Ginevra Peruginelli, Sebastiano Faro
Tommaso Agnoloni, Mimi Urbanc, Michael Ochsner

Editors

RESEARCH EVALUATION AND SCIENTIFIC AUTONOMY UNDER PRESSURE

Conference Proceedings

RESSH 2026 - Research Evaluation in the Social Sciences and Humanities Conference

Firenze, April 22-24, 2026

ENRESSH

igsg
ISTITUTO DI INFORMATICA GIURIDICA E SISTEMI GIUDIZIARI

Consiglio Nazionale delle Ricerche

Published in April 2026 by
Consiglio Nazionale delle Ricerche
Istituto di Informatica Giuridica e Sistemi Giudiziari

ISBN: 9788890576485
Doi: https://doi.org/10.32091/RESSH2026

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Cite as: Peruginelli, G., Faro, S., Agnoloni, T., Urbanc, M. and Ochsner, M. (eds.) (2026). *Research Evaluation and Scientific Autonomy under Pressure: Conference Proceedings*. RESSH 2026 Conference, Florence, 22–24 April 2026. Florence: CNR-IGSG. https://doi.org/10.32091/RESSH2026

La Conferenza RESSH 2026 a Firenze

Questo volume raccoglie gli atti della Conferenza RESSH 2026, ospitata a Firenze dal 22 al 24 aprile nell'ambito della serie di conferenze RESSH promosse da ENRESSH – European Network for Research Evaluation in the Social Sciences and Humanities, la rete internazionale che riunisce la comunità scientifica impegnata nello studio della valutazione della ricerca nelle scienze umane e sociali.

Questa edizione è organizzata congiuntamente da ENRESSH e dall'Istituto di Informatica Giuridica e Sistemi Giudiziari del Consiglio Nazionale delle Ricerche (CNR-IGSG), che curano anche il volume. RESSH 2026 riunisce oltre 170 autori provenienti da 30 differenti paesi attorno a un tema di stringente attualità: la valutazione della ricerca e l'autonomia scientifica nell'era delle trasformazioni geopolitiche, delle crescenti preoccupazioni per la sicurezza e della onnipresenza dell'intelligenza artificiale.

I contributi raccolti spaziano tra prospettive concettuali, empiriche e di policy, affrontando questioni cruciali quali la libertà scientifica, l'integrità della ricerca, la diversità delle pratiche valutative, l'impatto delle metriche sulle carriere e sui percorsi accademici e il ruolo crescente degli strumenti di intelligenza artificiale nei processi di valutazione. Il programma di RESSH 2026 si articola in due interventi di apertura, tredici sessioni tematiche, una sessione poster e un evento collaterale dedicato alla presentazione di un volume sui temi di interesse della Conferenza, riflettendo la ricchezza e la pluralità del dibattito in corso a livello internazionale.

Un ringraziamento sentito va ai componenti del Comitato scientifico e di programma, ai presidenti delle sessioni, alle due relatrici invitate e a tutti i revisori, il cui lavoro ha contribuito a definire la struttura e la qualità del programma. La più profonda gratitudine è rivolta agli autori: sono infatti i loro contributi a dare vita a questo volume e ad alimentare un dibattito vivo, ricco e stimolante, testimonianza del valore e della qualità del loro lavoro.

TABLE OF CONTENTS

CONFERENCE COMMITTEES 1

INTRODUCTION TO THE RESSH 2026 CONFERENCE / INTRODUZIONE ALLA CONFERENZA RESSH 2026 5

KEYNOTES 11

Funding Research in Time of Crises 12

Open Science, Research Assessment Reform, and Autonomy of Research(ers). A Critical Discussion 14

1.1 ARTIFICIAL INTELLIGENCE AND RESEARCH EVALUATION 17

Artificial Intelligence in Scientific Publishing: Possibilities, Threats, and Ethical Boundaries 18

Negotiating Institutional, Commercial and Machinic Autonomy: Generative AI, Research Evaluation and the Politics of Digital Infrastructures in the UK 21

Human Judgement in the Age of AI: Peer Review Practices and Trust Challenges in the SSH: Insights from the University of Zadar, Croatia 25

Objective vs. Subjective Evaluation in the Context of AI: What is "Objectivity" in Evaluation? 29

1.2 RESEARCH EVALUATION REFORMS, GOVERNANCE AND SCIENTIFIC AUTONOMY 33

Confronting the Reforms of Italy's Universities. The Experience of the Rete delle Società Scientifiche 34

The Impact of Global and European Political, Technological, and Epistemic Transformations on the Belgian Federal SSAH Research Ecosystem 35

Opening Science and Research Information by Law? 40

Research Evaluation as Academic Financialization: How Banking Logics Reshape the Social Sciences and Humanities 43

2.1 QUALITY, VALUES AND CONCEPTUAL FRAMEWORKS 47

Shielding Transformative Niches: From Value Capture to Collective Capabilities 48

What We Talk About When We Talk About Quality 50

Development of a Values-Based Approach to Research Evaluation and Impact in the Arts and Humanities 53

2.2A RESEARCH EVALUATION PRACTICES ACROSS INSTITUTIONAL AND PERIPHERAL CONTEXTS 57

From Commitment to Practice: Implementing ARRA Commitment 2 in Career Progression Procedures at the National Research Council of Italy 58

