Best practice example(s) of alternative assessment approaches

Deliverable 3.1

Work Package: 3 – Alternative assessment approaches

Version: 1.0 (25 February 2022)

Editor: University of Padova

Arqus Research & Innovation Project

(Grant agreement No 101017448)

Contents

1. Introduction

1.1 About the project

The Arqus Research & Innovation project (Arqus R.I.) is a H2020-funded project that aims to enhance the research and innovation dimensions of the member institutions of the Arqus European University Alliance[1]. The Alliance brings together seven longstanding prestigious comprehensive research-intensive universities that share extensive experience in joint projects and a similar profile as internationalized institutions with deep regional engagement.

The Arqus Research & Innovation project aims to complement and build on efforts already underway to foster scientific cooperation within the Arqus European University Alliance, and to seek synergies between the education, research and innovation dimensions of the alliance.

This project is organised around three major challenges, all intended to nurture sustainable strengthened collaboration among partners.

- **Challenge 1** addresses Joint Science and Innovation, with a drive to building an Arqus research community, and to designing a roadmap for transformational excellence.
- **Challenge 2** focuses on New Perspectives on Research, seeking: (1) to strengthen the human capital in the partner institutions discerning and helping the future evolution of research assessment practices, to make them fitter to attract, retain, and recognise top talent by the extent, depth, and diversity of individual research projects; and (2) to re-think knowledge transfer, in particular, but not exclusively, in the Social Sciences, Humanities and the Arts.
- **Challenge 3,** which calls for Openness to Society, seeks and promotes the cultural shift brought about by Open and Citizen Science.

1.2 About WP3 - Alternative Assessment Approaches

Work Package 3 (WP3 in the sequel) is part of "Challenge 2: New Perspectives on Research" of the Arqus R.I. project, to address aspect (1) of that challenge, with focus on transformational excellence, seen from the perspective of understanding, recognition, evaluation, and reward, which are the dimensions of research assessment approaches. This report is the external product of such effort at the end of Year 1 of the project.

The actors in WP3 have no mandate to make policy in the regard of research assessment approaches, whose responsibility strictly pertains to the governance of the individual institutions at the local level, and to the corresponding political bodies at regional, national and continental levels. The role of WP3 is more simply to feed reflection on the shortcomings of the current research assessment practices, particularly in the (cultural and normative) context of the project partners, and to identify directions of interest, which might determine desirable and sustainable changes to such practices at some if not all of the levels of the concerned authorities.

This report is devoted to presenting results pertinent to Goal 3.1 of WP3: "Sharing best practice on assessment systems and criteria for academic careers". The relevant work has been carried out as part of Task 3.1: "Non-bibliometric assessment criteria analysis", devoted to fostering critical analysis of

---

[1] https://www.arqus-alliance.eu/

assessment practices in place at the institutions of the Alliance as determined by local, regional, national policies, or happening outside of the Alliance, but with sufficient merit to be deemed of interest.

The reasons for directing work to Goal 3.1 and Task 3.1 as above recalled reflects the observation that questions are being raised at various levels, from individual researchers to institutions, to groups of them even across national borders, about the fairness and the accuracy warranted in the recognition of researchers' output and its rewarding for recruitment and promotion [2]. A range of initiatives have been undertaken in various parts of the world, including in Europe, to attempt to mend the perceived distortions and help researchers understand and adjust to the anticipated changes.

One common trait of the shortcomings of the dominant model of research assessment is the inward and single-sided nature of the traditional criteria. Built on the notion of peer-review (intrinsic and essential to science per se) and often exceedingly quantitative, the dominant criteria are ill-equipped to capture the diversity of research work and products, to positively relate to the world outside, with its variety of stakeholders (socio-economic bodies, innovation actors, policy makers, and the public at large), and to reflect the role that academic researchers play in it. The acknowledgement of this deficiency and the quest for corrective measures have driven the activities performed by WP3 for the production of this report.

At the same time, the stringent limitations in the time, effort, and authority availed to its operation required the WP3 team to carefully scope the work to be conducted so that it would produce results (especially stimuli) of useful value to the respective institutions and, possibly, to the global debate.

In order to understand the scope given to the work reported in this document, it is important to recall the exclusion and inclusion principles specifically adopted in that regard.

First and foremost, we had to come to a sound and shared understanding of the notion of "alternative" that reads in the very title of WP3: "Alternative Assessment Approaches".

Streamlining complex differences into simple numbers that can easily be compared is tempting. Drawing quantitative indicators obviously help to draw comparisons, capture differences, and observe change gradients. And it has done so well in certain circumstances, as noted in Appendix A.3 to this document. The downside of that approach, however, has been to force a subtle yet strong effect of unidimensional standardization on the type of research being conducted and of the products being exposed by researchers. Such standardization fails to recognize and reward diversity, which instead is natural to research, and effectively discourages it.

For this reason, we rejected all interpretations of "alternative" that merely implied or suggested "different calculations" of quantitative indicators for the same set of traditional (standardized) research products. Conversely, we favoured those approaches that entail explicit support for alternative research practices, viz. diversity, most particularly (but not exclusively) as they emanate from Open Science initiatives. With this, we do not want to contrast "alternative to traditional" research as if the latter were to be abandoned or disgraced. Actively embracing diversity is naturally inclusive as opposed to sectarian. It widens the spectrum of practices and products, and, like in nature, it fosters a richer and more robust ecosystem.

To use its scant resources at best, WP3 has chosen to focus its attention on pertinent experiences that have made some inroad into partner institutions, as well as on practices that, in our view, reflect proceeds

---

[2] For a recent and comprehensive discussion of defects in the assessment of research and researchers, see for example: https://www.leru.org/files/Publications/PP-APathway-Execituve-Summary.pdf.

of the current global debate on new and multidimensional quantitative and qualitative criteria for research assessment, for products and careers, in different disciplinary and institutional contexts.

In keeping with the scope of the funding program to which this project belongs, WP3 focuses especially on practices that promote the implementation of "science with and for society", helping to widen the focus of research efforts and the recognition of research products, from traditional forms of publication, to other "modes of delivery", public outreach, as well as knowledge transfer.

In seeking to promote the widening of the recognition of research efforts and products that qualify for assessment procedures, we decided to conduct our investigation in two complementary directions:

- On the one hand, we reflected on the top-down side of the matter of research assessment approaches, where institutional policies are conceived and promulgated to promote alternative assessment approaches that contemplate a wider understanding and valuation research work.
- On the other hand, we also followed a bottom-up route, singling out researcher-level reflections, initiatives or efforts to conduct or promote research that can be categorized as Open Science, public outreach, and knowledge transfer, to help reason on how it can be considered, fairly and accurately, in novel assessment approaches.

Best practices in both regards, selected via shared reflection, self-analysis, solicitation, and promotion by the partner institutions, but not limited to partner actors, were examined in the making of this report. A series of webinars was organized to this end, which addressed different target groups within each institution (senior management and decision-makers, senior academics, early stage academics, research support staff).

To support the understanding and the scrutiny of the surveyed initiatives, we developed two types questionnaires, which we administered to all of the Alliance partners. A top-down questionnaire investigated the existence of alternative approaches to research assessment in place, currently or prospectively, at the local level. A bottom-up questionnaire helped investigate alternative practices of doing research. Both questionnaires are reproduced in Appendices A.1 and A.2 respectively, while their outcomes are summarised and discussed in the core part of this document.

The remainder of this document is organized as follows:

- Section 2 discusses and justifies the scope of this document and the work that has come to feed it
- Section 3 presents and reviews exemplary top-down initiatives with bearing on Alternative Assessment Approaches for any of the Arqus Alliance institutions.
- Section 4 presents and reviews exemplary bottom-up initiates undertaken at any of the project partners, which we deem to have useful potential to feed the formation of Alternative Assessment Approaches.
- Section 5 draws conclusions from this work and outlines directions for future work in this ambit.
- Appendix A presents supplemental material.

By reviewing situations and initiatives that may promote reflection and stimulate improvement actions at the Arqus Alliance partners, this document also feeds the production of deliverable report D3.2, entitled "An internal Arqus discussion paper on alternative approaches potentially leading to input to the ongoing debate at national and European level", and not intended for public circulation. Specific text boxes in the sequel of this report mark and present questions that arise from the initiatives discussed in the vicinity of

that box, and that the project team will pose to the senior officials of their own institutions. The D3.2 deliverable report will collate the responses obtained to those questions and draw conclusions and prospects from them.

2. Scoping the study

2.1 A view of the general context (with attention to Open Science)

The theme of academic research assessment, for the purposes of recruitment, promotion, or funding, has gained considerable attention in the last couple of decades, in part owing to the numerous shortcomings of the current practices. Testimonies of such attention include the (1) European University Association resources on Academic Career Assessment [3], which encompass DORA, the Declaration on Research Assessment, and its SPACE rubric [4], and (2) the Open-Science Career Assessment Matrix (OS-CAM) [5] by the Directorate-General for Research and Innovation at the European Commission, corresponded by equivalent national initiatives, for example in Norway [6].

The main shortcomings observed in traditional assessment practices, exclusively centred on quantitative bibliometric indicators, manifest in the narrowness, bias, and distortions that they may incur. Efforts have been made in some countries to correct or remedy some of those defects. Those efforts may provide useful guidance or insight for others to consider. We shall discuss lessons learned from some such efforts in Sections 3 and 4 of this report.

The main strand of corrective forces applied to the status quo proposes a shift toward a balance between quantitative and qualitative assessments, in addition to greater awareness of the risks of bias, discrimination, and other forms of unfair treatment.

Attention to Open Science is one major driver behind said shift, especially for a notion of Open Science understood as *a set of good practices, principles and goals that aims to reduce barriers in all aspects of the research process for the benefit of research and society. Open Science encompasses transparency, accessibility, reproducibility, comprehensibility, trustworthiness, participation and inclusiveness in all parts of the research process. Open Science increases the efficiency of research by making scientific knowledge findable, accessible, interoperable and reusable, thereby accelerating progress and discoveries for the common good* [7].

Open Science as outlined above is a multi-faceted concept, which can be pictorially depicted as shown in Figure 1. For the purposes of this document, is understood to include Citizen Science, Open Education, and Open Innovation, which are primary dimensions of research, which very much concern the researchers themselves.

