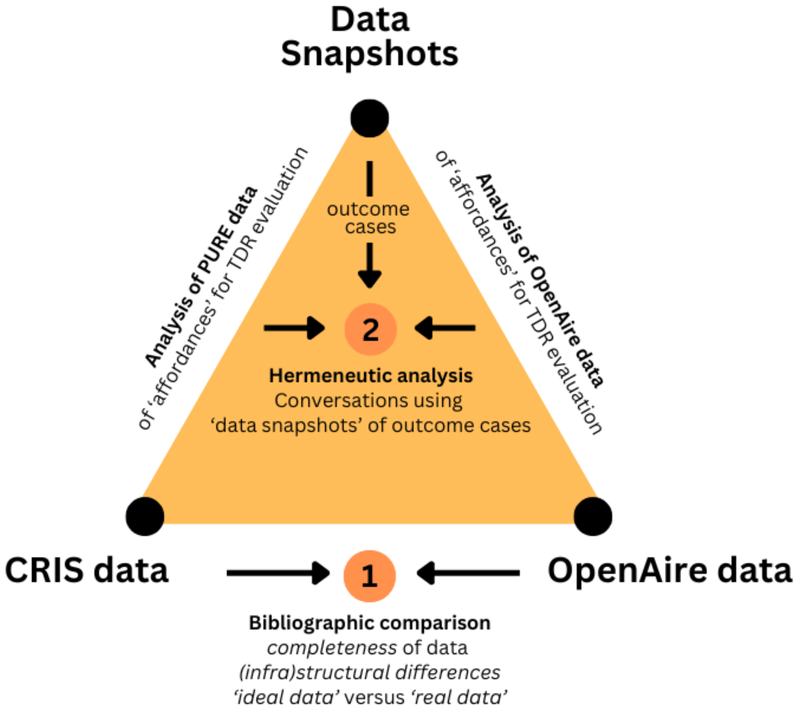

Figure 5: Overview of analyses of UU pilot

In the next section, probing, we will present the results from both analyses. We will first describe how we retrieved the necessary data from the CRIS and OpenAIRE respectively, which we follow up on by describing the datasets. Lastly, we compare both datasets. In section 7.2, we describe the hermeneutic analysis ('data snapshots').

7. Probing

7.1. Comparison of data sources

In this chapter we discuss the comparison between *OpenAIRE* and CRIS data. We describe data retrieval efforts, main findings from the comparison and how both sources afford for the evaluation of transdisciplinary research. In each section we also add notes, reflections and limitations on the process which serve as learnings from the comparison.

Analytically, the OpenAIRE and CRIS corpus were analysed ***potentially*** and ***actually***. We made this distinction because we learned that the *evaluative potential* of each corpus (i.e. what claims can be made with each corpus) is different than the *actual evaluative potential*. That is, because while both datasets offer an extended taxonomy for monitoring research outputs (and also research activities in the case of the CRIS), the data is often not populated, or at least not enough for evaluative purposes. This resulted in two analyses that will be presented below: a comparison between the *potential extent of data* and the *actual extent of data* between OpenAIRE and CRIS.

7.1.1. Data retrieval[15]

We focused the comparison between OpenAIRE and CRIS data on the period 2020–2025, as this timeframe aligns with the upcoming research evaluation of the Copernicus Institute of Sustainable Development. In this context, the comparison also functions as an exploration, or mock-up, of the

[15] See Appendix 9.2 for further methodological notes

evaluative potential offered by the two data sources. In order to compare OpenAIRE data on the department with CRIS data on the department, we first had to get access to the two datasets.

For the OpenAIRE dataset, the first issue was that we had to determine the parent-child relations between different organisational units of Utrecht University so that individual items or researchers can be linked to the department unambiguously in the data architecture of OpenAIRE. After a couple of iterations and meetings with the OpenAIRE team, one solution to this problem seemed to be the use of another OpenAIRE service, namely OpenOrgs. This service has as its goal to aggregate information on research organisations from a variety of different research organisation registries, such as ROR[16].

While this step entailed a manual disambiguation of different instances of that Copernicus Institute name, the administrative processes surrounding this intervention were complex. What in essence was a matter of logging into a portal and classifying items that related to varying search terms of the department's name in a database; became rather complex because it involved processes attached to the OpenOrgs curation process specifically, including one-on-one training and the signing of a 'volunteer contract'.

In a second step, the updated hierarchies between research units was imported ('dumped') into the new version of the OpenAIRE Graph. Based on this new version, in March 2025, OpenAIRE colleagues extracted a corpus listing all information of researchers being affiliated to the Copernicus Institute of Sustainable Development in the period between 2020 and 2025. The corpus was a collection of .json files, which we subsequently converted to a readily interpretable format. For the conversion, the OpenAIRE scheme was used.

After extraction, the OpenAIRE dataset included 1406 entries, all of which were publications. From this dataset, we deduplicated 310 entries[17]. The final dataset included 1096 unique items.

Limitations and notes:

- One limitation is that the request may not have been non-ambiguous to OpenAIRE. Meetings took place and so on to clarify the request, but we cannot be sure that our request was translated correctly, technically.
- Later in the process, it became apparent that 298 entries of the initial 1406 entries were published before the requested time period. Given that the analysis was based on the initial dataset including the older entries, we can assume a slight overestimation of OpenAIRE data.

For the CRIS dataset, the data was retrieved via Research in Context Graph (Ricgraph)[18], an application that retrieves the data from CRIS. The data listed all information of researchers while being affiliated to the Copernicus Institute of Sustainable Development in the period between 2020 and March 2025. Similarly to OpenAIRE, the json files were converted into an interpretable format using python. The total number of entries was 3866, of which 289 appeared to be duplicates. After deduplication[19], the final dataset included 3577 unique items.

Limitation and notes:

- Limitation for retrieving data is that we don't have reference material for triangulation to assess the completeness of the data

[16] See https://ror.org/
[17] See appendix 10.2 for deduplication technique description
[18] see Janssen 2024. Ricgraph was used after being recommended to us by the central administration of research information of Utrecht University either as a user interface for data retrieval or as a tool itself that can perform different analyses. We used it for the former purpose.
[19] See appendix 10.2

- The Research in Context Graph application was not used before in a mock-up and working with the application and retrieving the CRIS data resulted in valuable user feedback and learnings for the developers of the application.

7.1.2. Overview of unique items and overlaps between datasets

Figure 6 shows the distribution of unique OpenAIRE entries (red), CRIS entries (blue), and entries found in both datasets (green).

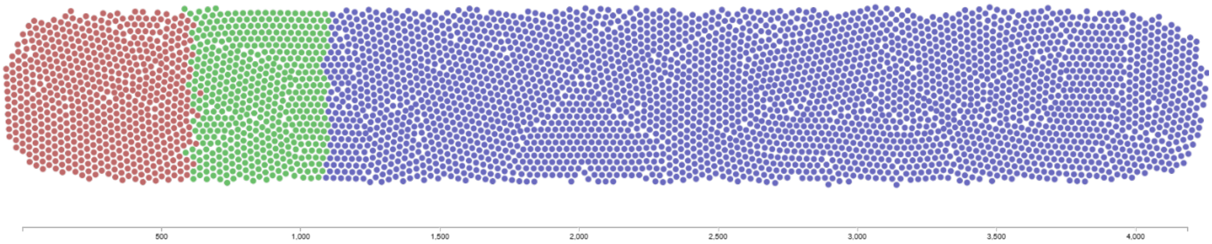

Figure 6: Unique and Overlapping Entries between the CRIS and OpenAIRE based on corpuses on the Department of Sustainable Development at Utrecht University[20].

7.1.3. Comparison of 'potential' data

Below is an overview of the *potential* items that the CRIS and OpenAIRE monitor. For every table, a description of the rows is given per classification.