Reforming Research Assessment within an Italian National Research Center: Hurdles and Path Forward 62

The Internationalization of Science in Latin America: A Comparative Assessment Among Seven Universities of the Montevideo Group 64

Research Activity in Peripheral University Contexts: The Case of a Spanish University Campus Located in North Africa 69

Beyond Traditional Metrics: Key Value Indicators for Value-Based Research and Innovation Evaluation 73

2.2B CO-CREATION, COLLABORATION AND OPEN BIBLIOMETRIC INFRASTRUCTURE TO ACCOUNT FOR DIVERSITY 77

Journal Reputation and Scientific Autonomy in Open Access Publishing 78

Co-Creating Narrative Research Assessment: The COMON Experiment at Kyoto University 81

They Collaborate, But Do They "Co-Produce?" Examining Academic Scientists' Collaborative Patterns with Government 83

Discoverability and Open Access: Library Strategies and Impacts on Research Evaluation 87
From Third Mission to Knowledge Valorisation: Societal Impact and SSH Case Studies in the Italian VQR 2020–2024 92
2.3A MULTILINGUALISM ACROSS DISCIPLINES AND RESEARCH EVALUATION 97
The Influence of Multilingualism and Output Formats on the Development of Research Topics in the Social Sciences and Humanities 98
Assessing the Visibility of Non-English Scholarly Literatures in Open Research Information Infrastructures: The Case of Japanese-Language SSH Publications 104
Multilingualism in Research Assessment: Insights from Comparative Surveys in Italy and Hungary 109
Multilingualism and Metadata: How Minority-Language Protections Influence Metric and Technology Skepticism in Canadian Research Evaluation 114
Exploration of the Pasteur's Quadrant in the Social Sciences 118
2.3B CAREER, GENDER AND EFFECTS OF EVALUATION 121
Reimagining Academic Career Assessment: Insights from Global Reform Initiatives to Broaden Evaluation of Academic Talent 122
Promotion as a Temporal Regulator of Scholarly Work in the Social Sciences and Humanities 125
How Do Humanities Researchers Evaluate the Importance of Research Outputs? Evidence from a Questionnaire Survey in Japan 128
Research Inactivity in Social Sciences: An Incomplete and Biased Metric 132
2.4A CRITERIA, EXCELLENCE AND DYNAMICS IN RESEARCH EVALUATION 137
What Is Inclusive Excellence? Reflections on Excellence and Equality in Research Evaluation 138
See You in Court! Judicial Dynamics in Responsible Research Evaluation 141
Under Evaluation Pressure: How Research Evaluation Shapes Academics' Community Engagement Practices 145
Stability and Change: Disciplinary Heterogeneity and Regional Dynamics in the Research Evaluation of the Humanities and Social Sciences in Japan 149
2.4B GOVERNANCE AND POLITICS OF RESEARCH SYSTEMS 153
University Journals in a Stratified Evaluation Regime: Autonomy, Evaluation, and State Control of Scholarly Publishing in Ukraine (as a Case) 154
Strategic Funding Priorities, Scientific Sovereignty, and the Political Economy of Science: The Brazilian Experience in Science, Technology, and Innovation Policy 158
From SDGs to ISO Standards: An Opportunity More than a Constraint 162
Loneliness as an Evaluative Challenge for Public Institutions: Towards a Framework for Recognising Co-Created Research Contributions 164
2.5 SOCIETAL IMPACT AND PUBLIC ENGAGEMENT 169
Advancing Research on the Knowledge Exchange of the SSH with Society 170
The Future of Research Societal Impact in Universities and Public Research Organizations: What Changes We Need? 174
Beyond Outputs: Measuring Public Engagement in Social Sciences 177
Integrating Societal Impacts into Evaluation Frameworks: Introduction to and SSH-Related Comments on the Report from the CoARA WG Transformative Research Assessment 182
3.1 OPEN RESEARCH AND INFRASTRUCTURE 187
Infrastructural Constraints on the Move towards Open Research Information Systems: Analysing the Invisible Work of Metadata Curators 188
Current Practices and Future Challenges in Open Peer Review among Canadian Scholarly Journals 192