---

[3] https://eua.eu/component/tags/tag/85-academic-career-assessment.html

[4] https://sfdora.org/resource/space-to-evolve-academic-assessment-a-rubric-for-analyzing-institutional-conditions-and-progress-indicators/

[5] https://op.europa.eu/en/publication-detail/-/publication/47a3a330-c9cb-11e7-8e69-01aa75ed71a1/language-en

[6] https://www.uhr.no/en/front-page-carousel/nor-cam-a-toolbox-for-recognition-and-rewards-in-academic-careers.5780.aspx

[7] This definition of Open Science is drawn from a the "Openness Position Paper" that has been produced by Action Line 6.2 of the Arqus Alliance in their Erasmus+ project, to which this report aligns fully.
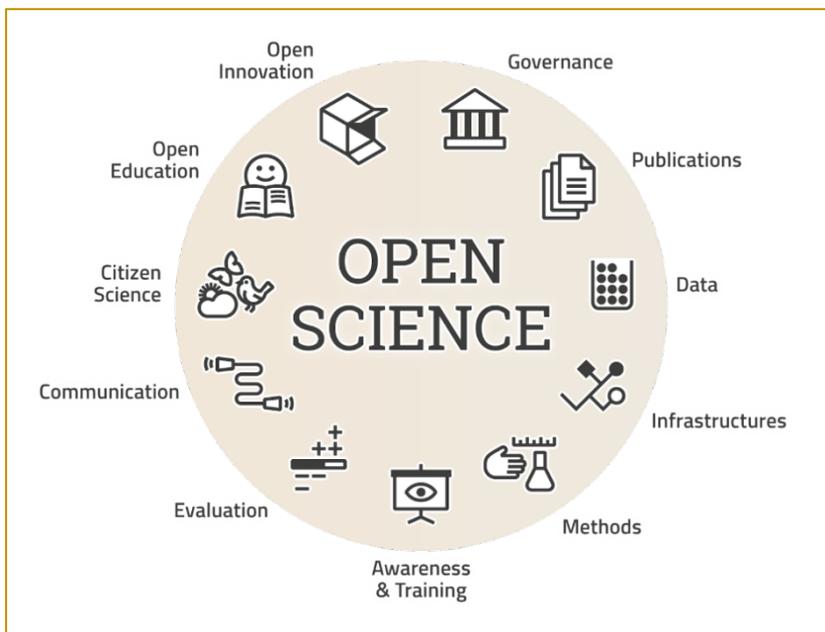
Figure 1: A pictorial representation of the multiple facets encompassed by the notion of Open Science.

- *Citizen Science* promotes the cooperation of science and society to generate new knowledge, favouring the inclusion of citizens' perspectives in the formulation of research questions, thus increasing the societal relevance of research products[8], [9]. Citizen Science as a method markedly emerged around 2010, when digital platforms made it possible to make volunteer contributions in science more visible and viable[10].
- *Open Education* widens access and participation to the learning process by means of high-quality, easy-access resources accessible to everyone.
- *Open Innovation* interlaces actors from the business, academic, civil society, or government sector in the conception of innovation, facilitating knowledge circulation, addressing societal challenges, and facilitating economic growth.

Making science even more publicly accountable and accessible has been a key and constant concern in science-related policies at large. Accountability to taxpayers was one of the political needs used to justify the introduction of the Research Assessment Exercise started in the UK in the late 1980s[11]. An issue made relevant also by the general reduction in university fundings.

While such need is still present to date, a new challenge has emerged: closing the gap between science and public opinion in order to avoid misconceptions and the spread of misinformation. Admittedly,

---

[8] F. Heigl, B. Kieslinger, et al. (2019). Opinion: Toward an international definition of citizen science. Proceedings of the National Academy of Sciences of the United States of America, 116(17), 8089. https://doi.org/10.1073/PNAS.1903393116

[9] C. Kullenberg, D. Kasperowski, What Is Citizen Science? – A Scientometric Meta-Analysis, PLOS ONE, 2016, https://doi.org/10.1371/journal.pone.0147152

[10] C. Kullenberg, D. Kasperowski, What Is Citizen Science? – A Scientometric Meta-Analysis
Published: January 14, 2016, https://doi.org/10.1371/journal.pone.0147152

[11] V. Bence, C. Oppenheim. The Evolution of the UK's Research Assessment Exercise: Publications, Performance and Perceptions (2006). Journal of Educational Administration and History 37(2):137-155. https://doi.org/10.1080/00220620500211189

however, this is such a complex phenomenon, which takes place simultaneously at several levels (individual, groups, societal) and involves a range of actors, that Open and Citizen Science policies might be only a part of a greater action plan.

The European Commission has recently adopted an agenda centred on Open Science and Citizen Science initiatives[12], to foster the accessibility of scientific findings and their transfer to the economic environment as soon as results are available. Such policies are inserted in the long-term goal to transform the European society into the most competitive and dynamic knowledge-based economy in the world, capable of sustainable economic growth with more and better jobs and greater social cohesion[13]. The primary aim of these policies is to facilitate knowledge transfer, while the secondary aim is to make science more inclusive towards citizens. The political need to embrace new frameworks for making and communicating science is also a response to anti-scientific and credulous trends that frequently take place in contemporary society.

Among other things, Open Science practices seek the reproducibility and replicability of research findings[14], which makes scientific procedures more transparent and scrutable by actors, both scientific or not. Open data and publication in Open Access journals are typical practices inside the Open Science framework. Reproducibility of results from experiments, especially from life sciences like medicine, psychology, chemistry and biology, and increasingly so for computer science, has become extremely difficult[15]. The remedy to such deficiencies needs to raise the bar on the definition, the scrutiny, and the promotion of high-quality data. This effort cannot be limited to offering authoritative portals for open data, such as Springer Nature [16] and IEEEDataPort [17], would need as much peer review as the corresponding research paper. The "Findable Accessible Interoperable Reusable" (FAIR) principles [18] stipulate the base expectations to be promoted, sought and assured by institutional repositories: initiatives to this end, taken by institutions member of the Arqus Alliance, are discussed in Section 3 of this document.

Open and Citizen Science practices have been often, mistakenly, considered outside the classic boundaries of how science is done and evaluated. Indeed, they may be regarded as alternative to the traditional way of doing and organizing science. This denotation, however, instead of drawing negative prejudice, suggests different methodological approaches with respect to traditional science, which require attention, understanding and recognition in ex-ante and ex-post academic evaluation practices. Some of these concerns have been captured in a range of initiatives reviewed in Sections 3 and 4 of this document.

---

[12] https://ec.europa.eu/programmes/horizon2020/en/h2020-section/science-and-society

[13] European Parliament – Directorate for the Planning of Parliamentary Business, European Council, 23-24 March 2000, Lisbon, 1s2000en.pdf (europa.eu)

[14] S. Crüwell, J. van Doorn, et al., 7 Easy Steps to Open Science: An Annotated Reading List, PsyArXiV Preprints, https://doi.org/10.31234/osf.io/cfzyx

[15] A poll of 1,500 scientists conducted by Nature in 2016 reported that 70% of them had failed to reproduce at least one other scientist's experiment (including 87% of chemists, 77% of biologists, 69% of physicist and engineers, 67% of medical researchers, 64% of earth and environmental scientists, and 62% of all others), while 50% had failed to reproduce one of their own experiments, and less than 20% had ever been contacted by another researcher unable to reproduce their work (Baker, 2016).

[16] https://www.springernature.com/gp/open-research/about

[17] https://ieee-dataport.org/

[18] M. Wilkinson, M. Dumontier, I. Aalbersberg. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

Research assessment concerns recruitment, funding, and promotion. The awarding of funding naturally correlates with the dynamics of promotion and career progression. To regulate the assignment of funding to groups and individuals, several European countries have adopted Performance based Research Funding Schemes (PRFSs). These are complex systems designed to evaluate universities and public research centres in a given country according to outputs and outcomes rather than processes and structures[19] using bibliometrics, peer review, or a combination of the two. In recent years, several PRFSs have expanded their set of evaluation criteria, in the intent of making them better capable of "sensing diversity", thereby measuring a wider range of significant aspects of research work. Paradigmatic in this regard is the case of the Research Assessment Exercise in Lithuania, where scientific activity is measured through three domains (academic achievements, visibility and impact; academic leadership; other academic and social activities), each reflected by multiple indicators. Similar practices in this regard have been adopted by members of the Arqus alliance, as we discuss in Section 3.

As of late, a wave of criticism has risen against the effects more than the principles of PRFSs[20]. The central argument of the critique is that such systems tend to concentrate resources in those institutions that are already performing well, directing further funding to already well-fed groups, also favouring incremental research within the boundaries of existing theoretical and methodological paradigms, again disfavouring diversity and discouraging long-term research[21]. Similar observations were noted in the proceedings of the webinar-and-debate series run by WP3 in the Year 1 of the project, and summarised in the Appendix to this document.

On the whole, in working on this report, we have contemplated Alternative Assessment Approaches as intentions, practices, procedures and lessons learned connected with the recognition, evaluation, and dissemination of a spectrum of research products and scientific profiles wider than classic peer-reviewed papers published in journal venues with high ranking according to established classification criteria. Such Alternative Assessment practices should be regarded as a tool designed especially to foster transparency, diversity, interdisciplinarity, and to promote innovative ways of organizing and communicating science.

2.2 Zooming in on the Arqus R&I partner institutions

Before reasoning on which alternative assessment approaches might show potential to trigger improvement actions in the Arqus Alliance institutions, we thought we should first appreciate the status quo of those partners with regard to the various dimensions of the subject matter.

In essence, we wanted the WP3 members to help themselves and their hierarchy to gain a fresh factual understanding of what those institutions currently do, formally and informally, in the regard of the research assessment, the rationale and the dynamics of it as far as they can be discerned via informal dialogue and testimonies drawn from their scientific personnel or research support staff.

We performed that reconnaissance work in three distinct and complementary steps.

---

[19] D. Hicks, Performance-based university research funding systems, Research Policy, 41(2), 2012, 251-261, https://doi.org/10.1016/j.respol.2011.09.007
[20] D. Hicks, P. Wouters, L. Waltman, et al. Bibliometrics: The Leiden Manifesto for research metrics. Nature 520, 429–431 (2015). https://doi.org/10.1038/520429a
[21] D. Gillies, How Should Research be Organised? An Alternative to the UK Research Assessment Exercise, From Knowledge to Wisdom: Studies in the Thought of Nicholas Maxwell, Ed. Leemon McHenry, Ontos Verlag, 2009, 147-168, https://discovery.ucl.ac.uk/id/eprint/16521/1/16521.pdf

- Early in Year 1 of the project, we submitted questionnaires to the Arqus R&I partner universities, seeking to determine, broadly: (1) who, for individuals or governing bodies, is tasked or enabled to carry out reflections on Alternative Assessment Approaches at those institutions; (2) who such reflections involve for level of seniority in post, whether only tenured professors or also on-tenure researchers from PhDs upward; and (3) which practices and actions for alternative assessment are favoured in those efforts. The output of those questionnaires is summarized in Section 2.3 below as part of the background to the subsequent contents of this document.
- We had interviews with experts and emblematic figures from the midst of partner institutions, whom we could easily reach out to, to help identify elements of good practices that could be replicable at the local level across the Arqus Alliance. Each invited expert was requested to give a presentation on a relevant pre-determined topic, coherent with their specific background and expertise. In order to help steer and focus those presentations, we developed two questionnaires and forwarded them to the invited experts ahead of the webinar: one questionnaire concerned their awareness of and view on top-down initiatives in place at their institutions (see Appendix A.1) , the other the same for bottom-up actions (see Appendix A.2). The top-down questionnaire is structured in three sections investigating the expert's personal beliefs on Alternative Assessment Approaches, the state of the art about their adoption at the country and university level (institutional context), and recommendations for their implementation. The bottom-up questionnaire includes two sections, which concern the functioning of third-mission and outreach activities and the perceived degree of transferability of the corresponding practices. At the end of the webinar series, we held two debates with three speakers each, to discuss our preliminary findings from said webinars, and the pros and cons of the practices that were most discussed. The output of this line of activity is summarized in contributions to Sections 3 or 4 of this document, as appropriate.
- We invited partners to volunteer in-depth reflections on top-down (institutionalized) or bottom-up (voluntary) practices that they have recent experience with in the ambit of interest to WP3. We did not want to produce a full inventory listing of initiatives, as we had no business satisfying the pride of particular organizations. We wanted examples that had traits of transferability and might be deemed to stimulate the interest of other institutions in the Arqus Alliance and perhaps also outside of it. We present and discuss such examples in Sections 3 and 4, which form the central part of this document.