<u>OpenAIRE</u>

Category	Column (ideal)	Content
Identifier	id	OpenAIRE interal Identifier
	doi	Digital object identifier
	originalIds	Other identifiers for OpenAIRE internal use
	pids	persistent identifiers
Content	mainTitle	Title of the output
	subTitle	Subtitle of the output
	description	Summary of the output
	subjects	Keywords related to the content of the output
	language	Language the output is in
	coverages	The time and/or locations covered by the publication
	geoLocations	Geolocation information
Output	type	See output types
	sub_type	See output types
	formats	See output types (https://www.iana.org/assignments/media-types/media-types.xhtml)
Open Access	isGreen	Whether the output has at least the Green Label for Open Access
	isInDiamondJournal	Whether the resource is in a Diamond journal
	BestAccessRightCode	The code describing the level of Open Access (according to COAR https://vocabularies.coar-repositories.org/documentation/access_rights/)
	BestAccessRightLabel	The label describing the level of Open Access (from COAR label)
	openAccessColor	The Open Access colour of the output (levels: gold, hybrid or bronze)

[20] Visualisation was produced using RAWGraphs (Mauri et al. 2017)

	embargoEndDate	Date when embargo ends and result becomes open access
Indicators	downloads	Number of downloads an output has
	views	Number of views an output has
	citationCount	Number of citations an output has
	impulse	Initial momentum of output, based on citation network
	influence	Popularity of output over the whole period, based on citation network
	popularity	Current popularity of output, based on citation network
Publication details	publisher	The name of the entity that holds, archives, publishes prints, distributes, releases, issues, or produces the resource.
	publicationDate	The date on which the resource is published
	journal (from container)	Name of journal or conference
	volume	Volume in which the resource appears
	startingPage	Page on which the resource starts
	endingPage	Page on which the resource ends
	collectedFrom	Information about the sources from which the record has been collected
	dateOfCollection	Date on which the output is collected for the document containing the data
	lastUpdateTimeStamp	Timestamp of last update of the record in OpenAIRE
	sources	A Reference to a resource from which the present resource is derived
	version	Version of the result
Funding	publiclyFunded	Whether the resource is publicly funded
	Project (definition)	List of projects (i.e. grants) that (co-)funded the production of the research results
	acronym	Abbreviated project title
	project_title	Project title
	funding_stream	Stream of funding (e.g. for European Commission can be H2020 or FP7)
	jurisdiction	Geographical jurisdiction (e.g. for European Commission is EU, for Croatian Science Foundation is HR)
	funder_name	The name of the funder (e.g. European Commission)
Author	authorFullName	The full name of an author
	authorOrcid	The ORCID identifier of an author
	authorRank	The rank of an author
Groups belonging to output	countries	The different countries related to the output
	communities	The OpenAIRE communities the output is part of
	contributors	Contributors for the result
	organisations	List of organizations in an affiliation relation with the research results

Table 6: Potential aspects for evaluation of the OpenAIRE dataset

What is noticeable are the multiple identifiers that OpenAIRE uses. Column "id" is a local identifier for OpenAIRE, whereas the OriginalIds are identifiers of other sources that OpenAIRE uses. The PIDs have the same function, although there the scheme is also provided, so that other users can access it as well.

There are four overarching output types, each pertaining to a set of subtypes. These include:

a. Publication
b. Research data
c. Research software
d. Other

Of these, a new categorisation of output types has been created to increase granularity and better divide the subtypes over the overall output type. These include:

a. Publication
b. Thesis
c. Non-textual
d. Event/activity
e. Membership
f. Other
g. Collection
h. Conference object
i. Data management plan
j. Data paper
k. External Research Report
l. Internal report
m. Project deliverable, milestone and proposal
n. Research
o. Software paper
p. InteractiveResource
q. Model
r. PhysicalObject
s. Research Activity Identifier
t. Research Object
u. Bioentity

In general, the combination of the title, description and subjects can provide a general idea of what the output is about. Unfortunately, there is no concrete information on what 'coverages' entails exactly, but we presumed it to describe timestamps and geolocation of a certain output. The separate geoLocations output, on the other hand, appears novel. Nonetheless, we are not sure what it aims to describe concretely. Type and subtype are interesting, especially with regards to transdisciplinary research as some relate to activities that researchers may engage in (e.g. 'event/activity' or 'membership', which may tell about collaborations or other external relations that researchers may have.

In terms of open access, there are various columns aiming at describing whether a publication is in a green, bronze, gold, hybrid and/or in a diamond journal.

Furthermore, the indicators that are provided have the classical citation scores and additionally the views and downloads of any given item. 'Impulse' describes the influence and the popularity that the citation scores are elaborated on.

Publication details appear to be classical scientometric values. There are, on the other hand, 3 columns describing the collection process of OpenAIRE. These are dateOfCollection, collectedFrom and lastUpdateTimeStamp. The columns can be relevant, but are similar to other columns, such as instances and sources for the CollectedFrom column, and dateOfCollection resembles lastUpdateTimeStamp. Also, there appears to be considerable overlap in terms of where the data is retrieved from. OpenAIRE provides data on the sources, instances and collectedFrom. This makes it complicated to use, since it is unclear which of these columns would be the best to use and why.

The funding columns of OpenAIRE appear extensive. The projects relating to the output are clearly given through the project_title, as well as the funding_stream, funder_name and jurisdiction relating to the projects.

In terms of the author section, OpenAIRE has only the fullname and (potential) ORCID. Even though organisations are also provided in the data, there is no connection between the author and the organisations.

For the groups belonging to the output, there are the countries, contributors, communities and organisations. The contributors and organisations are not the same, but in terms of function they appear to resemble one another. These points are mainly relevant for making organisational connections, which may be of importance to transdisciplinary research.

CRIS

Category	Column (ideal)	Content
Identifier	the CRISID	Identifier used in the CRIS
	uuID	Identifier used in the UU
	ExternalID	Identifier that is not the CRIS or UU
	ExternalIDSource	Source the identifier belongs to
	PrettyURLIdentifiers	Variable resembling "Title"
	AdditionalExternalIDs	Additional identifier
	IDSource	Source the identifier belongs to
Output	Output_type	See ideal output (output)
	Output_category	Professional area in which the output is produced (academic, professional, other)
	Event_name	Name of the activity
	Event_type	See ideal output (activity)
Content	Title	Title of the output
	Abstract	Abstract of the output
	Keywords	Keywords associated with the output
	BibliographicalNote	Variable resembling "event_name"
	Language	Language the output is in
Open Access	OpenAccessPermission	Whether the output is open access (not known how this is identified)
	Visibility	Whether the output is publicly visible
	Workflow	?
Publication details	Publication_year	Year of publication
	Publication_month	Month of publication
	Publication_day	Day of publication
	publisher_name	Name of the publisher
	publisher_type	Type of the publisher (?)
	journalAssociation_name	Name of the journal
	journalAssociation_type	Type of the journal (?)
	Publication_status	Published or not
Authors	totalNumberOfAuthors	The number of authors working on the output
	Person_the CRISID	The the CRIS identifier of an author
	CorrespondingAuthor	Whether an author is the corresponding author of the output
	Person_uuID	The UU identifier of an author
	Person_fullName	The full name of an author
	Person_firstName	The first name of an author
	Person_lastName	The surname of an author
	Person_role	The role of an author

	external_person	Whether the person is external to the UU
	Person_organisation_uuID	The UU identifier of an organisation an author belongs to
	Person_organisation_name	The name of an organisation an author belongs to
	Person_organisation_type	The type of an organisation an author belongs to
Groups belonging to output	OrganisationalUnits_uuid	The UU identifier of the overarching organisation the output belongs to (e.g. research programme)
	OrganisationalUnits_name	The name of the overarching organisation the output belongs to (e.g. research programme)
	OrganisationalUnits_type	The type of the overarching organisation the output belongs to (e.g. research programme)
	commissioningBody_name	The name of the organisation that commissioned the output
	commissioningBody_type	The type of the organisation that commissioned the output
	ManagingOrganisationalUnit _Value	The name of the organisation managing the output (e.g. a department where a report is stored)
	ManagingOrganisationalUnit _Type	The type of the organisation managing the output (e.g. a department where a report is stored)

Table 6: Potential aspects for evaluation of the CRIS dataset

A peculiar aspect about the identifiers in the CRIS dataset is that there are both the ExternalIds and the AdditionalExternalIds: they contain the same type of data, but split up. It is not clear as to why this is the case. Moreover, the PrettyURLIdentifier appears to resemble the “title” column, but with hyphens between the words.

For the Outputs, there is the output type and event type. The distinction between the two is interesting, as it relates to the difference of having a work as output, as opposed to having an event/activity as output.

In terms of content, the main items are represented: There is a title, abstract and keywords to identify the general content of an output. Moreover the BibliographicalNote provides additional information regarding funding and collaborations with people/organisations. This resembles the section of Funding and Acknowledgements at the end of a scientific publication.

The OpenAccessPermission column is the only one describing open access statuses. The possible values here are not known. However, the column on visibility provides information on whether there any restrictions for viewing the output.