Open Science and its Role in Research Evaluation: Cultural Change, Open Infrastructures and Scientific Autonomy at the University of Milan 195
Open Science and New Research Assessment Criteria: A Two-Ways Path 198
EU and National Policies on Research Security: Implications for Open Science and Researchers' Freedom 203
3.2 BIBLIOMETRICS, LLMs AND RESEARCH INTEGRITY 207
Topic Concentration in SSH Funding Portfolios: A Large Language Model Assisted Quantitative and Qualitative Analysis 208
Can Large Language Models Validly Identify Comparable Peers for Researcher Benchmarking? 215
Evaluating a Machine-Learning Based Reviewer Matching Tool across Disciplines: Early Evidence of Field-Sensitive Performance and Implications for SSH 221
Paper Mills: Objects of Fraud and Mirrors of Managerial Research Governance 224
3.3 RETHINKING INDICATORS, METRICS AND EVALUATION TOOLS FOR SSH 227
Counting the Uncountable? Bibliometric Infrastructures, Interdisciplinarity, and Responsible Research Evaluation: The Case of CNR-ISPC 228
Bibliographies and the bibliometrics of books. 232
New tools for a fair research assessment in humanities 232
Points of Friction between CoARA Principles and Current Evaluative Frameworks at the National, Institutional and Individual Level: Preliminary Results from the RIF4HSS Project 248
Next Generation SSH-Evaluation: From Indicators to Ontologies 251
POSTER SESSION 255
Latin America & Caribbean Retraction Watch: A Proposal Web Directory 256
Inclusive Co-Evaluation in Social Innovation and Research: Case Study and Methodological Innovation of a Training Program in Entrepreneurial Skills with Refugee Women 260
The Commitment of the University of Florence Library System to Open Science and Excellence in Research 262
Towards a Strategic View of Research Assessment: The CoARA National Chapter of Italy 2026-27 263
Perspecti: A Comprehensive Research and Citation Database for Scholarly Works in the Humanities and the Arts 265
CoARA Expect: Videos to Promote Responsible Researcher Assessment 267
AIbox – Dashboarding Algorithmic Analysis Tools for Application Processing in a Research Funding Organisation 268
Diamond Open Access in Italy: Evidence and Implications for Research Evaluation 269
Cultural Heritage as Research Output: Glam Wikimedia Projects at the CNR Central Library 271
Retractions in the Social Sciences and Humanities: Descriptive Evidence from Retraction Watch Notices 273
Finnish Career Assessment Matrix (FIN-CAM): Supporting Responsible Researcher Assessment 275
COLLATERAL EVENT - BOOK LAUNCH 277
"The Knowledge of Humanities" by Andrea Bonaccorsi (2026): First Monograph in the Brepols' Series "Research on Research in the Arts, Humanities and Social Sciences" 278
INDEX OF KEYWORDS 281
INDEX OF AUTHORS 285

CONFERENCE COMMITTEES

Conference organisers

Institute of Legal Informatics and Judicial Systems – National Research Council (CNR-IGSG), Italy

ENRESSH – European Network for Research Evaluation in the SSH

Programme committee

Ginevra Peruginelli, Institute of Legal Informatics and Judicial Systems – National Research Council (CNR-IGSG), Italy

Michael Ochsner, ENRESSH – European Network for Research Evaluation in the Social Sciences and Humanities

Mimi Urbanc, Research Centre of the Slovenian Academy of Sciences and Arts (ZRC SAZU), Slovenia

Sebastiano Faro, Institute of Legal Informatics and Judicial Systems – National Research Council (CNR-IGSG), Italy

Tommaso Agnoloni, Institute of Legal Informatics and Judicial Systems – National Research Council (CNR-IGSG), Italy

Maddi Abdleghani, Sorbonne Université, CNRS, Groupe d'Étude des Méthodes de l'Analyse Sociologique de la Sorbonne, GEMASS, France

Scientific committee

Michael Ochsner
President of ENRESSH
FORS, Switzerland

Mimi Urbanc
Chair
Research Centre of the Slovenian Academy of Sciences and Arts (ZRC SAZU), Slovenia

Ginevra Peruginelli
Chair of Organising Committee
Institute of Legal Informatics and Judicial Systems, National Research Council (CNR-IGSG), Italy

Abdelghani Maddi
CNRS and Sorbonne University, France

Alesia A. Zuccala
University of Copenhagen, Denmark

Alex Rushforth
University of Leiden, Netherlands

André Brasil
University of Leiden, Netherlands

Andrea Bonaccorsi
University of Pisa, Italy

Elena Giglia
University of Turin, Italy

Emanuel Kulczycki
Adam Mickiewicz University, Poznan, Poland

Emanuela Reale
National Research Council, Italy

Francesca Di Donato
Institute of Computational Linguistics "A. Zampolli", National Research Council, Italy

Gemma Derrick
University of Bristol, United Kingdom

Geoffrey Williams
Université de Bretagne Sud, France

Gunnar Sivertsen
Nordic Institute for Studies in Innovation, Research and Education, Norway

Jack Spaapen
ScienceWorks, Netherlands

Janne Pölönen
Federation of Finnish Learned Societies, Finland

Jelena Brankovic
Robert K. Merton Center for Science Studies, Humboldt-Universität zu Berlin, Germany

Jon Holm
Research Council of Norway, Norway

Julia Olmos Peñuela
University of Valencia, Spain

Kamila Lewandowska
University of Warsaw, Poland

Lai Ma
University College Dublin, Ireland

Laura Rovelli
CLASCO/FOLEC, Argentina

Lucia Cespedes
Erudit, Canada

Lynne Bowker
Université Laval, Canada

Mikael Laakso
Tampere University, Finland

Mira Söderman
Federation of Finnish Learned Societies, Finland

Nataša Jermen
The Miroslav Krleza Institute of Lexicography, Croatia

Paola Galimberti
University of Milan, Italy

Reetta Muhonen
Tampere University, Finland

Sebastiano Faro
Institute of Legal Informatics and Judicial Systems, National Research Council (CNR-IGSG), Italy