2.3 Returns from questionnaires submitted to partner institutions

Table 1 summarises the overall picture drawn from such enquiry, which reports the Arqus R&I project partner institutions as active at various levels of intensity in reflections on Alternative Assessment Approaches, although with considerable difference in the institutional figures who carry out such reflections. For some partners, discussions are taking place in both informal and formal occasions, involving practitioners, university board members and administrative staff from central service offices. What this birds-eye picture reveals is that the partners of the Arqus Alliance are actively reflecting on Alternative Assessment Approaches, which is a solid and promising starting point for further initiatives in this domain.

Table 1: How reflection on Alternative Assessment Approaches is carried out at the Arqus Alliance institutions.

	At your institution, reflection on Alternative Assessment Approaches is carried out at:		
Partner Institution	Informal meetings among interested practitioners	Rector's board members	Administrative (central service office) staff
Bergen	×	×	×
Granada	×		
Graz	×	×	×
Leipzig	×		×
Lyon	×	×	×
Padova		×	×
Vilnius		×	×

The next question addressed in the questionnaire tried to determine how much the respective academic communities were involved in reflections on Alternative Assessment Approaches. The corresponding findings are reported in Table 2, which shows that the Arqus R&I project partners appear to be quite inclusive. While all involved institutions report engaging tenured faculty into this kind of discussions, half of them also declare to involve early-career researchers (PhD candidates, postdoc researchers, and on-tenure researchers) in aspects of that process. The involvement of tenured faculty is not a surprise, since they are entitled to occupy governing positions inside academia. The engagement of early-career researchers may be read as a very positive indicator, although the extent of that engagement varies considerably for the nature of the discussion opportunities and the scope of them. Some partners report that the involvement of those figures for them happens at the institutional level, in specific committees that include on-tenure representatives. Those profiles hold the most vulnerable positions in academia. They are the ones who experience the highest level of uncertainty on the development of their professional career resulting from arising from fixed-term employment contracts and contracting opportunities for positions in the teaching and research positions. For these reasons, their involvement is a commendable practice to consider when institutions reflect on practices for alternative assessment approaches. At the same time, it is observed that early-career researchers are naturally inclined and quick at taking up innovative practices like OA publishing and OS, with distinct interest in open research artifacts, both for their own use and for result dissemination.

Table 2: Who is tasked/invited to reflect on Alternative Assessment Approaches at the Arqus Alliance institutions.

	At your institution, reflections on Alternative Assessment Approaches involve:			
Partner Institution	PhD candidates	Postdocs	On-tenure researchers	Tenured Faculty
Bergen	×	×	×	×
Granada	×	×	×	×
Graz	×	×	×	×
Leipzig			×	×
Lyon	×	×		×
Padova				×
Vilnius				×

When asked about practices that can contribute to cultural change in research assessment, partners were asked to review their actions devoted to scientific communication (outside dimension) and discussions or focus groups among researchers (inside dimension). The majority of partners indicated various initiatives connected to the outside dimension, with structured support (e.g. dedicated offices) for dissemination and valorisation activities towards the general public.

Table 3 summarises the kind of Alternative Assessment initiatives that the Arqus Alliance institutions are favouring or are adopting among: (a) narrative curricula[22]; (b) attention to ability to obtain funds from external sources; (c) Open Access publication; (d) open data and artifacts for reproducibility. Some partners are adopting narrative curricula for selection procedures of academic personnel. Most frequently, Narrative Curricula describe future research intentions in application for professorship, as a tool of ex-ante evaluation of future contributions to the discipline. They may also serve as ex-post evaluation tools, as self-report of past activities in a given time frame for the purposes of classic research evaluation systems.

---

[22] As of late, the term "narrative curriculum [vitae]" has become known in the context of research assessment to signify that the quality and impact of the research carried out by an individual should be deduced *not solely* from by that person's academic publication profile, but also in a variety of other activities, including contributions to the societal debate, economic valorisation of research results, and as many others as required to cover all pertinent facets of research. The recognition of "narrative curricula" in research evaluation should allow researchers to make personal choices in the way they shape their research trajectory, and should allow them to describe it *narratively* as opposed to quantitatively.

Distinct attention to the ability to obtain funds from private partners clearly emerges from the analysis of recruitment procedures.

As regards OS practices such as publication in Open Access journals and open data, the situation at Arqus R&I partners appears to be promising. OA publication, especially in non-paid venues or official non-commercial repositories is broadly encouraged. Some universities have developed, or are developing, infrastructure dedicated to sustain such a change of paradigm. Most partners have a university policy for Open Access publication, paired with services from central offices and economic resources. When looking at practices for OS, such as open data and accessible coding for reproducible research, however, only half of the partners reported positively.

Table 3: Which kind of Alternative Assessment Initiatives the Arqus R&I institutions are actively promoting.

Partner Institution	Your institution favours use and adoption of:			
	Narrative curricula	Attention to ability to acquire funds from external sources	Publication in Open Access repositories or non-paid OA journals	Publication of research artifacts (data and software)
Bergen		×	×	×
Granada		×		
Graz	×	×	×	×
Leipzig	×	×	×	
Lyon	×	×	×	×
Padova		×		
Vilnius			×	×

3. Sample of top-down initiatives that foster Alternative Assessment Practices

3.1 Lyon

3.1.1 Research assessment in the National Action Plans in France

The French Plan for Open Science, with a budget of 15.8 million euro, was first introduced in 2018 by the Ministry of Higher Education, Research and Innovation with the aim to support OA practices in scientific publishing. Coordinated by the Committee for Open Science (which brings together the Ministry, the universities and research performing organizations and the scientific community) and the plan has enabled some substantial progress: the percentage of open access scientific publications in France has risen from 41% to 56% since the issue of the plan, calls for projects to promote OA publication and provided support for international structuring initiatives were launched, the French National Research Agency (*Agence nationale de la recherche*, ANR) and other funding agencies now require the projects they fund to make the publications available in OA and to draw data management plans coherent with that intent.

The Second French Plan for Open Science[23] is an extension and renewal of the earlier plan following changes in the international context (e.g., the launch of Horizon Europe). It has been publicly announced in July 2021, and it will take effect in 2024. With this new plan, France is continuing the trajectory initiated by the Digital Republic Act of 2016 and confirmed by the Research Programming Law of 2020, which includes OS as one of the missions of researchers. This plan includes a wide range of actions to promote OS and the diversity of scientific production in the evaluation of research (path 4, measure 11)[24]. The launch of this plan is accompanied by international events such as the Paris Open Science European Conference[25] organised by the French Presidency of the European Union on 4-5 February 2022, and participated by the Ministry of Higher Education. The plan carries structural changes such as the inclusion of OS principles and best practices in the High Council for Evaluation of Research (Hcéres) research evaluation platform[26], strengthened cooperation between the Hcéres and the Committee for Open Science, the abandonment of all references to journal impact factor and H-index in the call for project proposals and in the application forms, the promotion of the use of narrative curricula.

The Second French Plan for Open Science has considerable potential for impact on the organization of French scientific communities, especially for its paths 1, 3, 4. The plan actions also have impact potential on the society at large by the adoption of transformative practices that make OS the default principle of innovation.

---

[23] https://www.ouvrirlascience.fr/second-national-plan-for-open-science/.

[24] Path 4: "Transform practices to make open science the default principle". Measure 11: "Value open science and the diversity of scientific productions in the assessment of researchers, of projects and of universities and research performing organizations". The two other measures to achieve this path 4 are "Develop and value open science skills" and "Triple the budget for open science" (the other three paths are respectively entitled Generalising open access to publications, Structuring, sharing and opening up research data & Opening up and promoting source code produced by research).

[25] OSEC: https://www.ouvrirlascience.fr/paris-open-science-european-conference-osec/.

[26] https://www.hceres.fr/en/missions

- Path 1 promotes the "diamond model" of OA publishing[27] via the obligation to publish with OA procedures all research results obtained from public funding.
- Path 3 addresses how science is done by introducing active policies for the promotion of open-source software artefacts, recognizing software and source code as research products that can be covered by open-source licenses.
- Path 4 promotes long-term actions devoted to implement and foster recognition of OS skills in educational and career pathways, with likely transversal impact across academic disciplines.

After the publication of this second plan, many French institutions signed the DORA declaration and expressed their desire to initiate a reform of research evaluation. A national working group (GT Evaluation DORA) has been created to act as a taskforce to accompany research institutions and universities in the reform. The CNRS has organized a conference in November 2021 to present the reform of research evaluation.

For the time being, the first reform actions are converging towards the abandonment of the Impact Factor and the ingress in a more qualitative form of evaluation, with adoption of narrative curricula. This approach aims to give recognition to a broader understanding of science production over and above traditional scientific publications, and including interdisciplinary research, naturally with major impact on early-career researchers.

An important challenge associated with this reform regards the training of researchers in qualitative evaluation, which carries an important change of paradigm for them. To sustain this dynamic, the "Best practices for research-supportive HR measures"[28] guidance was published in Winter 2021 as part of the National Action Plan for the Improvement of French Participation in European Research and Innovation Funding Schemes (PAPFE, approved in mid-2018). One aspect of interest in that guidance are the incentives for researchers to participate in the PCRI and to coordinate international collaborative projects, including promotion and monetary reward via profit-sharing schemes.

3.1.2 Awarding of research excellence

The conferment of medals and prizes at national level, to award non-traditional forms of research excellence may also operate as an effective lever of change.

A notable example of that effort emerged in 2021, is the "CNRS Medal for Scientific Mediation", an award supported by the Ministry of Higher Education, Research and Innovation, to prize actions, whether one-off or permanent, personal or collective, by women, men, scientists, research support staff, to promote science in society.

The "Science avec et pour la société" token, also launched in 2021[29], is awarded to higher-education institutions and research establishments that promote synergy between science and society. Prerequisites to that recognition include strategic governance-level commitment, strong partnership with local stakeholders in culture and industry, wide range of actions and targets (schools, the media), participatory research, and a self-evaluation process.

---

[27] Where the publication of scientific works is free of fee charges for authors.

[28] https://www.enseignementsup-recherche.gouv.fr/fr/guide-de-bonnes-pratiques-rh-pour-favoriser-la-participation-des-chercheurs-et-chercheuses-des-82423.

[29] https://www.enseignementsup-recherche.gouv.fr/fr/le-ministere-de-l-enseignement-superieur-de-la-recherche-et-de-l-innovation-annonce-les-premiers-81946.

The ANR funded calls for project proposals, with a considerable volume of funding, to promote better recognition of researchers' activities and their role for society, under the general umbrella cap of "*Science with and for society - Scientific, technical and industrial culture*"[30].

3.2 Bergen

3.2.1 Norwegian Career Assessment Matrix (NOR-CAM) initiative

NOR-CAM[31] is the Norwegian adaptation of EU's OS-CAM mentioned earlier. NOR-CAM was developed by The Norwegian University Association and published in June 2021, with UiB taking part in that effort.