The Workflow column appears to be for internal use in the CRIS, but this has not been verified. The values in this column can be “Validated", “Approved”, “For approval" and “Entry in progress", suggesting that it regards the workflow of the CRIS rather than that of the output.

The PublicationDetails again mainly contain classic variables, such as publisher and journal names. Additionally, the ‘publication status’ is provided, providing information on whether the output is already published or not. Presumably this applies to academic publications only. Whether preprint-publishing are part of this classification is unknown.

The authors section is extensive. There are three columns for the names of the person (first, last and full). The Person_role and external_person provide useful additional information on the authors. Lastly, the authors can be connected to specific organisations.

The last category contains the groups related to a respective item (Organisational_Units), funding of an item (commissioning_Body) or its management (ManagingOrganisationsBody).

7.1.4. Comparison of 'actual' data

We also analysed and compared the 'actual' extent to which data on the department was provided by both the CRIS and OpenAIRE. To assess this, we calculated the number of occupied rows by counting the total rows and maximum number of unique rows. This might lead to an overestimation of the number of occupied rows for the unique titles, but it provides a general indication for how occupied the dataset is.

<u>OpenAIRE</u>

Most of the columns of the OpenAIRE ideal data were occupied, and of those most had 500+ instances. The specific occupancy of the columns can be found in the following table, with underneath a description about the interesting findings.

Classification	Column	Maximum items (n)	Actual Occupancy (%)
Identifier	id	1400	78.21
	doi	672	100.00
Content	mainTitle	1400	78.21
	subTitle	138	100.00
	description	1314	83.33
	subjects	1290	84.88
	language	1400	78.21
	coverages		
	geoLocations		
Output	type	1400	78.21
	sub_type	1400	78.21
	formats	944	100.00
Open Access	isGreen	618	100.00
	isInDiamondJournal	618	100.00
	BestAccessRightCode	1362	80.40
	BestAccessRightLabel	1362	80.40
	openAccessColor	571	100.00
Indicators	downloads	94	100.00
	views	94	100.00
	citationCount	1400	78.21
	impulse	1400	78.21
	influence	1400	78.21
	popularity	1397	78.38
Publication details	publisher	729	100.00
	publicationDate	1397	78.38
	journal (from container)	562	100.00
	volume	549	100.00
	startingPage	400	100.00
	endingPage	270	100.00
	dateOfCollection	1400	78.21

	version		
Funding	publiclyFunded	618	100.00
	acronym	478	100.00
	project_title	622	100.00
	funding_stream	623	100.00
	jurisdiction	532	100.00
	funder_name	609	100.00
Author	authorFullName	1400	78.21
	authorOrcid	582	100.00
	authorRank	1400	78.21
Groups belonging	countries	1248	87.74
to output	communities	1400	78.21
	contributors	712	100.00
	organisations	1400	78.21

Table 7: Openaire output extent

Below we present some observations:

- Only 7 of the 138 the subtitles are not in the main title. This means that subtitle is generally unoccupied and for only 7 it provides additional information.
- Language is often also "undetermined". This is peculiar, since in other columns, any missing information translates into an empty cell.
- In OpenAIRE, there can be multiple same articles. This is because the subtype of an output is based on the information that OpenAIRE obtained from the instance of the output. Since there are multiple subtypes and multiple instances, it is difficult to assess what the most accurate subtype is.
- The formats of the data are derived from a list of media types an output can have (https://www.iana.org/assignments/media-types/media-types.xhtml). However, there are also formats that are not included in the list of media types. These generally contain only numbers.
- The information that exists on OpenAccess is mainly BestAccessRight.
- For the publication details, only the publication date and date of collection are complete. The other publication details are only half filled.
- AuthorORCID: only 582 rows with authors also provide at least one author with ORCID.
- In the organisations, there are duplicate organisations, where some organisations have a list of pids, and other organisations have no pids.

CRIS

It is not possible to assess the completeness of this dataset, it cannot be compared with the total amount of outputs since that total cannot be assessed manually. There are 289 duplicate titles, where most of the duplicates have the title like "Author Correction", "Corrigendum" or "Editorial". Of these rows, most of the data of the ideal columns is empty in the real data.

Classification	Column	Maximum items (n)	Actual Occupancy (%)
Identifier	CRISID		
	uuID	3866	92.52
	ExternalID		
	ExternalIDSource		
	PrettyURLIdentifiers		
	AdditionalExternalIDs		
	IDSource		
Output	Output_type	3866	92.52
	Output_category		
	Event_name		
	Event_type		
Content	Title	3866	92.52
	Abstract		
	Keywords		
	Language		
	OpenAccessPermission		
Open Access	BibliographicalNote		
	Visibility		
	Workflow	3866	92.52
Publication details	Publication_year	3866	92.52
	Publication_month	2864	99.23
	Publication_day	1266	98.66
	publisher_name		
	publisher_type		
	journalAssociation_name		
	journalAssociation_type		
	Publication_status	3866	92.52
Authors	totalNumberOfAuthors		
	Person_the CRISID	3866	92.52
	CorrespondingAuthor	3866	92.52
	Person_uuID	2096	99.14
	Person_fullName	2096	99.14
	Person_firstName	3861	92.64
	Person_lastName	3864	92.57
	Person_role	3866	92.52
	external_person	2721	98.71
	Person_organisation_uuID	2039	99.26

	Person_organisation_name	2039	99.26
	Person_organisation_type	2039	99.26
Groups belonging to output	OrganisationalUnits_uuid		
	OrganisationalUnits_name		
	OrganisationalUnits_type		
	commissioningBody_name		
	commissioningBody_type		
	ManagingOrganisationalUnit_Value		
	ManagingOrganisationalUnit_Type		

Table 8: CRIS output extent

Below we present some observations:

- A lot of useful information appears to be missing. It is unclear how this could be the case.
- The identifiers are empty, with the exception of the UUid. External Ids would be especially useful, since these can connect the CRIS database to other databases.
- The column of event type is empty. However, from the column output type it is evident that there are still some activities present. It is unclear how activities identified from the column output type differ from that of the column event type.
- From the content, the only remaining column is the titles. The abstract, keywords and bibliographic notes are empty.
- The OpenAccessPermission and visibility are also empty. The workflow is the only column that still remains, but this column contains the least interesting information.
- It is remarkable that the publication date and status are present, but that publisher and journal is missing from the data.
- In contrast to the other columns, the person and organisation data is still completely present. The only columns that are not 100% filled are Person_uuid, Person_fullName and organisations. These have above 2000 rows each.
- The external person column has 2721 rows. This means that for the other rows, none of the authors was external.
- Lastly, the total number of authors is not mentioned, but this can be derived from the rest of the data.
- Organisational Units is empty, but perhaps it could be that this information is contained in the Person_organisational_units.

7.1.5. Overall comparison of data sources

For the comparison of data sources between CRIS and OpenAIRE, we analysed the key differences, strengths and weaknesses. This was done by documenting the processes of data retrieval, processing and analysis. In that sense, the comparison does not only relate to the corpuses themselves, but also to their embeddedness in the (organisational, technical) systems of origin. For the datasets themselves, both 'ideal' and 'real' were taken into consideration.

Differences:

- OpenAIRE tends to focus on traditional bibliometric aspects of research information. The dataset contains primarily publication data, such as indicators, publishers or journals, and open science features primarily as open access publications.
- OpenAIRE includes funding data. In particular whether or not a resource is publicly funded, grants, project titles, funding streams and name of funder. It is unclear how OpenAIRE collects this data.
- OpenAIRE does not connect researchers to organisations. Rather, it maintains a full list of authors and a full list of organisations. Especially since OpenAIRE does not appear to use author identifiers other than ORCID, connecting authors to each other or to organisations through OpenAIRE data is complicated.
- OpenAIRE does have an extensive list of data categories for the publication, making it very useful to describe the contents of that publication and what organisations and funds are related to it. Additionally, OpenAIRE has identifiers available for publications, which can be useful for interconnecting databases.
- The CRIS is much more focused on different output types, and connections between the output types, the persons and the corresponding organisations. The inclusion of the role of a person and whether they are external or internal to Utrecht University also emphasises the focus on individuals.
- In comparison to OpenAIRE, the CRIS invests on different identifiers, output types and output categories, and the persons related to the output. For our purposes, the distinction between outputs and events is particularly interesting.
- Data on organisations in the CRIS include their roles related to an output, differentiating between commissioning, creating and managing organisations. This enables insight into what the output was and how it came to be.
- A key difference between OpenAIRE and the CRIS is that OpenAIRE harvests the information from online sources, such as publication repositories, research data archives, or funder databases, whereas the CRIS obtains the data directly from the authors and extends this data by relying on automatic population from e.g. Scopus (e.g. metadata).