Tim C.E. Engels
University of Antwerp, Belgium

Tommaso Agnoloni
Institute of Legal Informatics and Judicial Systems, National Research Council (CNR-IGSG), Italy

Vas Virag
HUN-REN, Research Centre for Natural Sciences, Institute of Molecular Life Sciences, Hungary

Zehra Taşkın
Hacettepe University, Institute of Informatics, Turkey

Zoe Hope Bulaitis
University of Birmingham, United Kingdom

Local organising committee

Institute of Legal Informatics and Judicial Systems, National Research Council (CNR-IGSG), Italy

Ginevra Peruginelli
Sebastiano Faro
Tommaso Agnoloni

Massimo Cristofaro
Fulvia Ciurlia
Gerardo Giardiello
Teresa di Candilo
Sara Conti
Silvana Scarpi

INTRODUCTION TO THE RESSH 2026 CONFERENCE / INTRODUZIONE ALLA CONFERENZA RESSH 2026

The RESSH Series and ENRESSH

RESSH 2026 is the latest edition of a conference series initiated and sustained by the European Network for Research Evaluation in the Social Sciences and Humanities (ENRESSH). Since its founding, ENRESSH has established RESSH as the principal international forum for the critical examination of research evaluation in particular regarding the Social Sciences and Humanities (SSH), building a community of scholars, policymakers, and practitioners whose collective work has shaped both the field of research evaluation as well as the reform agendas now visible at European and national levels.

The association's credo is to support and integrate the diversity in views and voices of all stakeholders, disciplines, and countries and to thus promote a pluralist approach to research evaluation. While the association started on the basis of Social Sciences, Humanities and Arts scholars, interest in ENRESSH's approach to research evaluation has spawned the interest of scholars (and evaluators as well as policy makers) concerned with Science, Technology, Mathematics and Engineering (STEM) disciplines as well, given the rise of the same concerns and criticisms in those areas that have been voiced in the SSH for a long time. Similarly, while the root of ENRESSH lies in a European network that sought to influence European science policy, the network grew considerably and its membership now extends to several continents, reflecting the inherently global nature of research and progressively building a more inclusive international community. Against this background, ENRESSH retains its European identity while fostering international engagement and dialogue.

Besides regular smaller workshops organised since 2012 as an informal network of evaluation specialists in the Social Sciences and Humanities, the international RESSH conference has been an important forum among stakeholders engaged in research evaluation reform since the foundation of the international association in 2014. Former conferences took place in Rennes (2015), Antwerp (2017), Valencia (2019), Galway (2024) and Helsinki (2025).

CNR-IGSG and the 2026 edition

The *Istituto di Informatica Giuridica e Sistemi Giudiziari* (Institute of Legal Informatics and Judicial Systems) of the National Research Council (CNR-IGSG) is honoured to serve as local organizer of RESSH 2026 and as co-editor of this Conference Proceedings. These roles are embedded within the broader framework of ENRESSH and reflect for CNR-IGSG not only an organizational responsibility, but also a deeper intellectual affinity.

The Institute has occupied a specific position within the Italian and European research landscape, working on the structures, norms and technologies through which law, institutions and knowledge systems interact. Since 2016, CNR-IGSG has performed research studies on the evaluation of legal scholarship and, more broadly, on the methodological and institutional challenges of assessing research in SSH. The study of how legal science is produced, communicated and evaluated and of how evaluation criteria transform and sometimes distort scientific priorities, has been a part of the Institute's research agenda.

CNR-IGSG's involvement in the ENRESSH network and community reflects a long-standing commitment to these issues, situating the Institute within a wider international effort to critically examine and advance research evaluation practices in the SSH. In this perspective, hosting RESSH 2026 represents both a continuation of

ENRESSH's ongoing work and a context in which CNR-IGSG's own research trajectory converges with the broader aims of the network.

The Florentine Conference at the right moment

RESSH 2026 gathers scholars, policymakers, research managers and practitioners from 30 countries, a breadth of representation that is itself a statement about the global stakes of the questions under discussion. The diversity of national contexts, disciplinary traditions and institutional positions represents an intellectual resource as well as a condition for the kind of comparative, cross-contextual analysis that a field as politically and institutionally embedded as research evaluation requires. The contributions collected in this volume span conceptual, empirical and policy-oriented perspectives. They address key transformations in research evaluation across SSH disciplines and beyond and reflect a wide range of methodological approaches, including case studies, theoretical work, quantitative analyses, surveys, legal analysis.

The theme of RESSH 2026 *"Scientific Autonomy Under Pressure: Rethinking Research Evaluation in the Context of Global Power Shifts, Security Concerns, and Artificial Intelligence Challenges"* captures an ongoing condition. Research evaluation is undergoing significant transformation, driven by multiple and often divergent forces. Geopolitical shifts are reshaping the conditions of international collaboration and influencing research priorities, so that judgments about what counts as valuable or legitimate research are increasingly intertwined with political as well as epistemic considerations. At the same time, security concerns, data sovereignty, dual-use research and the governance of open science infrastructures are introducing new constraints into academic practice.