NOR-CAM aims to help expand the basis for research assessment, acknowledging the rapid development and spreading of OS principles, the need to recognize a greater breadth of competencies in assessment, and the need to attenuate the dependence on quantitative publication metrics.

NOR-CAM is intended as a tool for academic institutions in recruitment and career development, as well as in assessment and evaluation practices for funding agencies and national authorities.

NOR-CAM enumerates a number of factors that should be considered when assessing general academic activity for results and competencies, which include practices and activities associated with the research process, pedagogical competences, innovation leadership and other experiences. As regards research output, the NOR-CAM matrix lists:

- Published works (not limited to scientific articles)
- Datasets
- Software
- Methodologies
- Artistic results
- Research reports.

As the publication of NOR-CAM is very recent, UiB have not yet adopted it as such in their own research management policies. The natural expectation is that it will be central in developing local guidelines for research assessment and leadership at the universities in the years to come. At an institutional level, it will typically be referenced in strategic documents and action plans for HR, in guidelines for recruitment committees, academic promotion processes and yearly employee interviews.

Shortly before the release of NOR-CAM, the UiB research policy guidelines were revised in the process of qualifying for the EU HR Excellence in Research (HREiR) award, which further deferred the adoption of the NOR-CAM per se. The priority plan at UiB is to revise their current guidelines to incorporate, among others, a checklist for Open Transparent and Merit based Recruitment (OTM-R) by the next qualifying round for HREiR.

3.2.2 Ministerial Framework for evaluation of research and education of institutions

---

[30] https://anr.fr/fr/detail/call/appel-saps-csti-generique-1819-science-avec-et-pour-la-societe-culture-scientifique-technique/

[31] https://www.uhr.no/en/_f/p3/i86e9ec84-3b3d-48ce-8167-bbae0f507ce8/nor-cam-a-tool-box-for-assessment-and-rewards.pdf

In 2021, a Working Group under the Ministry of Research and Education proposed a national framework for evaluation of quality in research and education of institutions.

While having different goals and premises than the assessment of individuals, the research evaluation of institutions has also been based to a large extent also on bibliometric indicators. One task of the Ministerial Working Group was to devise a more robust and diverse set of evaluation criteria. Unlike other Scandinavian countries (e.g. Sweden), in Norway there is no tradition for regular institutional self-evaluation of universities. Evaluation of research quality has been the responsibility of the Research Council of Norway (RCN), which has developed a system for evaluation of broad disciplinary areas (recently, Law, Humanities, Social Sciences, Pedagogy), assessing institutes or research groups within each institution where relevant research is carried out. The Working Group proposal recommends placing more emphasis on a wider range of institutional goals and strategies, and attribute higher relevance to research artifacts and other products, to include them in the assessment of research content and relevance/impact. For example, a recent RCN evaluation of Law research (JUREVAL, 2021) solicited the institutions themselves to identify relevant common evaluation criteria of quality and impact of research, and to expose institution-specific goals and strategies to assessment. In comparison to evaluation of other disciplines, JUREVAL was found to place less emphasis on international scientific output, and more on other forms of impact, like contributions to actual law development.

The NOR-CAM also provides recommendations for functionality in CRISTIN[32], short for "Current Research Information System in Norway", the national database of research products, to enable it to document a wider range of research output, and to offer CV functionalities for research assessment purposes. At present, in fact, Cristin focuses almost exclusively on scientific publications.

All universities and colleges, medical and research institutions in Norway are required to report their results through CRISTIN. The system maintains quality-assured lists of publishers and journals, and their Open Access status. The Ministry also requires universities to deposit journal articles in open repositories through CRISTIN.

75% of UiB's scientific articles are published in OA journals, and 85% deposited in open archives.

The lists of publishers and journals in CRISTIN are maintained by the Norwegian university association (UHR), and defines the scope of published results recognized by the Ministry, eligible for reporting to the results-based public financing systems for universities, medical and research institutions. The lists are further subdivided to identify internationally-leading journals and publishers within broad academic disciplines ("quality level 2"), prized with additional funding in the ministerial allocation systems. The system is central to Norwegian research assessment practices, even though it was developed solely to support results-based allocation of public funding to research institutions.

The CRISTIN system, established in 2004, has been highly controversial. It is commonly acknowledged that the system is unfit for research assessment at the level of individual researchers. However, for several reasons, over time the system has also gained wide acceptance and legitimacy. The system covers all scientific publishing, including non-English-language journals typically not indexed in international commercial services like Clarivate/Web of Science or Scopus. Therefore, the coverage of research products, particularly within SSH-disciplines, is much higher in CRISTIN than in commercial bibliometric indexes. Using CRISTIN data for comparison, we observe that only 25% of scientific publications in the

---

[32] https://www.cristin.no/

Humanities at our university are indexed by Clarivate. The lists of publication channels are actively curated by academic peers. In the last few years, this has become particularly important with the rise of predatory journals and publishers. Generally, the curated two-level list of scientific journals and publishers are preferred to systems based on algorithm-derived journal impact factor.

The CRISTIN system should be considered an alternative to traditional bibliometrics in research assessment. It has had a very large impact in Scandinavia and beyond, and it is a very mature system practice in Norway.

3.2.3 National Research Council policy on Open Science

In 2020, the Research Council in Norway published new policy[33] covering OS/CS publishing, research data, research innovation and action, research assessment. The policy clarifies what the Research Council should do and what research institutions should do collaboratively to stimulate more open research in the Norwegian research and innovation system.

The concrete measures outlined in the policy seek to stimulate an open manner of science. Some of those measures are implemented via the Research Council's own activities and funding instruments. Some of the problems and measures presented require collaboration between the Council and other stakeholders such as relevant ministries, the Norwegian Directorate for ICT and Joint Services in Higher Education and research (UNIT) and Universities Norway. Contingent on the thematic area and type of measure involved, cooperation is solicited among a range of stakeholders such as universities, colleges, committees on research ethics, councils and commissions, service providers, companies, and industrial and special-interest organizations.

The policy has established "as open as possible, as closed as necessary" as a core principle governing publicly financed research in Norway, and has been followed up by adjusting contractual terms to require OA publishing of results and FAIR access to research data.

3.2.4. National strategy on access to and sharing of research data

In 2017, the Norwegian Ministry of Education and Research published a national strategy concerning licensing and publishing of open data, which states that research data should be shared and reused more widely[34].

Since the publication of that regulation, all the major universities in Norway have established repositories, courses and services for FAIR open research data. Several national initiatives have been started, including further elaboration of technological barriers to FAIR publishing, privacy, licensing and IPR issues, and development of repository technology. The national strategy was central to developing RCNs policy for open science (described above). RCN now requires financed projects to develop data management plans, and to publish research data if possible. The strategy as such has had major impact in Norway.

3.2.5 Pursuit of excellence in HR support for Research Management

---

[33] https://www.forskningsradet.no/siteassets/forskningspolitisk-radgivning/apen-forskning/nfr-policy-open-science-eng.pdf
[34] https://www.regjeringen.no/en/dokumenter/national-strategy-on-access-to-and-sharing-of-research-data/id2582412/?ch=1

UiB has committed to the principles of the EC "European Charter & Code for Researchers"[35]. As part of that, in 2019, UiB received the "HR Excellence in Research" (HREiR) award for the period up to 2024. The basis for the award was an application in which UiB assessed its own practices in accordance with the principles of said Charter and implemented actions to offer better HR support in research. The actions have been collated in an action plan containing 15 items aimed at: career development, recruitment, the role of the supervisor, equal opportunities and diversity and administrative support for R&D employees.

The process of earning the HREiR badge has led to greater attention to HR research management, particularly concerning the provisions made for early-career researchers. An interim assessment and a revision of the Action Plan was presented to the University Board in November 2021, which shows that a number of actions has been implemented. Overall, the HREiR has been considered as a valuable aid by university management to improve the general conditions for early-career researchers.

In 2020, the university board adopted a new policy for early-career researchers, requiring faculties to offer programs for competence, career and talent development. Among other, those guidelines describe the annual employee interview as a tool for individual assessment and career development, and the need to focus on wider aspects of research than scientific publishing, in accordance with the principles in OS-CAM and NOR-CAM. As part of the HREiR action plan, UiB recently established UiB FERD, a career development center providing services for early-career researchers[36].

3.2.6 UiB policy for Open Science and Open Data

In 2020, the UiB board adopted a policy for OS. The policy covers OA to research publications, artistic research, and research data, open innovation, open educational resources, and citizen science.

UiB's policy for open science states that FAIR research data should be archived in relevant discipline-specific repositories as far as possible, e.g. Clarin ERIC, Elixir, Cessda ERIC.

UiB has its own local archival system for Open Research Data, established in 2019 and based on the open-source Dataverse platform[37].

UiB's activities within the OS policy areas are at very different levels of maturity and scope. In 2021, it has been decided to prioritize development work related to open learning resources and open data, in addition to maintaining already established systems for OA publishing[38].

Over the past two years, UiB has increased capacity for support and training in OS issues like open learning resources, FAIR data management, and OA publishing.

The "UiB Open Research Data" (UORD) is a repository for entirely open data, i.e. data where access is not limited by IPR, privacy policies etc. To date, 95 projects have archived data in the repository, for a running total of 66 000 data, metadata and script files. Over the short time span of its introduction, the UORD has become a valuable tool for long-term data management and dissemination.

---

[35] https://euraxess.ec.europa.eu/jobs/charter-code-researchers
[36] https://www.uib.no/en/hr/137241/hr-research
[37] https://dataverse.no/dataverse/uib
[38] https://www.uib.no/en/foremployees/142184/university-bergen-policy-open-science

3.3 Graz

3.3.1 Professorial appointment procedures

Like all higher-education and research institution, the University of Graz seeks to recruit highly qualified and promising scientists to strengthen its positioning in the vast spectrum of research fields of their strategic interest. The University of Graz is also aware of its societal responsibility and places much attention to the education of students and the formation of early-career researchers. On account of that, the university has designed an appointment procedure – adapted to the varying career level of the candidates -- with focus not limited to research but also considering teaching, training, attention to gender equality, and societal engagement.

The appointment procedure process is defined in accordance with the Austrian University Act and the statutes and directives of the University of Graz. The amendment of the University Act in 2015 enabled new career opportunities, for tenure-track national and international candidates, and for highly-qualified tenured associate professors already in post at the University of Graz to rise to the rank of full professor.

The appointment process starts with the definition of profile requirements by a committee composed of 6 professors, 2 representatives of the non-professorial teaching staff, and 2 students. The profile requirements, approved by the Rector, specifies:

- The strategic orientation of the professorship position
- The employment requirements
- The expected qualification profile (portfolio) of the applicant
- The achievement targets to be met.

Applications for promotion to the full-professor position are required to include:

- a list of previous research projects and other relevant forms of scientific collaboration;
- a description of intended future research activity and
- a list of the five most important publications.

Applications for tenure-track positions are required to include:

- a report (exposé) on future projects and objectives in the field of research
- a list of the three most important publications.

One important highlight of such requirements list is the attention to quality more than quantity in the scientific production.

For the selection of tenure-track candidates, the evaluation committee singles out five applicants to invite to a hearing. The application documents are also forwarded to experts (with at least one of them external to the university) who assess the applicants on the basis of the profile requirements[39]. After the hearing, the committee makes a recommendation for the appointment. The Rector eventually selects the candidate after considering the opinion expressed by the department to which the position is assigned. For the recruited candidate, a qualification agreement is then agreed upon, based on the profile

---

[39] Translated from the German term *Anforderungsprofil*

requirements for the post. Upon successful achievement of all the expectations set in the qualification agreement, the person is promoted to full professor.