Strengths

- A primary strength of OpenAIRE is that it is interconnected with the total OpenAIRE database for projects. For many outputs, OpenAIRE can identify research projects and grants, and has more data on this internally
- A primary strength of the CRIS data is that it contains information on the role and position of a person
- Furthermore, CRIS data has an extensive set of output types in the event/activity category which can be of relevance for evaluating transdisciplinary research.

Weaknesses

- The biggest weakness of OpenAIRE seems to be its completeness. The difference between 3577 items on the CRIS and only 1400 items on OpenAIRE data seems concerning. This also came back in the data-interviews (see data snapshots (7.3) and evaluation (8)).
- One weakness of the OpenAIRE database for our purposes is that it does not connect authors to respective organisations. I.e. ‘author’ is not an entity in the OpenAire graph.

This means that in order to assess research collaboration in general and transdisciplinarity research specifically using the OpenAIRE database, an organisation-level analysis can be performed but not an individual-level analysis. The extent of a collaboration can therefore not be assessed, as it is unclear how many people from how many organisations are involved. For example, if one of the contributors is a researcher who works at an academic and a non-academic organisation, that would appear the same as when one contributor works at an academic organisation and another one works at a non-academic organisation.

- The indexed outputs in the OpenAIRE database are the largest extent academic[21]. This presents a significant challenge when evaluating transdisciplinary research, which may result in non-academic outputs such as e.g., policy reports.
- There is only limited information on event types in the OpenAIRE database.
- Quite a large selection of the information on the OpenAIRE database is for internal use and relate primarily to identifiers.
- While the OpenAIRE corpus has useful information on research outputs, funding and organisations, other parts of the database are not intuitive and require the use of other OpenAIRE data, or data from other data sources to become sensible.
- The main weakness of the CRIS is that no funding data is available for a particular output, and neither what research project the output is related to.
- Another weakness of the CRIS is that it requires manual curation particularly of non-traditional outputs such as events which can be labour intensive and incomplete.

Finally, both datasets lack full text, as well as concrete citations or references that are used in the work[22]. This would potentially enable further analysis on the transdisciplinarity of the outputs.

7.2. Potential for evaluation of transdisciplinary interactions

As we were interested in testing how the two corpuses afforded the evaluation of *transdisciplinary research processes* between academic and non-academic actors, we listed possible interactions (see Table 5), which were based on observations of research practices within the department relating to transdisciplinarity. These interactions thus describe research processes and activities, such as ‘academic co-authorship’ between an employee of Copernicus Institute and an external ‘non-academic’ author, , which arguably can be used for evaluating transdisciplinary research. We were particularly interested to learn how OpenAIRE data and CRIS data afford their evaluation.

In what follows, small paragraphs will elaborate on each (or collections of) items:

- **Co-authorship on either academic or non-academic output**

Although the OpenAIRE database allows for assessing whether an academic and non-academic affiliation appear on a single item, it is difficult to assess the co-authorships of academic and non-academic affiliated persons in the OpenAIRE database. Outputs can have organisations or contributors from both academic and non-academic backgrounds, but it is not possible to assess this on an individual-level. Only the organisations linked to the output can be used. At the same time, it would be possible that there is one author who is affiliated with both an academic and a

[21] There are instances of ‘reports’, but the vast majority are academic objects (e.g. articles, theses, books, DMPs, etc.)

[22] At least in the data that we have received by our OpenAIRE colleagues, as the OpenAIRE Graph does have references to citation relations (see https://graph.openaire.eu/docs/data-model/relationships/relationship-types).

non-academic organisation. This would mean that the output did not enable the collaboration between the two organisations, but that there is an author who already was connected to both organisations, who now collaborates on the output. In consequence, the number and range of non-Copernicus and non-Utrecht University organisations related to an output is an interesting datapoint, although highly contextual that cannot be understood by relying on the data point alone.

The CRIS does allow this analysis. Persons are considered external or not, and they are each linked to an organisation. This makes the identification of whether a person has an academic or non-academic affiliation possible. In extension, it can determine whether there was a collaboration between an author with both an academic and non-academic affiliation.

- **Share research project (w. grant)**

Sharing here specifies that non-academic actors and academic actors are both partners in a research project Share research project with grant funding. This can, in theory, be analysed using OpenAIRE, as funding data is collected by harvesting metadata directly from research funding agencies or by full-text mining of articles when they are made available[23]. Currently, direct import of metadata comes from exclusively Euro-American agencies. In the CRIS data that we have of the department, no data on funding is available, so it’s not possible to assess whether the research project also received a grant. The CRIS has the organisation type “research programme”, which could describe a research project.

- **Interacted in research project (non-formal partnership)**

OpenAIRE does, currently, not allow the analysis of non-formal project partners. The persons that are found in the data that OpenAIRE provided are direct contributors to an output (i.e. authors). There is no other category where persons can be found.

In the ideal CRIS, the ‘BibliographicalNote’ may contain additional information, such as acknowledgements of interlocutors acknowledgements and may be therefore used to at least document informal partners (after revision). In the real data, no ‘BibliographicalNote’s were present.

- **Shared presentation and presenting to a non-academic audience**

While ‘presentations’ as output are listed in *ideal* OpenAIRE data and constitute presentation files sourced from Zenodo, the *real* OpenAIRE data not the CRIS data included any presentations. In the CRIS, there is an output that is a podcast which could be considered as presentation to a non-academic audience. At the same time, understanding the presentation as an *interdisciplinary event* is also not viable with the data available.

- **Interactions between diverse expertise**

While interactions between scientific disciplines may be approached bibliometrically by way of delineating fields and analysing their interconnections, a non-scientific extent of expertise (which is important in transdisciplinary research) is not an explicit phenomenon for monitoring in either CRIS or OpenAIRE.

- **Framing a problem**

Project and grant data is available in OpenAire, but not the proposal or agreement, through which problem framings of research could be approached in combination with corresponding project descriptions and the formal project partners. This means that OpenAIRE allows for the identification of some forms of problem framing in relation to projects. The CRIS only contains titles of collected outputs in the real data, making it impossible to accurately consider the framing of problems.

[23] See https://www.openaire.eu/blogs/openaire-research-graph-data-dump-of-funded-products

- **Information from and to stakeholders**

For the CRIS and OpenAIRE there exists no output of solely information sharing. In OpenAIRE, output a type of "contribution to newspaper or weekly magazine" exist. The CRIS contains the output type "report" and "podcasts", which could have the function to share information to stakeholders.

- **Meaningfully mentioned or cited together in (full-text) of output**

Neither OpenAIRE nor the CRIS allow for this. The full text is not available and there are no citations known for the outputs. For OpenAIRE, the doi is available, making it possible to link the outputs to another database and get the full text from there. The CRIS only uses internal IDs and does not disclose dois.

- **Advancing of / handling of matters of concern**

OpenAIRE contains grant data, the description, subjects and title. Combined, these can give an indication as to what the problem statement and the anticipated results are. the CRIS only contains the title, and therefore does not allow such an indication.

7.3. Data snapshots (hermeneutic additions)

A large part of the comparative work introduced before focused on existing categories and affordances of data that is commonly used for research assessments. With this exercise, we wanted to understand what kinds of work of researchers is and is not captured by both the CRIS and OpenAIRE.

For this, we reached out to five researchers across the different sections of the institute and seniority. We invited them into a conversation where we presented these researchers, individually, with printed visualisations of their data in both OpenAIRE and the CRIS. Using these printed sheets as conversational prompts about how they understand their research work, and in particular impact-related work, in view of the visualisation in front of them. Then, we 'sketched' and 'extended' these visualisations until the researchers would recognise themselves satisfactorily in them. Find below a blurred example of these sheets.