Technological developments are also playing a central role. The growing integration of digital tools across the research cycle is altering evaluation processes, while raising questions about their reliability and implications. In parallel, growing evidence of research misconduct, including fabricated data, paper mills, predatory publishing and the misuse of generative tools, has intensified scrutiny of the scientific outputs, highlighting the limits of purely technical countermeasures.

What connects these transformations is their impact on the concept of scientific autonomy. This concerns the ability of researchers, institutions and disciplines to define their own priorities, standards and modes of knowledge production.

This volume's contributors consistently uphold scientific autonomy as a core principle, even as the political, technological and institutional environments in which science operates continue to evolve.

The program and its sessions

The program of RESSH 2026 is organized across two keynotes, 13 thematic sessions distributed over the conference days, a poster session and a book launch as a side event.

The technological transformation of evaluation is addressed in Session 1.1, where the integration of artificial intelligence into scholarly communication foregrounds questions of objectivity, bias and epistemic control. The institutional and political dimensions of research assessment reform are examined in Session 1.2, which situates

evaluation within broader governance frameworks and highlights tensions between autonomy, regulation and systemic change.

The poster component, introduced in Session 1.3 through a dedicated flash presentation slot, constitutes an integral part of the conference's epistemic architecture. It enables the presentation of diverse research contributions, methodological approaches and empirical materials, while fostering distributed forms of scholarly interaction. In this respect, it contributes to expanding the evidentiary and discursive scope of the conference, supporting inclusiveness and plurality in research evaluation debates.

The conceptual foundations of evaluation are developed in Session 2.1, where core notions such as quality and values are critically addressed. Issues of institutional diversity and inclusion are explored in Sessions 2.2A and 2.2B, which address, respectively, the variability of evaluation practices across diverse contexts and the role of collaborative and infrastructural arrangements in accounting for that diversity. The question of diversity is further extended in Sessions 2.3A and 2.3B, focusing on multilingualism and on the effects of evaluation on careers, gender and academic trajectories.

The criteria, notions of excellence and evolving dynamics shaping research evaluation are examined in Session 2.4A, while Session 2.4B situates research evaluation within the broader political and governance structures of national research systems, including questions of scientific sovereignty, societal impact and regulatory frameworks.

The relationship between research and society is addressed in Session 2.5, which considers the challenges of conceptualizing and assessing societal impact and public engagement.

The infrastructural requirements for a responsible research assessment are the focus of Session 3.1, which highlights the practical and organizational dimensions of open research systems. The relationship between evaluation, integrity and emerging technologies is explored in Session 3.2, where new risks and systemic effects associated with computational tools are brought into view. Finally, Session 3.3 addresses the limits of existing metrics and advances alternative approaches to the evaluation of SSH research.

Taken as a whole, all the sessions do not converge on a single model of reform. Rather, they collectively frame research evaluation as a contextual and evolving institution, whose transformation requires simultaneous attention to technological infrastructures, conceptual frameworks, governance arrangements and the diverse practices through which scientific knowledge is created.

The Conference in keywords

The keywords index that closes this volume functions as an implicit map of the field's current policy agenda. Read across, it reveals a dense cluster of terms associated with ongoing reform initiatives and actions such as CoARA, the Barcelona Declaration, responsible research assessment, open research information, pointing to a community that is not merely analyzing evaluation systems, but actively engaged in their transformation. These keywords reflect a shift from critique to implementation, where principles are increasingly translated into institutional practices, infrastructures and policy tools. The recurrence of terms such as scientific autonomy, researchers' freedom and knowledge sovereignty, alongside geopolitics and research security, further

indicates that research evaluation is now widely recognized as a domain of strategic governance shaped by geopolitical, technological and security-related pressures.

A second cluster maps the diversity agenda as a central policy priority. Keywords such as multilingualism, bibliodiversity, diversity of research outputs and the diversification of career paths signal a growing effort to align evaluation criteria with the actual epistemic practices of the SSH, where monographs, non-English publications and field-specific output formats remain underrepresented in dominant metric-based frameworks. In this sense, diversity is not only a normative commitment, but also a corrective to evaluation systems that risk narrowing the scope of recognized knowledge production.

A third cluster, including paper mills, predatory journals, retraction notices, research integrity and questionable journals, captures the integrity dimension of current reform efforts. Here, the focus shifts from individual misconduct to the systemic vulnerabilities of the research ecosystem, increasingly exposed by technological developments and intensified publication pressures. This highlights the need for coordinated policy responses, improved monitoring infrastructures and stronger links between evaluation practices and integrity safeguards.

Particularly relevant, from a policy perspective, is the density of terms related to careers and their human costs: career stage, academic promotion, career progression, inactive researcher, gender, inclusive excellence and social inclusion. These keywords underscore that evaluation systems function as powerful allocation mechanisms under conditions of increased competition and scrutiny, shaping research outputs as well as career trajectories and working conditions. Their effects are unevenly distributed, often reinforcing structural inequalities, call for more inclusive, context-sensitive and responsible approaches to assessment.

A note of thanks

These Conference Proceedings represent the collective intellectual effort of all those who submitted and reviewed contributions that helped shape the RESSH 2026 program. We are deeply grateful to the members of the scientific and program Committees for their intellectual guidance, to all session chairs and the two keynote speakers for their generosity with time and expertise. The contributors (a total of 173 authors), from diverse institutions play a central role in this volume, bringing their research, questions and critical energy to this conversation and shaping both its scope and its intellectual depth.