The process for the selection of candidates to full professorship is similar in principle but also involves external reviewers. The qualifications profile serves as a guide for evaluating applicants and for justifying the appointment proposal. The appointment decision considers a **teaching assessment**[40], which includes an evaluation and a demonstrative lecture by the candidate, and, possibly a bibliometric analysis, dependent on the scientific discipline. After final selection endorsed by the Rector the employment negotiation starts, which sets agreed target measures to achieve, including, for example, societal impact, for the first six years of the appointment.

At the beginning of 2020, the appointment target agreements were redesigned and the distinction between the following experience stages as a full professor was implemented:

- Professorship initial employment
- Professorship follow-up
- Professorship – high performer.

Prior to this revision, all candidates, regardless of their career level, were asked the same questions and given the same targets. The new procedure allows much more tailoring to the specific objectives of the professorship position, and thus more adapted development goals.

The person selected for a professorship position is required to prepare a concept for future research projects and teaching undertaking on the basis of a tailored set of questions. The personalized goals for the next review period reflect on both the concept and the profile requirements of the position. A target achievement review takes place annually. In case of non-attainment of contractually agreed targets, a detailed analysis of the reasons is undertaken. If the professor and the university concur that the observed shortfall is not attributable to him or her personally, the variable component of the salary is not reduced or suspended.

Important highlights of target agreements include explicit consideration of the candidate's level of experience and future research plan, and attention to societal impact and interaction. The base structure of the agreement instrument is given top down; its concrete contents and targets however have an increasingly bottom-up individually-tailored nature. Especially important is the annual review of achievements, which allows detecting problems in a timely fashion and attempting to cure them promptly.

Reflections are being made on where the above procedures may be improved in the future, possibly considering OS works and products (data, software, methodologies, inclusion of CS), transdisciplinarity, and more attention to demanding qualitative goals.

3.3.2 Research evaluation procedures

---

[40] To increase attention to the teaching quality dimension, the selection procedure includes a Teaching Skills Assessment (TSA). In additional to preliminary check-box information, the TSA entails a demonstration lecture given during the hearing, and specific criteria designed to measure the candidate's didactic skills.

Societal responsibility is anchored in the strategy of the University of Graz. Initiatives that present scientific achievements to the broader public, and thus promote knowledge transfer to society have existed for many years. Cooperation with industry is being actively intensified to make the most recent scientific results readily available to industry and society.

To this end, the University of Graz adopted the in-house Research Information System portal[41] already in 2007. The RIS platform also stores third-mission activities such as oral presentations, non-academic dissemination events, media contributions, and publications intended for a wider audience. The Research Evaluation procedures applied to the research units of the university assesses societal engagement and networking activities in addition to scientific accomplishments and measures on the promotion of early stage researchers.

The Research Evaluation is part of the university's quality management system and helps steer planning and decision-making processes. Academic research units and subject areas are evaluated at least every seven years, with focus on development and profiling, optimisation, improvement.

The Research Evaluation is carried out using a uniform, but subject-specific process, which comprises the following stages:

- **Narrative** self-evaluation, including strategic orientation and attention to **third-mission** efforts
- Peer visit and peer review
- Implementation workshop.

The experience gained in the first two evaluation rounds since the inception of the system was used in the third cycle of research evaluation, which is now coming to an end. The critical review of it has given priority to: the profiling of the unit to be evaluated, its future viability, the promotion of early-career researchers. The Research Evaluation process has evolved from a simple level of accountability to being a communicative, participatory improvement process. In some evaluation units, experts from outside the field are also brought in to provide a broader perspective in the assessment of the future potential of the research topics, also considering their relevance to society and the research policy of the institution. Those experts are also required to give feedback on internal and external cooperation and assess the coherence of the strategic analysis of the unit being evaluated.

Research evaluations are effort-intensive and their proceed critically depend on the openness and trust of all those involved in it, and the actual prospects (financial, strategy) of implementing proposed changes.

Research evaluations may have long-term effects and also more immediate ones, such as, for example, the identification and implementation of specific training initiatives or faculty development processes, institute mergers, the opening of additional professorial or management positions; the creation of more social space.

Important highlights of the Research Evaluation procedure are that is narrative, participatory, attentive to third-mission efforts, open to expects from other research fields than those being considered, and designed to self-improve from lessons learned in the application of it.

Reflections are being made on where the above procedures may be improved in the future, possibly considering OS works and products (data, software, methodologies, inclusion of CS), and transdisciplinarity.

3.3.3 Transdisciplinarity (webinar #5)

---

[41] https://online.uni-graz.at/kfu_online/wbForschungsportal.cbShowPortal

The fifth webinar, which took place on 10 September 2021, centred specifically on *the drive to the understanding, recognition and promotion of transdisciplinary research* at the University of Graz (UG). The invited speaker was Holger Hoff, Transacademic Interface Manager, at the Wegener Center for Climate and Global Change, Faculty of Environmental, Regional and Educational Sciences, UG. Focus in this webinar was the rationale of UG's for its attention to transdisciplinarity and the lessons learned to date from it.

Attention to transdisciplinarity, as a systemic manifestation of OS/CS) is present and visible at UG. Transdisciplinarity is anchored in the University's Development Plan as a strategic priority and a strategic goal. It aims to seek solutions to societal challenges via continuous dialogue between researchers and stakeholders.

- The level of societal engagement achieved and the knowledge transfer efforts made feature within the assessment criteria for the appointment of new research staff.
- The institutional research portal at UG enumerates third-mission activities for all researchers individually and cumulatively.
- UG offers regular training courses related to OS/CS and makes calls for OS/CS projects.

The current situation has a number of known challenges:

- Career incentives are limited, since high-level scientific publications are still difficult to derive from transdisciplinary work.
- Coupled with that, there appears to be limited academic reward and prestige from trans-disciplinarity, which – most evidently – is not and far off the mainstream as yet.
- Opportunities and uptake of OS/CS vary much across scientific disciplines and university departments, hindering a uniform approach and a common understanding and recognition.

Interesting and positive trends can be observed:

- There are increasing funding opportunities (e.g., Horizon Europe) and growing societal demand, which should mobilize the interest of researchers.
- The younger generation of researchers (and of proposal evaluators) appears to be more interested and involved in OS/CS than their seniors, which may allow for progressively wider adoption of OS/CS practices.

3.4 Leipzig

3.4.1 Research assessment in the context of German federalism

Higher education as well as the academic research system is not centrally organised in Germany, which makes them highly fragmented. Both themes fall under the responsibility of the regions (Länder), regulated by their higher education laws. Under the Saxon Higher Education Freedom Act (*Sächsisches Hochschulfreiheitsgesetz*), a Higher Education Institution (HEI), for example, is committed to setting up a system to ensure the quality of its work as well as the quality of research and teaching, which it has internally and externally evaluated at appropriate intervals. The central bodies of a HEI shall regulate the details, in particular informing those concerned about the purpose and content of surveys and evaluations as well as the procedure for evaluating research, in a set of regulations. Thus, there are no national or regionally prevalent indicators or procedures that should be used for research assessment. This arrangement gives HEIs in Germany an element of freedom and relative autonomy. Therefore, HEIs determine their own assessment procedures and methods in light of the local context as well as best practices at international and national levels. The university sets up target agreements with the faculties, which are agreed upon an individual basis.

Federal institutions such as the Federal Ministry of Education and Research (*BMBF*), the German Council of Science and Humanities (*Wissenschaftsrat*) or the German Research Foundation (*DFG*) provide scientific policy advice, guidelines and initiatives for the development of science and the promotion of relations between the research community, society and the private sector. Some examples are the Strategy for Open Access in Germany[42], DFG Guidelines for Ensuring Good Research Practice[43], or the Recommendations for the Evaluation and Governance of Research Performance[44] elaborated by the German Council of Science and Humanities.

In Germany, both federal institutions and various foundations provide a considerable volume of funding to HEIs and non-HEIs. In so doing, they establish general requirements for the evaluation of institutions, researchers and projects that institutions and researchers should fulfil in order to be awarded funding.

3.4.2 German Research Foundation evaluation standards and guidelines

German Research Foundation (*DFG*) is the biggest central independent research funding organization in Germany. The DFG promotes excellence by selecting the best research projects on a competitive basis and facilitating national and international collaboration among researchers. The qualitative scientific review of project proposals is an integral part of the DFG's funding process. In 2020, approximately 24,700 written reviews were received from nearly 16,800 reviewers, 22 percent of whom were women, moreover a third of all DFG-reviewers were based at research institutions outside Germany[45].

In 2019, against the backdrop of international debate on research integrity, the DFG adopted the Code of Conduct Guidelines for Safeguarding Good Research Practice. The Code defines fundamental principles and guidelines that cover the key steps of good practice throughout the research process. This Code is aimed at both researchers and institutions.

Among such principles as professional ethics and responsibility of the heads of research work units, a principle on dimensions of performance and assessment criteria are defined. According it to assess the performance of researchers, a multidimensional approach is called for. In addition to academic and scientific achievements, other aspects, such as involvement in teaching, academic self-governance, public relations, knowledge and technology transfer, contributions to the general good of society may be taken into consideration during assessment procedure. It is stated that research performance is assessed primarily on the basis of qualitative measures, while quantitative indicators may be incorporated into the overall assessment only with appropriate differentiation and reflection. Where provided voluntarily, individual circumstances stated in curricula vitae – as well as the categories specified in the German General Equal Treatment Act (Allgemeines Gleichbehandlungsgesetz) – should be taken into account during assessment process. It also states that high-quality research is oriented towards criteria specific to individual disciplines.

---

[42] https://www.bmbf.de/SharedDocs/Publikationen/de/bmbf/1/24102_Open_Access_in_Deutschland.pdf;jsessionid=982C7613C7E8E698CB4CBDDD066A81C7.live382?__blob=publicationFile&v=5

[43] https://www.dfg.de/download/pdf/foerderung/rechtliche_rahmenbedingungen/gute_wissenschaftliche_praxis/kodex_gwp_en.pdf

[44] https://www.wissenschaftsrat.de/download/archiv/1656-11.pdf?__blob=publicationFile&v=1

[45] https://www.dfg.de/en/dfg_profile/facts_figures/evaluation_studies_monitoring/evaluation_standards/index.htm

In the regards of the transition to Open Science, the DFG deems it essential that research data, project-related publications and research software be freely available to the general public after the project is completed.

3.4.3 Awarding of research communication

The *Community Prize* is a new competition format within the framework of the "Research in Germany" initiative of the Federal Ministry of Education and Research that combines competition and cooperation. The competition aims to increase the international visibility and networking of German universities and research institutions. Individuals engaged in academic administration and research at Higher Education Institutions and Research Institutions in Germany can compete for funding for project ideas in the area of international research marketing and participate in an online vote on the entries made in the competition. Competing entries are presented in the form of short video clips uploaded on a digital voting platform. Nearly 500 individuals and research institutions registered as contestants on the platform in 2020. About 350 of them took part in the voting.