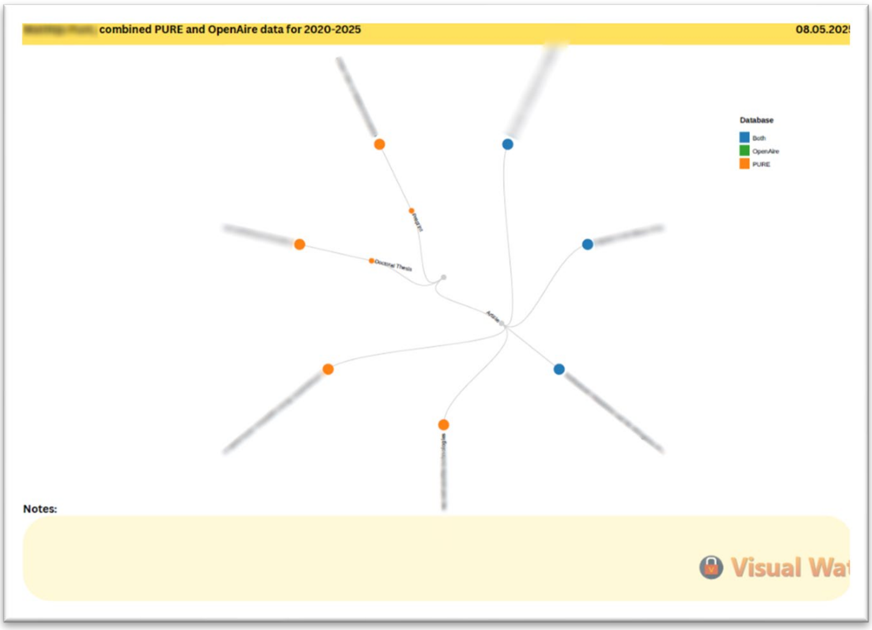

Figure 7: Blurred example of printed visualisation of researcher data for interviews

Figure 8: Blurred example of printed and filled in visualisation of researcher data for interviews

7.3.1. Results

#1 Assistant Professor	**Note**: Immediate reflection was that the listed items were only 'very formal' aspects of their work.
Existing data: Doctoral thesis Preprints Articles	**Added aspects:** Supervision of PhDs at other university Peer reviewing Engagement in advisory committee Involvement in facilitation of scholarly impacts Non-academic speaking: - as panellist - as speaker

#2 Postdoctoral researcher	**Note**: They had immediately the feeling that they don't relate to this [the visualisation of their CRIS and OpenAire data], nor as well as "scores and metrics on Google Scholar" [unprompted]. Stressed that they don't want to be motivated by that. They described that intrinsic motivation is very important to them and started at Copernicus Institute exactly for transdisciplinary research and the kind of work that is available at the institute.
Existing data: Abstracts Review Articles Articles	**Added aspects:** Communications: - facilitating a community of practitioners - developing grey literatures: (e.g. workbooks, guides, worksheets, blogs) Supervision of PhDs and MScs Developing courses for professionals 'Direct Impact' projects: - mapping, building and maintaining network - Organising series of workshops and events - Long-term collab. with societal partners

#3 Researcher on project funding	**Note**: They edited and added the data into the CRIS themselves based on prompts from the secretariat. Felt like the proportions between the web publication and the sheet do not match in terms of impact.
Existing data: Web publications Posters	**Added aspects:** Global South collaborations: - academic collaborations - non-academic collaborations Building capacity of Global South junior Academics: - funding allocations through collaborations Communications:

	- grey literature - newsletters - conference papers - social media Policy interventions - building a database for African case studies

#4 PhD	**Note**: Privacy-related issues inhibit the publication of data. Has a focus of finishing the PhD, which is why the data that exists makes sense to them (Seeing 'paper one, two and three') of the PhD project. If they had more time, they would develop high school teaching materials, but the PhD takes too much toll.
Existing data: Articles Book	**Added aspects:** Working in an interdisciplinary research project Mental load of proposal-writing

#5 Assistant Professor	**Note**: They mentioned that they got frustrated with the CRIS and stopped adding items. The secretariat would send annual emails, collect data of employees and add this data themselves. They recognised considerable duplicates and wrongly classified outputs from the CRIS and OpenAire. They also mention an aspect of temporality: one item was a preprint, then became a published and peer-reviewed paper. It appears twice in the data. The category 'review articles' suffers a lot from misclassifications ranging from a literature review to an opinion paper. Felt uncomfortable with pub data.
Existing data: Abstracts Review articles Articles Comments/Letters to the Editor Preprints Report Web publications Posters Papers Letters	**Added aspects:** Scientific community contributions: - peer reviewing - editing / editorships - managing activities PhD Supervision Education Articles in newspapers Initiating scientific reflections in quarterly report of the particular scientific field (journal).

Table 10: Results and notes from data snapshots

7.3.2. Conclusions

In this section, we briefly outline some key aspects that emerged from the sketching-interviews. These can be read as indicative of topics and items that would deserve more inquiry in assessment of transdisciplinary research. In any case, these aspects are interesting as these show the boundaries of what actually can and cannot be grasped by research information infrastructures, whether proprietary or not.

Temporality of research outputs
Some respondents noted that duplicate items in the visualisation (and/or item list) change while retaining the same title. This is exemplified by one respondent who noted that at first, they published an article as a preprint, which then became an article after going through peer review. This was the case multiple times.

Misclassifications
There were multiple misclassifications that were spotted by the respondents, or at least resulted in confusion. This was particularly so in reference to the CRIS's output type 'Review Article', to which respondents had a wide array of publications attached that they did not understand as 'review' articles per se.

OpenAIRE coverage
What is striking is that in this small sample of five researchers, the use of OpenAIRE resulted in one added item that was not documented by the CRIS. This exceeds the ratio that we presented previously multiple times.

(Transdisciplinary) grey literature
Two of the respondents who engage strongly in transdisciplinary research with different communities mentioned a vast array of *objects* that become produced; but mechanisms are not in place (and the question is whether there should be) to capture these objects. This includes workbooks, or guides that are written for practitioners.

(Extra-)academic collaborations
Researchers that engage in collaborations, both academic and non-academic, tended to flag that these were missing in the representations that we presented them. This was in part due to choices in visualisation, as data of co-authors, for instance, exists in the datasets. We find it important to flag nonetheless; especially those collaborations that are non-academic, as these are untraceable as of now.

Administrative processes
Being from different sections of the department, these researchers also experiences different ways of actually documenting their work. While one once 'stumbled' on the CRIS and added items, including web publications, another respondent mentioned that the secretariat periodically collected the section's research works and documented these *for* the researchers themselves.

Invisible work
In what has been mentioned as 'invisible work', most respondents spoke of 'scientific community work' as an integral part of their work that needs different evaluation processes and methods to recognise and make rewardable. This work includes editorships or peer reviews, for instance. For transdisciplinary research practice in particular, invisible work also spans community building and maintenance of long-term relationships, which some respondents flagged.

Additionally, what we find worth noting is that two interviewees, when being presented with their the CRIS and OpenAIRE data (which spanned doctoral theses, articles and preprints), declared that they do not see themselves in this visualisation and avoided adding to this visualisation. As a result, they started a totally new sketch; and one of them also on a different sheet of paper. Upon following-up, our interpretation was that for them, academic objects are only *documentations* of what they understand being their 'actual' (transdisciplinary) research work, which they spoke of as 'direct impacts'.

8. Evaluation: discussion and recommendations

We will use this section as a space to share reflections and discuss recurring aspects. Whilst it does not follow SCOPE per se in that we don't evaluate our interventions, the idea is to capture learnings from the pilot for both OpenAIRE (and more widely other GraspOS research information services and tools) and the teams dedicated to the design and conduct of future research evaluations within the pilot institute.

Of course, there are several observations that we could address here, but we find it important to go back to the starting point of GraspOS and our pilot project. This is why we will discuss each SCOP(E)-step individually after an initial general reflection.

The way we read the inquiries that we conducted in each step is that they contribute to an increasingly clear picture of the (social, organisational and technical) complexity in which departmental research evaluations happen.

Finding open science therein is a fragmented affair. In our case, we found explicit associations to open science through the university-wide strategy and the strategy evaluation protocol as formulated by the Dutch knowledge institutions. From a 'bottom-up' perspective, open science concerns existed, albeit implicitly and mainly articulated themselves in aspects of societal impact, transdisciplinary research and generally concerns over societal value creation through science.

As all levels referred to transdisciplinarity[24] as one of the major concerns for doing science openly, we focused our efforts on its evaluation and tried to understand how GraspOS tools and services may play a role in it. After an assessment of existing tools and services, we decided to test with OpenAIRE data as an addition to the data sources that are already used for the departmental evaluation of the department, which is now, in May 2025, in planning. We also did so because our team may have direct, sustainable GraspOS impacts through being involved in that evaluation.