The Editors
Ginevra Peruginelli, Sebastiano Faro, Tommaso Agnoloni, Mimi Urbanc, Michael Ochsner

Firenze, April, 16, 2026

KEYNOTES

FUNDING RESEARCH IN TIME OF CRISES

GABI LOMBARDO
European Alliance for Social Sciences and Humanities (EASSH)

Keywords: European research funding, FP10 research policy, European Union, CoARA

Designing a new European Research Funding Programme

On 16 July 2025, the European Commission published its proposal for the Multiannual Framework Programme (MFF), the 7-year funding plan for the EU (European Commission, 2025). Through a democratic process of consultations, the plan is discussed, revised and eventually approved by the other two parties the council of Member States (MS) and the European Parliament (EP). The research and innovation funding is part of this process and there is no other research funding programme in the world that has a similar level of debate and discussion where every voice counts.

In this financial cycle, defence and security have leap-frogged to the top of the list of priorities caused by the relentless waves of new global instability. Therefore, while in more normal times, defence and security are issues for the MS, negotiations for the funding of the EU are more complex and dominated by shifting global and domestic politics and by the urgency to rise European competitiveness (e.g., Draghi, 2024; Heitor, 2024, Letta, 2024). Hence the EC proposal is linking much more closely the European Competitiveness Fund (ECF) and the Framework programme (FP) for Research and Innovation.

It is this context of 'crisis' that investment in research is battling to be heard. Compared to the current funding framework, all the indications are that the next FP will have less space for truly joint collaborative research across borders, across disciplines and addressing shared challenges that run across the EU. While defence and security are likely to be taken out of the core and funded through other mechanisms (e.g., the ECF), the real debate will be between the investment in traditional research and innovation and the new competitiveness discourse, bundled in the ECF.

However, in the short and elusive proposal of the EC, it is difficult to understand - for the parties negotiating on behalf of the Member States and European Parliament – how this relationship between collaborative research and the ECF should be articulated. After all, for over 50 years, the investment for research and innovation has always been a stand-alone programme, as it was designed so by the treaties. In this cycle that is no longer a stable prior assumption.

Even within the current EC proposal a more familiar tension is being played out over the balance between research and innovation. Innovation seems to be understood only in terms of AI and deep tech. So what is happening to other forms of innovation, what space is there for the innovation that provides social benefits, and creates 'social value' to our societies? What about the innovation that fosters corporate sustainability and social progress beyond economic returns? Even creative and service industries, previously acknowledged as an emerging backbone of European economy beyond

manufacturing seems to be overshadowed by the deep-tech obsession of European policy making.

The research community plays a role in the wider debate and in the negotiation. Scholars organised in university associations, academies, disciplinary alliances are investing greatly to have their voice heard at this critical time. The European Alliance for SSH is working in this space, at the interface of the research community of Art, Humanities and Social Sciences and policy makers. In particular, this community plays a critical role because of the recent changes in the political narratives where the European social values are no longer taken for granted.

We live in complicated times and the research community today more than ever must engage with the design of the European research funding programme, because research goes beyond peers' recognition, but constitute the ground for a collective social responsibility at time of crisis. The next Framework Programme is not just a funding instrument. It is a political signal. It is a statement about what we believe, what we value and what kind of Europe we are committed to building. Let's ensure it reflects the full complexity, and humanity, of that task.

References

Draghi, M. (2024). *Report on EU Competitiveness*. European Commission. Available at: https://commission.europa.eu/topics/competitiveness/draghi-report_en

European Commission (2025). *The 2028–2034 EU Budget for a Stronger Europe*. Available at: https://commission.europa.eu/strategy-and-policy/eu-budget/long-term-eu-budget/eu-budget-2028-2034_en

Heitor, M. (2024). *Align, Act, Accelerate: Research, Technology and Innovation to Boost European Competitiveness*. Publications Office of the European Union. Available at: https://op.europa.eu/en/publication-detail/-/publication/2f9fc221-86bb-11ef-a67d-01aa75ed71a1/language-en

Letta, E. (2025). *Report on the Future of the Single Market*. European Commission. Available at: https://single-market-economy.ec.europa.eu/news/enrico-lettas-report-future-single-market-2024-04-10_en

OPEN SCIENCE, RESEARCH ASSESSMENT REFORM, AND AUTONOMY OF RESEARCH(ERS). A CRITICAL DISCUSSION

FRANCESCA DI DONATO
National Research Council - Institute of Computational Linguistics "Antonio Zampolli" (CNR-ILC), Italy

The objective of this contribution is to emphasise the critical interrelationships between open science, research autonomy and the ongoing reform of research evaluation, which involves institutions and researchers. The issue is examined from a philosophical-political perspective and is based on a historical reading of the problem which, when contextualised, helps us to reflect on the implications for the present day.