One of the winners of the 2020 competition was the Research Centre Global Dynamics (ReCentGlobe)[46] at Leipzig University. ReCentGlobe is member of a worldwide network active in global-area studies, which produces a variety of blogs, podcasts, YouTube videos, open educational resources and numerous other formats on the subject of current global social issues.

3.4.4 Research assessment in the light Human Resources development at UL

One of Leipzig University's stated goal is to expand staff development for early-career researchers. The grand goal includes improving existing integration, support and qualification systems and establishing new ones. Worth of mention in this regard are the Pre-Doc Award, the establishment of Graduate Schools, and the implementation of the Leipzig Tenure Track Programme (LTTP) and the Leipzig Researcher Development Programme (LRDP).

Fair and just appointment processes based on adequate evaluation of the researcher's performance, which take account of the ongoing transformation of the research culture, are a priority for Leipzig University. Although not discussed as a single specific strategic issue, knowledge transfer, science communication, and stakeholder engagement become increasingly important in the transformation of the research assessment process in its various instances Tenure-Track-Regulations *(Tenure-Track Ordnung)*, Quality Assurance Concept *(Qualitätssicherungskonzept)*, Leipzig University's Staff Development Plan (Personalentwicklungskonzept), Interim Evaluation Regulations (Zwischenevaluationsordnung für Juniorprofessor:innen).

One notable element of that transformation is the **planning meeting** that is required to take place in the LTTP. This procedure testimonies Leipzig University's desire to offer young researchers attractive, reliable and timely career prospects on the path to tenured positions. To this end, the tenure track professor (TTP) is structurally integrated in the overall strategic concept of personnel development, based on a transparent, quality-oriented procedure.

---

[46] https://www.youtube.com/watch?v=XAxX_XlAWgQ

The planning meeting is where the performance goals of the TTP are discussed (by the candidate and the Dean of the corresponding faculty) and set in an evaluation agreement on the basis of which the subsequent performance evaluations will take place.

The performance goals address the following four categories: research, teaching, knowledge transfer, stakeholder engagement. Two to three performance targets per category are set. The performance goals should be science-specific, i.e., measurable, qualified, verifiable and clearly formulated. This approach favors a more encompassing notion of assessment, which extends from quantitative research vision to a qualitative and broader academic perspective.

Below are some selected examples identified during the planning meeting and included as further evaluation criteria in evaluation agreement at UL:

Research: Supervision of doctoral and postdoctoral students; Management of research projects; Cooperation with other institutions (internal, national and international) through research projects, publications, conferences, etc.

Teaching: Participation in continuing education didactic programs; Developing, testing and scientifically supporting innovative learning formats (e.g., business game, blended learning, digital resources); Curating specific aspects in course programmatic (e.g., quality culture in teaching and learning, digitalization, internationalization

Knowledge Transfer: Transfer of scientific content into society through media contributions or interviews; Science communication for wider public (school and other public institutions); Cooperation with the Science Night, Kinderuni, open-day school labs; training for school teachers; News releases on institutional homepages; Participation in expert reports/studies for external consultation (e.g., ministries); Publications on pertinent blogs, portals, podcasts.

Stakeholder engagement: Regular and visible activity in national and international professional societies; Co-operation with and management of professional associations.

Some relevant lessons can be learned from the deployment of the above-described procedures so far.

1. The agreed performance goals (assessment criteria) appear to be highly differentiated across scientific research disciplines and disciplinary areas, even within faculties.
2. Attention to quantitative indicators of publication in peer-review journals, book chapters, monographs remains high across all disciplines, with some contextual differentiations among them.
3. Traditional evaluation criteria (such as bibliometric indicators and the ability to obtain research funding) continue to be used on the premise of being easy to determine and specify. Interestingly, those criteria are considered important by the researchers themselves and are given preference in the agreed evaluation criteria.

The evaluation agreement is a relatively new instrument for research evaluation at UL. It is expected that it will evolve in the future to reflect the interests of all stakeholders and the reform of the research assessment system.

3.4.5 The role of the Research Service in research assessment (webinar #3)

The theme of the webinar, that took place on 26 July 2021, focused on the analytical tools for research assessment at UL and thus provides context to the top-down assessment practices at this institution. The

invited speaker was Marc Kaulisch, external project manager of the "leuris:Analyse" project at UL. The focus of this webinar was on the research service's role in research assessment. The highlights of this speaker's contribution can be summarized as follows.

The UL Department of Research Service provides support to faculties and central research units for defining targets and measurable indicators for research analysis and assessment, as well as support university with implementation of digitization strategy and open science strategy.

UL develops its own research information system leuris, short for "LEipzig University Research Information System", which serves as a source of information on research products/outputs within the university and as a basis for data-driven research management in order to identify current research directions as well as outstanding scientists at an early stage and to promote their networking. Currently, leuris consists of several modules: leuris:Research Report, leuris: Personal Profile, leuris:Portal.

By enabling academics to present their manifold research outputs (research projects, publications, international collaborations, prizes and awards, scientific events, CV, research foci, teaching activities) and link their profile to ORCID leuris creates the open and trustworthy environment for research assessment. All research units and researchers are required to report their results via leuris. Therefore the research information collected on leuris provides valid data for benchmarking units and individuals. UL and its organizational units thus fulfil the legal mandate to regularly inform the public about its research activities and research results in accordance with Saxony's Higher Education Freedom Act. Furthermore, being in-house system it reflects and implements the factual requirements of researchers and university management. For example the possibility of integrating meta-data on research data as well as on open access into leuris is currently being worked on.

In 2020, the leuris:Analyse project was launched with the aim of developing new innovative tools and methods for research analysis and assessment in order to provide support to the Vice Rector and the Department of Research Service in the assessment process. As an end product, leuris:Analysis delivers meaningful and practice-relevant evaluations, graphics and tools on research information, key figures and indicators. Within the scope of the project, various sources of information (leuris database, WoS, Dimensions, DFG database, CORDIS, etc.) are used. Methods that encompass a wide range of indicators (publication type and open access type, SNIP percentiles of journals, citations and their derivatives, co-authorship networks) along with text mining (topics modelling and their analysis over time), as implemented in leuris:Analyse, can be considered as an alternative to traditional bibliometrics in research assessment. Thus, the number of publications is a minor component of the research analysis. Among the advantages of leuris:Analyse is that it overcomes the weaknesses of WoS or other bibliometric databases, where, for example, some SSH disciplines are not adequate represented or the metadata of some publishers are broken and cannot be indexed. The leuris:Analyse initiatives provide the opportunity to develop the long-term infrastructure for data-driven research services and operational controlling at the university level.

UL consolidates a proper assessment[47] based on wide spectrum of quantitative as well as qualitative indicators based on analyses of publications and acquisition of third-party funding. Recognizing the relevance of taking into account ever increasing number of new scientific products ranging from data sets,

---

[47] The intended interpretation of "proper" in this context is the internalization of the metadata required to establish the quantitative indicators of interest without externalized dependence on commercial enterprises such as Scopus, WoS, etc.

software, algorithms and protocols to citizen and open science projects in research assessment, the UL also acknowledges the significant challenges and practical problems involved: what should and could be measured, what the actual consequences are, and how complicated and expensive it would be to maintain the chosen alternative approach in the long term. Such practices should be embedded at national or regional levels and encouraged by the HEI through education and incentives.

3.5 Vilnius

The contents of this section reflect the proceeds of the webinar that took place on 3 September 2021, with a theme centred *on the research evaluation of institutions and individuals in Lithuania in general and at Vilnius University (VU) in particular*. The invited speaker was Vida Lapinskaitė, Director of the Department for Research and Innovation at VU, singled out as a senior carrier of a top-down institutional view of research assessment practices at VU.

The evaluation of research institutions follows rules defined by National Bodies, Regulations, Councils (Ministry of Education, Science and Sports; Law of Science and Studies; Research Council of Lithuania.

The assessment criteria employed for it concern the quality of R&D quality, its socioeconomic impact and perspective, centred on publications and other research results, international R&D projects, R&D contracts (governmental and business), license agreements (only patents/patent applications).

The allocation of central funding for research and education to the examined institutions depends on the outcome of that evaluation.

The evaluation of the research product of individuals considers scientific publications, projects, applied research activities (i.e. patents), internationality, presentations at conferences; teaching competencies (limited to teaching staff only); scientific achievements (visibility and impact); academic leadership (only for higher level positions) in research groups/projects, organization of scientific events, mentoring of early researchers; other academic and social activities (membership, institutional activities, activities as reviewers, public education and activities to promote the dissemination of science); potential for the future (minimal requirements defined at a national level, extra requirements defined at institutional level).

A critical view can be made of the status quo, signalling that:

- A transition to qualitative instead of quantitative evaluation of research is desired but not actuated. Admittedly, the present focus continues to be on quantity, without attention to quality.
- A wider spectrum of activities should be taken into account and they are not, as yet.
- For the evaluation of high-level researchers, the challenge is to recognize true excellence without being carried away by stereotyped or soft criteria.
- Transition to alternative assessment approaches must be made with extreme care, and attention to pros and cons.
- Extremism in the evaluation criteria should be avoided: it would not be healthy to replace bibliometric-criteria-only with exclusive focus on OS/CS.

3.6 Expected evolution of institutionalized assessment of research (debate #1)

On 17 September 2022, WP3 organized a debate on the theme in the heading, with a panel comprised of the following experts:

- *Eva Méndez*, Associate Professor at the Department of Library and Information Science, Universidad Carlos III de Madrid, Spain (outside of the Arqus Alliance)

- *Linas Tarasonis*, Vice Dean and Research fellow, Faculty of Economics and Business Administration, Vilnius University (VU)
- *Chérifa Boukachem-Zegmouri*, Full Professor at the Department of Computer Science, Université de Lyon
- *Massimo Castagnaro*, Full Professor, Department of Comparative Biomedicine and Food Science, University of Padova (former member of the National Agency for the Evaluation of Research).

The panellists were singled out owing to their direct involvement with the works of National Research Evaluation Agencies, and their consequent ability to reflect retrospectively on their wisdom, strengths, weaknesses and prospects.

The debate was run by posing the following three questions to each panellist and then triggering a discussion among them and with the audience.

Q1	How do you judge the outcome/effect of the top-down research assessment criteria that have applied in your country/institution/context so far.
Q2	What short- to mid-term evolution do you anticipate/expect/plan for such criteria.
Q3	What changes would you personally wish to make to such criteria after your personal experience of them.

The highlights of the debate may be summarised as follows.

Méndez

Q1: Spanish institutions and researchers have been taught to strictly follow the evaluation criteria and policies of the National Agency for Quality Assessment and Accreditation of Spain (ANECA). They have proven to be very good at that. Spain, Italy and France have very powerful national-level agencies in charge of setting and implementing assessment policies and procedures. Individual institutions have very little freedom on that matter. Consequently, it is very difficult to motivate researchers to open to new innovative ways of making responsible research, when they continue to be assessed on the basis of old traditional criteria, such as journals impact factor (JIF) and citation index.

Q2: In November 2021, the EC published the report "Towards a reform of the research assessment system"[48] to foster agreed institutional changes. Achieving changes will not be easy because doing so entails a cultural transition: the choice of what to measure for research assessment directly influences research culture and behaviours. The JIF must (and will) be abandoned, but also rankings should be abolished or should include at least the OS practices and the research social impact as evaluation criteria.