The main analysis revolved around the comparison of the CRIS and OpenAIRE data on the department with two focal points: one being the general comparison of *potential* data that is being monitored by each provider. The other revolved around the extent to which usable data exists on the department for each provider. To highlight valued research practices that we had observed in the preceding analyses, we did a projection exercise where we estimated to what extent data from both providers affords making claims on transdisciplinarity. Finally, to sensitise our findings, we conducted exploratory interviews asking respective researchers to reflect on their existing data in view of their work on societal impact, which highlighted further a key insight from our work at UU: namely that the CRIS and OpenAIRE data are serving particular use cases and are limited especially with regard to their appropriation for transdisciplinary research evaluation. We provide a number of discussion points below regarding what can and cannot be grasped by this data.

Grasping (aspects of) open science

In our case, we found that same department is implicated in different movements that advocate for reconfiguring the value registers by which the organisational strategy is formulated and research is conducted and evaluated. One major source of this complexity emerges through the sheer interdisciplinary character of researchers that share the department. Everyone makes sense of and values their research from with their own specific histories and field-specific epistemic practices.

[24] As mentioned earlier, there were various ways to refer to transdisciplinarity that all described similar concerns. Namely inclusive research processes across academic and non-academic actors.

As this diversity is sought to be accommodated in (Dutch) research assessments[25], the *uniformity* of data provided – for instance by way of rehearsing dominant categories such as forms of knowledge transmission (research outputs) – is limiting in actualising the disciplinary diversity of the department. We found uniformity at odds in at least the following dimensions:

Designed to work with research outputs

One major limitation of both datasets is that they have been shaped (and are the result of) infrastructures for research information that value *research outputs*. At the same time, we observe shifting value registers that highlight *processes* as ingredients of evaluation. For instance, we found strong emphasis on *activities, the intended and unintended results of events, relationships and collaborations* as key spaces where research happens in our observations on transdisciplinarity.

If we relate this to the data sources we compared, we can see some limitations. For instance, in OpenAIRE data categories, there is emphasis on *Content* as a matter of outputs; a dedicated 'Outputs' category; *'Open Access',* which again refers to outputs; *'Indicators'* that reinforce an output orientation (e.g. counting the number of downloads, views, citations, and calculating the relative popularity of an *output* in a specific field). *Publication details* and *'Groups belonging to output'* act similarly, where *'Publication Details'* can be understood as specifications of outputs. There are some sub-types of the OpenAIRE output category that transcend this object-focus. These include the sub-types 'Non-textual', 'Event/activity', and 'Membership'. Unfortunately, none of these subtypes were populated for the dataset concerning Copernicus Institute. As to the CRIS, the categories also seem to speak to outputs with some exceptions. Namely organisational data, person-specific data and affiliations.

This is not surprising as 'lock-in' happens in technological development[26] where research information infrastructures are shaped as much by the socio-political environment as they themselves shape research (assessment) cultures. They don't exist in vacuums, so to speak. At the same time, transcending this lock-in requires considerable effort, especially as the information infrastructures are not designed to accommodate for other logics of valuation. In the movement for responsible research assessment, this is bypassed with the argument that qualitative evaluative knowledge should be enrolled and that quantitative indicators serve particular use cases, which we agree with. This pilot leaves us wondering nonetheless how we can think monitoring otherwise beyond the standard repertoire that is covered computationally already.

The two categories that are monitored by OpenAIRE that are not directly associated with outputs are *'Funding'* and *'Author'* information. As mentioned previously, these two categories are interesting in that they may allow the monitoring of (at least) formalised collaborations. Although informal collaborations – which often involve civil society in particular – would not appear by that approach.

Transdisciplinary research

In view of what this kind of data can tell us about transdisciplinary research at the department, we need to realise that a lot of customisation or further developments would be needed by, for instance, expanding research funding data in order to make high-level claims about collaborations (although this wouldn't say anything about the *quality* of the collaborations) or find alternative strategies to monitor and capture them. As an example, currently, the existence of funding data in combination with organisational data (i.e. affiliations), as well as memberships and activities can (potentially) approximate evaluations of transdisciplinary research (or relations across actors, to be more specific), especially in combination with non-academic or quasi-academic outputs, such as

[25] And is also advocated for in Dutch national strategic interventions such as 'Room for Everyone's talent'. See https://www.nwo.nl/en/position-paper-room-for-everyones-talent.

[26] See, for instance Rip, A. and R. Kemp (1998), 'Technological change', in: Human choice and climate change: resources and technology (S. Rayner and E. L. Malone, eds.), Battelle Press: Columbus, 327-399

'External Research Report'. Nonetheless, care should be exercised as to what can actually be said about academic and non-academic collaborations, as this data usually refers to the former and the data is also targeted at monitoring academic achievements.

In this sense, it is sensible that currently, stories of societal impact, stakeholder engagements and transdisciplinary research are accounted for through short narrative cases that introduce the reader to a transdisciplinary research setting. Such qualitative descriptions can be enriched by, e.g. research outputs that are attached to these processes, but cannot sufficiently indicate the depth and quality of collaborations across academic and non-academic actors.

Overall, the emergence of transdisciplinarity comes with a shift of focus from *outputs* to *interactions* (Brenninkmeijer 2022). This presents a major challenge to both the CRIS and OpenAIRE, at least if they are to be appropriated for evaluation. Both datasets are not designed for monitoring interactions and as such only hardly accommodate informing claim-making about transdisciplinary research. Despite this finding, the CRIS demonstrates that potentially, *activities* and *events* could be documented to make sense of transdisciplinary encounters provided there is a reliable list of interlocutors of research.

Accuracy

Another recurring issue that we found was accuracy. In particular, the data that we had to our availability suffered from duplicates (for instance because of the temporal changes to a research output, as mentioned earlier) or simply because of ambiguous naming and sourcing. Another issue of accuracy is the right classification of research objects into categories of research products (e.g. paper, preprint, review paper, other research product etc.). In most snapshot interviews, this issue was addressed by our respondents. Additionally, working through the comparative analysis also showed misclassifications (e.g. all 'other research products' were, in fact, papers) and we expect other ones as well for which we don't have the knowledge to determine.

We want to speak of accuracy as a matter of coverage. It was striking to us that OpenAIRE data had – in terms of unique items – a rather limited coverage of items associated with Copernicus Institute. We are unsure why that is (and inquiring about that would have transcended the pilot's focus). Inherently, this observation results in caution for how OpenAIRE data may be used as a source in (departmental) evaluations. For instance, understanding OpenAIRE as an *additional source* to CRIS data (next to other sources), could be interesting following an analysis of exactly what kind of data OpenAIRE can offer in contrast to what data exists in the CRIS.

Control and customisation

A fourth and final aspect includes the possibility for customisation. This organisational consideration concerns a fundamental logic that is woven into each of the two data sources. That is, the CRIS can be conceived of as a data space that can be locally adapted as its information is collected locally by the researchers themselves. It is an intra-organisational affair. OpenAIRE in contrast offers an aggregation service by which the control of what data is submitted lies beyond Utrecht University (or the department in any case). This results in dependencies where instead of addressing dedicated employees of Utrecht University whose job is to maintain the CRIS for questions and requests for customisation, this process would have to go by employees of OpenAIRE. For instance, in order to increase accuracy of research output classifications, the departmental secretariats could prompt researchers to submit their activities, events and publications via the CRIS, which would result in real-time changes to the database that could be appropriated for evaluations. These organisational strategies for increasing data accuracy can only be done indirectly with OpenAIRE and would seem to import more dependencies (e.g. of future 'data dumps' as we have experienced with OpenOrgs in our classification of the department).

Finally, these considerations are telling of the *scope* and *scale* for which the CRIS and OpenAIRE can be used, respectively. While localised assessments ask for control over the kind of data that becomes monitored and by extension appropriated for evaluation, national-level, or international or European monitoring of open science, for instance, requires other data affordances for which e.g. OpenAIRE is probably better equipped.

9. **Conclusion**

The goal of this pilot was to critically examine what data and research information infrastructures is, and can be, used for evaluating open science at a departmental level of a university using the SCOPE approach. Based on a contextual assessment of how open science matters at the Copernicus Institute of Sustainable Development, Utrecht University, we learned that a key pillar of what is valued in open science at this institute relates to societal impact creation and transdisciplinary research. Taken this as a starting point for choosing options for evaluation we examined how data from Utrecht University's local repository and OpenAIRE data can together inform evaluation of transdisciplinary research from the institute. For this, we defined a framework that focuses on the evaluation of research *interactions,* rather than *outputs.*

In particular, we probed evaluation options based on data that the OpenAIRE colleagues provided us about the Copernicus Institute of Sustainable Development after we defined the parent-child relationships in OpenOrgs. This data was complemented and compared with data from UU's local repository. Both datasets were assessed for completeness and relevance for evaluation purposes with a specific focus on transdisciplinary research.