To develop this argument, the presentation will begin by focusing on some of the current challenges which pose pressure on research autonomy. These challenges for (open) science have been brought to the fore by the current geopolitical landscape and recent technological changes. The failure of the cosmopolitan ideal and the return to the ideal of secrecy and sovereignty are evident, with an unclear role for pluralism and epistemic diversity. The reliance on fragile digital systems (sustainability is a problem) and the dominance of big private companies over research platforms, tools and data are also apparent. Furthermore, the rise of misinformation and the problem of assessing research quality become more important (Leonelli, 2026). In addition, in the context of global power shifts and geopolitical challenges, it is imperative to consider the potential of research evaluation systems as instruments of power. The question that arises is: who determines the practices and content that are worthy of recognition? And to what extent can the pressure to produce high-impact research (driven by nations or institutions) undermine scientific integrity or reduce researchers' autonomy?

The subsequent discussion will then proceed to consider how research autonomy is conceptualised in the Agreement on Reforming Research Research Assessment (ARRA). The recent reform promoted by CoARA enshrines the autonomy of research (of institutions and researchers) as one of its core principles. Indeed, the ARRA stipulates the direct involvement of individual academics and scientific communities in the definition of new criteria and processes, emphasising that criteria, data, infrastructures and processes must be in the hands of scholars (Arentoft et al., 2022). Furthermore, the concept of quality is not perceived as a static value, but rather as an ongoing, procedural process that serves to uphold the integrity of the scientific method. In essence, the notion of quality/qualitative assessment is intrinsically linked to the paradigm of open science, which is characterised as a persistent issue that remains unresolved (Di Donato, 2024).

The last part of the presentation will address the relationship between research autonomy and open science by examining a specific discussion that occurred in Prussia between the last quarter of the eighteenth and the first quarter of the nineteenth century. In the 18th century, universities in the German states experienced a period of significant challenges, precipitated by wars that resulted in significant population decline, a severe economic crisis characterised by rising living costs, low

purchasing power due to stagnant salaries, and a labour market crisis. Concurrently, there was a pervasive belief that obtaining a degree did not guarantee a higher standard of living (McClelland, 1986). In this context, Immanuel Kant and Wilhelm von Humboldt's university theories define the conditions for free and open science by establishing the role of research infrastructures (universities and academies) within the paradigm of scientific autonomy. In his essay on the Enlightenment (1784), Kant advocates a conception of freedom of public reason from institutions, only hinting at a future institutional evolution that would allow for freedom within them. A few years later, in a different political context, the philosopher tackles the problem of research quality by discussing the conditions for public research to be defined as such (Kant, 1798). In fact, according to Kant, research is considered to be autonomous and continuously subject to criticism by experts, whose status is grounded in scientific merit. Consequently, research cannot be other than open, collaborative and inclusive. From this perspective, scientific debate among peers plays a pivotal role. However, this is only possible if the political authorities guarantee the scientific republic and its infrastructures (namely, the university) complete freedom of research (Di Donato, 2006). Humboldt's (1809-10) conception of universities as emulating the open-ended nature of science posits that the state should foster and protect research autonomy by ensuring negative and positive freedom through a minimalist and reflexive regulatory framework, as well as recruiting teachers and academics in a pluralistic manner (Pievatolo, 2025). The conclusion will illustrate how this model can serve as a paradigm for open science both foreshadowing contemporary contradictions and suggesting how to address them.

References

Di Donato, F. (2006) 'Università, scienza e politica nel *Conflitto delle facoltà* di Kant', *Bollettino Telematico di Filosofia Politica*. Available at: https://doi.org/10.5281/zenodo.7543877

Di Donato, F. (2024) 'What we talk about when we talk about research quality: a discussion on responsible research assessment and open science', *Bollettino Telematico di Filosofia Politica*. Available at: https://doi.org/10.5281/zenodo.10890788

Humboldt, W. von (1809–1810) 'Über die innere und äußere Organisation der höheren wissenschaftlichen Anstalten in Berlin', in Markschies, C. (ed.) *Gründungstexte*. Berlin: Humboldt-Universität zu Berlin, pp. 229–241. Available at: https://doi.org/10.18452/4653

Kant, I. (1784) 'Beantwortung der Frage: Was ist Aufklärung?', in *Kants gesammelte Schriften* (KGS), vol. VIII. Berlin-Leipzig: W. de Gruyter. English translation in Reiss, H.S. (ed.) (1989) *Political writings*. Cambridge: Cambridge University Press.

Kant, I. (1798) *Der Streit der Fakultäten in drei Abschnitten*, in *Kants gesammelte Schriften* (KGS), vol. VII. Berlin-Leipzig: W. de Gruyter. English translation in Gregor, M. (1979) *The conflict of the faculties*. New York: Abaris Books.

Leonelli, S. (2026) 'AI versus open research: exploring the evolving tensions (opening keynote session)', *Oxford Forum of Open Scholarship*, 2 March. Available at: https://ora.ox.ac.uk/objects/uuid:df1b8963-7055-452f-b1d6-f56f4f0a119c/files/sn870zt679 (Accessed: 16 April 2026).

McClelland, C. (1980) *State, society and university in Germany 1700–1914*. Cambridge: Cambridge University Press.