Q3: The evaluation criteria should be adapted to the different career stages of researchers (as for example in the ERC scheme of things). Reward motivations should be strengthened: (1) money is actually already

---

[48] The report was meant to facilitate and speed up reform in research assessment (research projects, researchers, research units, and research institutions. It was produced from a 9-month consultation with European and international stakeholders, which yielded principles and recommended actions that could be agreed between research funding and research performing organisations, as they have the responsibility to define their criteria and processes to assess their researchers and research projects.

in place because OS is part of the evaluation criteria in the awarding of European project funding; (2) legislation is not sufficiently developed; (3) the quality of the research product must be preferred over quantity. The quantitative approach is rather easy to apply but it is most evidently not fair because it reduces researchers to a number; the qualitative approach is more expensive but also more seductive for researchers who will be active actors in making something valuable for the society.

Boukachem-Zegmouri

Q1: We are experiencing a change from a traditional way of disseminating research and the new vision boosted by the introduction of the Open Access. OS is a very hot topic in France. Until now, researchers have very well adapted to the drive of competitive publishing in a system based on journal rankings. Early-career researchers are practicing OS very naturally, *without* expecting any reward.

Q2: In the mid-term, JIF will still be with us, but it will be downsized and integrated with other criteria. In the Netherlands, for example, JIF has already been abandoned, with researchers are being rewarded on different premises. There are Evaluation Committees at several European institutions which take in account just the most relevant publications of a candidate (e.g., up to 5), thus making a clear drive to quality over quantity in the evaluation procedures. Moreover, the pandemic clearly stressed the importance of a dialogue between Science and Society. A new approach is essential to differentiate the disciplines and understand and recognize their trait and strength: SSH have been a victim of the JIF.

Q3: The evaluation should be adapted to the different career stages of researchers. Nowadays we observe that researchers very close to retirement appear to be more keen on doing OS and outreach activities than traditional research activities: clearly, there is something wrong in this phenomenon, which causes OS and third-mission efforts to pass as a luxury of the senior.

Side comment: At present in France PhD candidates are required to publish at least 2 papers in order to be allowed to defend their thesis. This approach teaches them to be more attentive to gaining visibility and reputation than to making good research. Replication crisis typically follows from this attitude.

Tarasonis

Q1: Researchers respond to incentives created by the system. Hence, all outcome depends on the nature of such incentives. In the last 10 years, Lithuania moved from a traditional framework based essentially on external peer-reviews, toward introducing the evaluation of other elements such as socio-economic impact. The lesson learned from that effort is that appropriate incentives have to be put into place to trigger change.

Q2: Lithuania has a mixed system (quantitative and qualitative), which allows new criteria to be included in it quite easily.

Q3: My personal hope is to go 100% into the new (mix of qualitative and quantitative) system. It is important to find adequate resources to perform sound and thorough evaluations. Collaboration among universities is essential to that end as well as to exchange good practices. In the last few years we witnessed that outreach activities and communication about research are fundamental aspects in the work of researchers, which have to be recognized and rewarded.

Side comment: in the economics sector, to publish top journals may take 4-5 years. Consequently, PhD candidates in economics are not required to publish: their evaluation is based on an assessment of the

potential of their research in the future. On the whole, research assessment is an extremely expensive practice for institutions.

Castagnaro

The assessment of universities and research at the national level in Italy started 10 years ago. There are two levels of evaluation: qualitative assessment of research products and their impact, based on publications for individual universities and departments; and the evaluation of research management practices of institutions for the allocation of competitive (reward-based) funding from the central government.

During the first multiannual assessment at national level, we at the National Agency (ANVUR) realized that, in some scientific areas, 8% of the researchers did not publish. This evidence provided a factual tool for performing comparative evaluation within one and the same discipline at national level. In my view, that was the most important result of the first campaign of nation-wide assessment that ANVUR run.

For the assessment of research management, we at ANVUR found that there was no transparency on the strategic plan of universities: focus was frequently set on delivering teaching than on deploying and sustaining research.

It is important to appreciate that the goal of any evaluation must always be decided up front, in order that the tool best fit to achieve it can be determined accordingly. This suggests that there is no impediment to making changes, as long as they are consciously decided.

I personally believe that we should move from a national evaluation system to a European one. However, if only 10% of the project proposals are awarded funding, we will not have a large enough basis to evaluate, which is a very serious problem.

One clearly negative ramification of the nation-wide evaluation system put in place for promotions is that early-career researchers appear to be more engaged and interested in their pursuing career than in taking teaching commitments. This is a serious problem because the tasks of universities in Italy revolve around on teaching *and* research[49]: sound balance must be preserved between the two.

---

[49] With third-mission responsibilities only beginning to appear on the horizon.

4. Sample of bottom-up initiatives that warrant Alternative Assessment Approaches

4.1 Lyon

A bottom-up reflection is taking place at present in France across the various research communities within disciplinary groups, at local or national scale, via surveys, interviews, and reports. For example, a recent work by Bastien Soulé at the University of Lyon[50] has shown factually how multi-disciplinarity was penalized by evaluation criteria siled around individual disciplines.

The main point of concern to the actors in that reflection effort is the lack of a global view on where to the transitions taking place on research assessment are leading. On the one hand, what is currently being in the discretionary circles of evaluation committees is not documented. On the other hand, the reforms undertaken are too recent to have accomplished any record of success.

Many institutions in France have manifested will to participate in the reform of the research evaluation system. Université Lyon 1 has signed the DORA declaration and is part of the national working group "GT Evaluation DORA" that seeks to provide more recognition to OS practices in the national research assessment system. Université Lyon 3 received the European "Human resources strategy for researchers" (HRS4R) award in 2018[51]. Overall, however, it is difficult to identify concrete actions that are causing real change in practice.

A number of initiatives can be mentioned however that actively aid the medium- to long-term transformation of research assessment:

- Since 2019, the library of the Université Lyon 1 has started a portal for the storage and retrieval of research data in chemistry and physics[52], as well as access to digital research-support tools. This is just one exemplar of similar initiatives for other research disciplines.
- The library of the Université Lyon 1 hosts the platform "Open Access & Frais de publication" that collects information on the Article Processing Charges and publishing policy of more than 13,000 journals as well as to the library agreements with publishers to facilitate Open Access publishing.
- The Lyon Saint-Étienne Editorial Pole, which brings together the universities Lyon 2, Lyon 3 and the Maison des Sciences de l'Homme Lyon Saint-Étienne, run the Prairial platform that does the same for the social sciences and humanities disciplines.

Since 2013, the Université de Lyon's "Science Store" at Lyon Saint-Étienne aims to liaise the research community with the general public, by enabling the latter to draw from the former in questions related to societal issues. This effort is part of the Université de Lyon's participation in the international "Living Knowledge" network. This effort is accompanied by the development of a laboratory of social innovation and the involvement in the "Responsible Research and Innovation" (RRI) campaign promoted by the European Commission. The Science Store has also participated in many European initiatives and projects (PERARES, EnRRICH, InSPIRES) centred on participatory and CS. It must be noted, however, that making contributions to the Science Store provides no direct boost to the research assessment of the contributors.

---

[50] https://www.cairn.info/evaluer-la-recherche-multidisciplinaire--9782910448301.htm
[51] https://euraxess.ec.europa.eu/jobs/hrs4r
[52] https://www.datacc.org/

4.2 Padova

4.2.1 Returns from practices in Open Science and Citizen Science (webinars #1 and #2)

A first webinar on this topic took place on 4 June 2021. The invited speaker was Mauro Masiero, Assistant Professor at UNIPD, founder of a spin-off specialized in the conservation and the promotion of forestland. The speaker was singled out as a carrier of testimony of thorough, organized and fruitful involvement with OS/CS initiatives while also being in the lowest rank of tenured professorship, thus naturally sensitive to the dynamics of career progressions.

The highlights of this speaker's contribution may be summarised as follows.

Positive returns from this line of action

- Networking and transferability: our spin-off participated in 2 Erasmus+ projects with other universities and spin-offs
- Acquisition of new skills and knowledge useful to broaden research and teaching activities
- Data collection for academic and non-academic publications.

Negative returns

- Engagement with the spin-off was no booster for academic career, which continues – in Italy – to be driven by the notorious *publish-or-perish* logic. Spin-off activities were not acknowledged in the academic job progression procedures although in 2013 the National Evaluation Agency (ANVUR), instituted by the Ministry of Research, formally included a *third-mission* dimension to the assessment with indicators that cover impact on research field as well as on social environment (i.e. nr. of people involved, trained, employed, non-scientific publications)[53].
- Researchers often lack communication skills to reach out to non-expert audience and have scant institutional support to mitigate their deficiencies.

These activities should be better evaluated in research assessment procedure.

The second webinar on this theme took place on 11 June. The invited speaker was Carlotta Mazzoldi, Associate Professor at UNIPD, principal investigator in a number of several Citizen Science projects and activities focusing on human sea culture and marine environment biodiversity and conservation.

The speaker was singled out as an exemplar of long-standing involvement with OS / CS initiatives, one level upper than speaker Masiero (webinar #1) in the ranking of tenured professorship, but evidently still fully exposed to the dynamics of career progressions.

The highlights of this speaker's contribution may be summarised as follows.

Positive returns from this line of action

- There appear to be increasing opportunities of profitable use of CS in competitive scientific projects and scientific publications, so long as researchers are able to combine the CS approach with more traditional ones

---

[53] This experience by the speaker shows a very-well known phenomenon whereby the researchers tasked to perform research assessments on behalf of the National Agency prefer to conform with *traditional* practices, with which they are more familiar for input, expectations, and output, and consequently feel less comfortable with the recognition and evaluation of "less traditional" products and profiles. Evidently, this phenomenon of "defensive conservatism" causes considerable lag in the actuation of even timid openings in the given regulations.

- CS naturally provides opportunity of data collection and education on conservation at the same time
- CS supports stakeholder engagement, which can be a precious ally in future activities
- CS teaches how co-management practices allow reduction of conflicts among stakeholders
- There appear to be Increasing institutional, national and international support, acknowledgement, interest and cooperation
- CS is an ambit that lends itself well to student involvement.

Negative returns

- It may be exceedingly hard to collect good and useful data from non-academic participants
- It is difficult to earn trust from domain stakeholders
- The professional duties of academic research leave insufficient room for temporally continuous and effort-intensive CS activities, which therefore inevitably become more affordable for tenured established researchers
- Choice of language to communicate with non-scientific actors and audience needs to be guided and supported from communication and dissemination experts, which academic institutions do not seem to possess enough (UNIPD currently does not)
- CS suffers from low rate of transferability of practices across domains.

OS/CS projects have a value per se at educational and research level, but these outcomes still do not appear to feed research assessment procedures and boost their outcome.

4.3 Anticipated impact of the drive toward Open Science (debate #2)

On 21 September 2021, WP3 held a debate with focus centred on *the influence or pressure exerted on research production by the requirements of external funding agencies (e.g., EU-level, private) and their (harmonic vs conflictual) relation with institutional research assessment policies*. The panel included:

- Elena Giglia, "Open Access Projects" Unit at the University of Turin, Italy (outside of the Arqus Alliance)
- Stefano Vassanelli, Associate Professor at Padova Neuroscience Center, UNIPD (with an important track record of funding acquisition).