Once we defined research processes relating to societal impact creation and transdisciplinarity as 'units' of evaluation, we saw a lack of relevant research information from both UU's local repository and OpenAIRE. That is particularly so with respect to data on involvement of non-academic stakeholders and when activities are highly collaborative and show high levels of engagement (e.g. 'co-designing a course for professionals' or 'co-producing knowledge with non-academic stakeholders'). On the one hand, this emphasises the importance of reflexive and intent use of quantitative research information for those purposes. On the other hand, it leads to new questions on how to inform alternative forms of evaluation such as narrative descriptions or impact case-that do justice to these types of activities. These questions were also discussed with researchers in the department using snapshots of relevant data which they complemented with a diverse range of outputs, activities and events that they valued as an integral part of transdisciplinarity.

During the process we provided real-life feedback to the researchers and engineers of OpenAIRE and those managing the local repository at Utrecht University, showcasing how they can adapt their tools and services to a changing research landscape. In that sense, the pilot allowed to sensitise the assumptions with which transition towards an Open Science-aware research assessment system is being approached and offers recommendations from the 'shop floor' of an academic institute focused on sustainability. In addition, the pilot also helped showcasing how transdisciplinary research processes can be an important unit of evaluation of an Open Science-aware research assessment system. Importantly, for the institute it generated valuable lessons about and experiences with OS-aware responsible research assessment. These lessons will be taken forward as part of future research assessment endeavours within the institute.

10. Appendix

10.1. Indicators used in the SEP self-evaluation

Target of assessment/monitoring	*Tools, services used**	*Data source(s) ***	*Challenges****
Composition of staff according to gender, nationality, age		*HRM System*	*See below*
	Share of female and male tenured staff		
	Share of Dutch, other EU, non-EU tenured staff		
	Nationalities of all staff and PhDs		
	Nationalities of tenured staff, non-tenured staff and PhDs		
	Age distribution tenured staff		
	Number of research staff (#) and research capacity (fte) according to function		
Funding and Expenditure		*(Beta) management system*	*See below*
	Total research capacity (fte) per year according to the source of funding		
	Total research capacity (fte) per year and in % according to source of funding		
	Total funding for research and education (kEUR) per year in %		
	Expenditure on research, (kEUR) per year		
	Expenditure on research, (kEUR) per year and in %		
	Prestigious grants and transdisciplinary projects (ten highlights); narrative	*Narrative; internal; questionnaire*	*See below*
Duration and success rate of the PhD Programme		*Management Systems*	*See below*
	PhD success rates, according to year of enrolment		
	External PhDs, by number, organisation and status		
	Type of first employer after graduation		
Research products for peers		*UU profile pages (internal); NARCIS (national)*	*See below*
	Research products for peers, per year and format		
	Publications between 2014-2020, by accessibility		
Use of research products by peers		*Bibliometric analysis with the support of the Utrecht University Library: indexed in Scopus and covered by SciVal)*	*See below*
	Number of citations		
	Average citation impact		
	Average number of citations per publication		
	Field Weighted Citation Impact (Top 10% and # and Top 1% and #)		
	Type of collaborations, based on affiliations of authors of articles, reviews and conference papers		

		Network visualisation of research collaborations		
		Research collaborations by number of joint publications with research organisations		
		Map of active research collaborations by countries of origin and number		
Marks of recognition from peers			*Form****** *the CRIS*	*See below*
	Six highlights of marks of recognition by category			
		Examples of awards and scholarly prizes		
		Examples of keynote lectures at major conf.		
		Organisation of international scientific conf.		
		Editorships and editorial boards		
		Memberships of academies		
		Visiting professors, visiting fellows		
List of awards by year, granting organisation and recipient			*the CRIS*	*See below*
	Awards (miscellaneous)			
	Nominations			
	Citation awards			
	Paper awards			
	Poster and presentation awards			
	PhD and MSc supervision, student awards			
Personal Grants			*the CRIS*	*See below*
Research products for societal target groups			*the CRIS*	*See below*
	Published policy reports (national, European, global)			*See below*
	Visits on Copernicus Institute website			*See below*
	Examples of (research) interaction with stakeholders		*Form; short text*	*See below*
	Example of research output for different audiences		*Form; short text*	*See below*
Use of research products by societal target groups				*See below*
	News mentions related to publications		*Altmetrics*	*See below*
	Newspaper items (national, international)			
	Radio appearances			
	Television appearances			
	Other news items			
	Mention of publications in Wikipedia articles			
	Mentions of publications in blogs			
	Social media mentions (Facebook and Twitter)			
	Mentions in policy documents			

	Publications co-financed by ministries in NL and abroad		
	Publications co-financed by the European Commission		
	Examples of co-creation, knowledge networks and platforms	*Narrative*	*See below*
	Examples of use of research products by societal target groups (international, national)	*Narrative*	*See below*
Marks of recognition from societal target groups		*Form*	*See below*
	Examples of marks of recognition by societal target groups (films, presentations, committee membership, etc.)		
Overview of PhD Courses offered		*Form*	*See below*
	PhD courses with contributions from Copernicus staff		
List of top publications Copernicus Institute 2014-2020		*Bibliometrics*	*See below*
	Highlights (cross-)topical themes to highlight breadth, interdisciplinarity, societal relevance, scientific impact, collaboration across career stages (top 10%, 5%, 1% best cited).		
Five detailed case studies to illustrate themes, interdisciplinarity, challenges and societal impact		*Internal; narrative; Form*	*See below*

Table 11: Indicators used in the SEP self-evaluation

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

***) include also if they are internal or external, national or international, if there are issues with quality of data, etc.*

****) limitations, restrictions, openness to interpretations, etc. to do with tools, services and/or data sources used*

*****Textual examples were collected by one of the team members by creating a shared document, meeting with senior staff members and asking them to share examples.*

******Underlined terms need to be confirmed by the respondents. We have asked, but no reply came in yet.*

10.2. Further methodological notes

<u>Deduplication</u>

Deduplication was based on the first 20 and last 20 characters of the titles. If these matched with the next item's title's beginning and end, the initial row was considered a duplicate and excluded from the original table.

<u>Data processing</u>

These are additions to he processing of OpenAIRE data in order to arrive at a workable and interpretable, deduplicated corpus of items attached to the Copernicus Institute.

<u>OpenAIRE</u>

1. Deduplication.
    a. In order to determine the duplicates, all titles are ordered from a-z. Then, the first 20 characters and last 20 characters of the title are taken. If these match with the next row, then the current row is considered to be a duplicate of the next row. In such a case, the current row gets a "true" at the end, and this row is excluded from the original table. This methodology is imperfect. For example, looking at the title starting with "The PMIP4 contribution to CMIP6", it is evident that there are 4 rows with the same title, yet the ending of the title has different characters and

therefore it is not observed. It is difficult to perform a quick analysis that is 100% accurate, and therefore this methodology is taken to be sufficiently functional to address that duplicate rows do in fact exist.

   b. In the duplicate rows, it happens that there exists data in one of the two rows that does not exist in the other row and vice versa. Therefore, in calculating the total occupancy of the table, the total number of instances are counted, but capped at 1095. This means that for any duplicate row, if only one of the two has a datapoint, the total will end up at 1095. However, if both duplicate rows contain a datapoint, then for the rest of the datapoints there will be an overestimation. This methodology is not perfect, but it is a simple calculation and still provides a good estimate for the total occupancy of the table.

2. Missing data based on triangulation
   a. In order to test the completeness of the data, we extracted OpenAIRE data from 2020-aug2025 and a sample of 21 titles from the CWTS database that then were compared to the titles of the data set retrieved from OpenAIRE.
      From this sample, five titles did not occur in the OpenAIRE data. This makes us question the completeness, and in extension quality, of the OpenAIRE corpus.

3. Notes on specific columns
   a. Subtitle is in about 90% of the cases only the part of the main title that comes after the ":", which makes it a duplicate of the main title. Only in 7 of the 138 occurrences with a subtitle, the subtitle was unique.

   b. Description sometimes has html coding ("<p>") or (???) coding ("<jats:p>"). Unnecessary.

   c. Language is very often Undetermined. Making a column for this is odd, since more often the values are left empty if the real data (e.g. if the journal is not known, it doesn't say "undetermined" but is simply left empty).

   d. Formats have a very strange format of the table. Most occurring are "application/pdf" and "image/pdf", but there are also some that have the format "text" and "pdf". Some of them are (page) numbers and others are variations on the "../pdf" or "text/..".

   e. BestAccessRight is always the same value, or empty.

   f. Starting page sometimes contains letters, and sometimes very high values (e.g. 129758).

   g. Author names have strange characters in them. Either not properly retrieved or the source material is wrong. Could be given attention.

   h. Funding data seems very accurate, especially when the focusing on the projects. The "publiclyFunded" column is very scarcely populated and mainly contains FALSE values, even when the funder is "EU".

   i. The contributors and organisations are poorly aligned. In the sense that a result with many organisations sometimes only has one contributor.

   j. The countries do not always overlap with the countries that could be derived from the "organisations" column. The second result (Title: Conceptual framework for the study of food waste generation and prevention in the hospitality sector), for example, has as countries Malaysia and Netherlands, where the UK should really also be in there (universities Leeds and De Montfort).

<u>the CRIS</u>

1. Obtaining API Key and authorisation
    a. See the the CRIS API documentation here[27]
    b. We received the API key from the person responsible for the CRIS at Utrecht University

2. Ricgraph
    a. Installed 'Virtualbox', and a virtual machine with Ubuntu 24.04.02
    b. Installed the make command
    c. Then followed Ricgraph documentation from point 1.1 until step 4.0[28]
    d. Afterwards, we checked the initiation file (ricgraph.ini) and edited the UU URL portal and the UU URL API key
    e. Edited harvest_the CRIS_to_ricgraph.py so that running it would provide the dataset of all entries from only 2025 to test for errors
    f. Afterwards, the harvest file was edited to harvest the data per year from 2020-2025.
    g. The jsons that existed in the /ricgraph_venv/harvest folder were transferred outside of the virtual environment for transcription to csv

3. Filtering Copernicus Institute (department)
    a. Filtering was done by first determining the organisation names that were part of the Copernicus Institute by filtering out all organisation names containing the word "university" and then manually looking through the dataset to check the names that could be part of the GEO faculty.
    b. This was done by including any words such as "geography", "innovation", "sustainability". Then, from that list, samples of authors were taken to assess whether that person was part of the CISD.
    c. If that was the case, then the organisation name that the person was ascribed to was also considered part of the CISD.
    d. This was later verified using the organisations csv file that was included in the harvest folder. This lead to both having 28 organisation names under the CISD, therefore having double verified that the data was correct.

10.3. Potential objects of evaluation of the OpenAIRE dataset

Category	Column (ideal)	Content
Identifier	id	OpenAIRE interal Identifier
	doi	Digital object identifier
	originalIds	Other identifiers for OpenAIRE internal use
	pids	persistent identifiers
Content	mainTitle	Title of the output
	subTitle	Subtitle of the output
	description	Summary of the output
	subjects	Keywords related to the content of the output

[27] See https://research-portal.uu.nl/ws/api/documentation/user-guide/api-keys.html. Accessed 14.05.2025
[28] See https://docs.ricgraph.eu/docs/ricgraph_install_configure.html. Accessed 14.05.2025

	language	Language the output is in
	coverages	The time and/or locations covered by the publication
	geoLocations	Geolocation information
Output	type	See output types
	sub_type	See output types
	formats	See output types (https://www.iana.org/assignments/media-types/media-types.xhtml)
Open Access	isGreen	Whether the output has at least the Green Label for Open Access
	isInDiamondJournal	Whether the resource is in a Diamond journal
	BestAccessRightCode	The code describing the level of Open Access (according to COAR https://vocabularies.coar-repositories.org/documentation/access_rights/)
	BestAccessRightLabel	The label describing the level of Open Access (from COAR label)
	openAccessColor	The Open Access colour of the output (levels: gold, hybrid or bronze)
	embargoEndDate	Date when embargo ends and result becomes open access
Indicators	downloads	Number of downloads an output has
	views	Number of views an output has
	citationCount	Number of citations an output has
	impulse	Initial momentum of output, based on citation network
	influence	Popularity of output over the whole period, based on citation network
	popularity	Current popularity of output, based on citation network
Publication details	publisher	The name of the entity that holds, archives, publishes prints, distributes, releases, issues, or produces the resource.
	publicationDate	The date on which the resource is published
	journal (from container)	Name of journal or conference
	volume	Volume in which the resource appears
	startingPage	Page on which the resource starts
	endingPage	Page on which the resource ends
	collectedFrom	Information about the sources from which the record has been collected
	dateOfCollection	Date on which the output is collected for the document containing the data
	lastUpdateTimeStamp	Timestamp of last update of the record in OpenAIRE
	sources	A Reference to a resource from which the present resource is derived
	version	Version of the result
Funding	publiclyFunded	Whether the resource is publicly funded
	Project (definition)	List of projects (i.e. grants) that (co-)funded the production of the research results
	acronym	Abbreviated project title
	project_title	Project title
	funding_stream	Stream of funding (e.g. for European Commission can be H2020 or FP7)
	jurisdiction	Geographical jurisdiction (e.g. for European Commission is EU, for Croatian Science Foundation is HR)
	funder_name	The name of the funder (e.g. European Commission)
Author	authorFullName	The full name of an author
	authorOrcid	The ORCID identifier of an author
	authorRank	The rank of an author
Groups belonging to output	countries	The different countries related to the output
	communities	The OpenAIRE communities the output is part of
	contributors	Contributors for the result
	organisations	List of organizations in an affiliation relation with the research results

Table 13: Complete overview of item categories in OpenAIRE dataset

11. References

- *Brenninkmeijer, Jonna. 2022. "Achieving Societal and Academic Impacts of Research: A Comparison of Networks, Values, and Strategies." Science and Public Policy 49 (5): 728–38. https://doi.org/10.1093/scipol/scac022.*

- *Fritz, Livia, and Claudia R. Binder. 2020. "Whose Knowledge, Whose Values? An Empirical Analysis of Power in Transdisciplinary Sustainability Research." European Journal of Futures Research 8 (1): 3. https://doi.org/10.1186/s40309-020-0161-4.*

- *Himanen, Laura, and IIRIS LIINAMAA. 2024. "GraspOS Deliverable D5.2: Pilot Findings and Progress Report," August. https://doi.org/10.5281/ZENODO.13629146.*

- *International Network Of Research Management Societies-Research Evaluation Group. 2023. "The SCOPE Framework." University of Melbourne. https://doi.org/10.26188/21919527.V1.*

- *Janssen, Rik D.T. 2024. "Ricgraph: A Flexible and Extensible Graph to Explore Research in Context from Various Systems." SoftwareX 26 (May):101736. https://doi.org/10.1016/j.softx.2024.101736.*

- *Laura Himanen. 2023. "GraspOS Deliverable 5.1 'Report on Pilot Setup, Current Practices and Initial Requirements,'" August. https://doi.org/10.5281/ZENODO.11091718.*

- *Mauri, Michele, Tommaso Elli, Giorgio Caviglia, Giorgio Uboldi, and Matteo Azzi. 2017. "RAWGraphs: A Visualisation Platform to Create Open Outputs." In Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter, 1–5. Cagliari Italy: ACM. https://doi.org/10.1145/3125571.3125585.*

- *Miedema, Frank. 2022. Open Science: The Very Idea. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-024-2115-6.*

- *Rafols, Ismael, Ingeborg Meijer, and Jordi Molas-Gallart. 2024. "Monitoring Open Science as Transformative Change: Towards a Systemic Framework." F1000Research 13 (April):320. https://doi.org/10.12688/f1000research.148290.1.*

- *Schrögel, Philipp, and Alma Kolleck. 2019. "The Many Faces of Participation in Science: Literature Review and Proposal for a Three-Dimensional Framework." Science & Technology Studies (Tampere, Finland) 32 (2): 77-.*

- *UNESCO. 2021. "UNESCO Recommendation on Open Science." UNESCO. https://doi.org/10.54677/MNMH8546.*

- *VSNU, KNAW, and NWO. 2020. "Strategy Evaluation Protocol 2021-2027." Den Haag.*

- *VSNU, NFU, KNAW, NWO, and ZonMw. 2019. "Room for Everyone's Talent. Towards a New Balance in Recognising and Rewarding Academics."*