The panellists were singled out owing to their familiarity and direct experience with the requirements and implications of Open Science. A third panellist, from outside of Italy, was selected and invited but eventually had to decline. The schedule of work in Task 3.1 did not allow postponing this debate until a replacement third panellist could be recruited.

The debate was run by posing the following three questions to each panellist and then triggering a discussion among them and with the audience.

Q1	From the standpoint of your professional role, in what forms – if any – do you see such pressure manifest?
Q2	Do you think that pressure on research actors is understood by all those concerned and that effective measures are being put in place to favour the valorisation of the research products resulting from such funded research?
Q3	Are you perceiving a conflict / divide in the research assessment measures between "traditional" (peer-reviewed, indexed) research products and "emerging, funding-driven" ones. Do you see sound ways to reconcile them?

The highlights of the debate may be summarised as follows.

Giglia

Premise: With Horizon Europe (HE), the EC made a bold step forward in the implementation of OS in research, introducing OS evaluation in the applications for research funding as a means to pressure PIs to design their research project including OS aspects. Funding agencies (e.g., UNESCO), research-performing organisations, service providers all over the world believe in OS very much. The National Research Evaluation Agency in Italy (ANVUR), however, currently is discouraging this transition, by continuing to anchor research evaluation to impact factor and journal rankings, and keeping researcher from the due-diligence of opening their research data.

Q1/Q2: Pressure is felt and understood, but the outputs vary much across the EU. Northern countries (UK, Netherlands, Sweden, Norway, Denmark, Germany) started to work on "OS thinking" back in 2016, creating infrastructures and awareness (training, information). Southern countries (Spain, Portugal, Greece,…) are opening to it just now, at a slower pace.

Q3: All disciplines have tools and infrastructures to implement their own style of OS. The problem though is that researchers continue to be assessed on impact factor et similia. The EC is trying to change the evaluation criteria. The current pandemic has taught us the importance of pre-printing, when the slow-paced traditional way of research publishing revealed its weakness: the first sequence of Sars-Cov2 would have been disclosed long after the global wave of the coronavirus, had the corresponding research gone through the normal publication process. It is not healthy (fair, encompassing, respectful of diversity) that a research be considered good only because *Nature* published it. I believe it is preferable that research should go through open competition and open review with other researches (ideas, methodology, data) on the basis of preprints.

Side comment: There is a frequent misconception about the relation between OS and patenting: HE does allow publishing (in open access) *or* patenting (for which no publication is required). HE is an impact-driven programme, where researchers must exploit their results. If they are not interested in direct exploitation, they must deposit them in a platform for exploitation by others. Also for publications, the EC created the platform Open Research Europe, which helps implement policies which make OS practicable and easy for researchers. As regard research data, there is a crucial difference between FAIR and open data: the EC rules for HE require data to be stored in a repository with a persistent identifier, while allowing those data to be protected through an embargo for exploitation. The overarching principle is that data must be as open as possible and as close as necessary.

Vassanelli:

Q1&Q2: I see a sort of uniformity among researchers in not fully understanding the importance of OS. Even if there has been a boost to OS requirements in the HE funding requirements by the EC, the output from EU countries for OS has not shown large differences to my view. Perhaps, in the future the Northern countries will be better equipped and prepared for OS, but they are not that much for now yet. On a local level, researchers must manage OS on their own, without institutional support.

Q3: Editors have huge power to decide about the researchers' careers even if this is filtered by peer-review and this is not correct. Quality assessment should consider also the researchers' skill to attract fundings and their capability to be innovative.

5. Conclusions

Instead of drawing conclusions and outlining future actions, this document lays ground for further internal reflection, effectively feeding into deliverable report D3.2 by way of posing specific questions to the senior management of the individual partner institutions, which emanate from the contents of this report. D3.2 will collate the answers to those questions and elaborate comprehensively on commonalities, differences, and opportunities for joint effort in the regard of the transformation of research assessment and the consequent adaptation of the profile of academic researchers.

The remainder of this section, therefore, presents those questions, providing the background and rationale to each of them as it reflects from the work carried out in WP3 in the Year 1 and the contents of this document.

The structure of this document differentiates between *top-down* and *bottom-up* views of the dynamics (for reflections, intellectual forces, trial experiments) that may sustain changes to the research assessment processes. The top-down dimension is often 'parachuted' on the researchers, without them being listened to and possibly even being understood for skills, and specialties and vocations, and outsourced to central autocratic bodies. The bottom-up dimension, where it is allowed to exist, is naturally closer to where research actually takes place, but may be very difficult to comprehend and manage for any institutions.

This dichotomy warrants the following question to be posed to the senior management of the partner institutions.

> **Question #1**: To what extent do the research evaluation criteria at your university attempt to respond to the different research disciplines, the diversity of research products, and the corresponding research culture? Are you planning or anticipating changes toward a more bottom-up design of research evaluation?

The core part of this document elaborates on various manifestations and implications of *Open Science* and its various ramifications. This notion warrants the following question.

> **Question #2**: Does your university regard the transition to Open Science as a strategic priority? If so, how is it going to influence your research assessment approaches? (If not, why?) What are the major challenges and obstacles that you anticipate in that transition?

Section 3 of this document reports initiatives of some Arqus Alliance institutions to favor the adoption of *narrative curricula*. As yet, however, there is no uniformity and general consensus on going that way. The drive toward narrative curricula may be seen as a reaction to the many drawbacks of having research evaluation depend on quantitative indicators. A serious (as opposed to superficial) transition is a very important endeavor. Reluctance to undertaking it may therefore be comprehended, although not applauded. This conundrum warrants the following question.

> **Question #3**: Is your university prepared to adopt narrative curricula in place of profiles centered on quantitative indicators in research evaluation? If so, what quality indicators do you consider using? How do you plan, for effort, complexity, and scale, to establish fair and systematic measures of qualitative evaluation? Do you fear risks of bias arising from this approach, which ones in particular?

Section 3 of this document also reports initiatives by some project partners to accompany the career of their research personnel with periodic *personalized objective-based evaluations*. This process is an

important and defining undertaking for an institution, and may arise from a local and autonomous initiative or respond to central requirements. This approach reflects practices that, however commendable, are still rare to see on a continental scale. Supposedly, moving from a totally uniform scheme of rewarding (with the attraction of its mechanistic and therefore impersonal spin) to one that allows individual tailoring without risking arbitrariness is a very complex and effort-intensive endeavour. The position of individual institutions in the Arqus Alliance with regard to this issue therefore warrants the following specific question.

> **Question #4**: Does your university contemplate the adoption of personalized objective-based evaluation of the academic achievements of your scientific personnel attached to career- or salary-grade compensations? If not, for what reasons? If so, is that effort part of a national-level policy of change or it is your own internal initiative? What challenges do you anticipate in the implementation of any such initiative?

Several members of the Arqus Alliance report having deployed own portals intended to store research products for various purposes, from periodic research evaluations (of individuals, units or entire establishments) to dissemination. It is easy to realize that different objectives pose different requirements (research products for traditional research evaluation have near-zero intersection with dissemination, outreach, third-mission artifacts) and therefore adopting a single portal for them all is unsound. At the same time, multiplying the spectrum of platforms on which individual researchers have to exhibit their products increases their effort for dubious and uncertain benefits. Furthermore, imagining that the simple act of uploading research artifacts makes an impact toward dissemination and stakeholder engagement is illusory at best. Engagement and science communication are complex endeavors that are counterproductive if treated naively, and that take training, competence, and specialization to be carried out soundly. What the Arqus Alliance institutions are doing to arrive at a mature position on determining what research products need exposing, how to do so, and how to evaluate and reward them, warrants the following question.

> **Question #5**: Does your university adopt a research-product repository portal, and for what purpose (research evaluation, independence from commercial solutions, outreach, other)? Do you consider such an infrastructure adequate to meet the emerging Open Science requirements? Is populating them simply an obligation for researchers or does it carry some motivating reward?

Appendix A: Supplemental material

Appendices A.1 and A.2 reproduce the two questionnaires we submitted to the speakers we invited to our webinar series and subsequent debates. We report them here to allow the reader to appreciate the drive we attempted to give to the reflections contributed by the invited speakers.

A.1 Top-down questionnaire: institutional perspective

1. About the presenter

- In which way are you involved with research assessment?
- What do you regard as the principal criteria currently in use at your institution to assess research?
    - To what extent do they include non-bibliometric elements? [Which ones? At which level? For which specific purpose?]
- What do you regard as the strengths and the weaknesses of the current research assessment criteria?
- What do you personally think of alternative assessment approaches (AAA) that value Open Science, Citizen Science, Science Communication and Third Mission and Outreach Activities?
    - Do you consider them desirable, despicable, ill-based, ...? Why?

2. Views on the national context

- To what extent, are AAA as above defined, contemplated in your institution's strategy?
    - If so, how does that reflect in the researcher assessment criteria?
- Are AAA a matter of national policy?
    - What autonomy do universities in your Country have in the way they evaluate research?
- Do you know of any attempts at local or national level to implement such AAA or any other forms of them?
    - To the best of your knowledge, what worked, what didn't work in them and why?
- Who are / should be the drivers for a shift towards AAA?

3. Recommendations

- What do you reckon is needed to implement AAA successfully and sustainably at local and continental (European) level?
    - Can you think of any concrete steps?
    - What do you consider the main obstacles in that direction?
- What do you think should be measured in research assessments to favour high quality, impact and diversity?

A.2 Bottom-up questionnaire: researcher's perspective

1. About the presenter

- How long have you been running / involved in third mission and outreach activities?
- Did your projects stem from a top-down (e.g. institutional interest) or bottom-up (e.g. individual interest) initiative?
- How many projects of that kind (known by you) have been activated in your institution?
- Are there measures in your institution that favour the emergence of researcher's interest in third mission and outreach activities (hubs, funding, competitions)?
    - What current incentives, if any, motivate you to continue your efforts in this ambit?
    - Can you think of (other) effective measures that could boost incentive?

2. Functioning of third-mission and outreach projects

- In your view, to what extent third-mission and outreach practices differ (e.g., in organization, execution, distribution of effort, resource constraints) from classic research projects? Can you highlight the principal differences?
- How do you approach, involve and interact with local stakeholders in your third mission and outreach projects/activities?
  - Which format do you use for communicating project goals and results to stakeholders?
  - How does the technical academic language relate to the stakeholders' communication needs?

3. Institutional context (local, national and international)

- Do you regard third mission and outreach projects as career-enhancing? How far do they help the researcher to stand out from those doing only classic research?
- Do you feel the current institutional context and national/local funding schemes are helping in developing good practices concerning third mission and outreach activities?
- Are you aware of national assessment criteria for third mission and outreach activities?
- Do you know of similar experiences abroad? If yes, have you discussed good practices about open/citizen science and third mission with colleagues involved in these projects?

4. Transferability

- How far do you think the practices that you have applied or that you have seen applied in third mission and outreach activities are transferable to other scientific ambits? If you think they are not, can you elaborate on the degree of specificity of these practices?

5. General recommendations

- In your view, which achievements in research overall should receive more recognition?
  - What is currently not or insufficiently included in the career assessment?
  - How would you like your researcher's work to be assessed? Which aspects should be more valued?
- In your view, how does the (local, national, international) research community assess third mission and outreach activities? Have you discussed good practices in assessment with colleagues involved in third mission and outreach activities?