Pievatolo, M.C. (2025) 'Science as "a problem not yet fully resolved": universities and the public use of reason between Kant and Humboldt', *infoclio.ch Conference 2025: Open Science in History*, Bern. Zenodo. Available at: https://doi.org/10.5281/zenodo.18070386

1.1 ARTIFICIAL INTELLIGENCE AND RESEARCH EVALUATION

ARTIFICIAL INTELLIGENCE IN SCIENTIFIC PUBLISHING: POSSIBILITIES, THREATS, AND ETHICAL BOUNDARIES

LEENA WAHLFORS, JANNE PÖLÖNEN
Federation of Finnish Learned Societies, Finland
MIKAEL LAAKSO
University of Tampere, Finland

Keywords: Artificial intelligence, Scholarly publishing, Scientific publishing, Peer review, Editorial office

Artificial intelligence (AI) has rapidly become, and is here to stay, as a part of all areas of life, including the scholarly publication process. AI can be used to screen manuscripts, to suggest peer reviewers, and even to make decisions about what gets published.

The study aims to examine how AI is used at different stages of the scholarly publication process, what benefits it may offer, and where the boundaries of ethical and responsible use lie. AI adoption raises concerns related to fairness, transparency, and accountability. Specifically, the study asks how AI can enhance the efficiency and quality of scholarly publishing—and at what point its use introduces ethical risks. Interesting questions to be answered also are when AI is consciously used in an ethically wrong way, when its misuse is due to ignorance, or when its use is even avoided unnecessarily.

In addition, we want to find out, whether publishers have policies on the use of AI—for example, how much, or in what ways, authors and peer reviewers, or editors themselves are allowed to use it.

Methods

For the study, a literature review was conducted to examine the use of AI in the scientific publishing process and to consider the beneficial and problematic uses of AI at different stages of the process.

The empirical data for the study consists of two sources: 23 structured discussions with the Publication Forum's evaluation panel, which took place at panel meetings in the fall of 2025, and a Lime Survey questionnaire sent to the same group. The panel discussions explored the panelists' experiences and knowledge of the use of artificial intelligence in the scientific publication process: whether they themselves have used artificial intelligence at any stage of the process in their role as a research manuscript author, peer reviewer, or editorial board member, and what knowledge or experience the panelists have of the use of AI by other actors at different stages of the publication process. The discussion proceeded with key questions related to the stages of the manuscript publication process, from manuscript submission to peer review, and concluded with considerations regarding publication or rejection decisions.

In a Lime Survey questionnaire sent to all panelists in September 2025, respondents were given the opportunity to provide open-ended answers to the same questions that

had been asked at the panel meetings. The survey was sent to panelists by email. By the time the survey closed in December 2025, 128 open-ended responses had been received.

Preliminary results

According to research literature, AI makes the work related to the scholarly publication process first step, the initial quality check, more efficient, but at the same time introduces new types of possible errors. AI tools may flag a text as fraudulently produced even when it has been written in accordance with good scientific practice or fail to detect a genuine problem. AI tools may also generate biases. So-called hallucinations and fabricated references are also a risk. (Hosseini & Resnik, 2025.)

There was a consensus among panelists that it is acceptable for editorial offices to use AI for tasks such as plagiarism detection and language checking. Nevertheless, decisions about whether a manuscript should advance to the next stage, for peer review, must always remain the responsibility of a human editor.

The panelists considered it problematic that editorial offices require researchers to disclose whether, or how, they have used AI, while the editorial offices do not nearly always operate transparently. Openness may soon become a competitive advantage for publishers, as one panelist pointed out. One panelist summed up the idea by saying: "we don't know, and that's why we have to suspect everything… an age of suspicion."

The panelists observed that editorial guidelines for researchers on the use of AI are inadequate and mutually inconsistent. They also raised, both from the perspective of peer reviewers and of researchers, the concern that it is not always possible to report AI use, because AI may be used unknowingly as part of available digital tools. Researchers' willingness to disclose AI use to the editorial office may also be reduced by fear that the manuscript will be rejected on that basis, even if the use was limited, for example, to language editing. Similar observations have been reported in the published literature. Weaknesses in guidance on AI-related research ethics likely reflect the rapid pace of AI development as well as gaps in knowledge (da Veiga, 2025).

Studies in the field highlight the role of AI in easing the increasingly laborious task of finding a suitable pool of reviewers, but also risks, including the under-identification of interdisciplinary topics. Algorithms are also unable to think critically, for instance, to weigh whether leading figures in a field may be too busy to devote time to reviewing. It is also possible that the system suggests entirely fictional individuals as reviewers (Ahn, 2024; Kousha & Thelwall, 2023; Farber, 2024). One panel member noted that the use of AI in finding peer reviewers is more relevant for younger researchers who do not yet have extensive networks.

The peer reviewer's task is to evaluate particularly the manuscript's novelty, significance, relevance, and scientific accuracy. AI can assist with routine and technical tasks. AI tools free up reviewers' time for what matters—assessing whether the research is new, important, and well justified. Human expertise, judgment, and responsibility cannot be replaced, as research in the field suggests. (Ahn 2024; Zhang et al., 2025)

The panelists stressed the same. It is important to be able to search for information, verify details, and uncover plagiarism, but peer reviews produced entirely by AI cannot be accepted. Data protection and confidentiality must be kept firmly in mind. If the editorial office or a peer reviewer enters the manuscript under review into an AI system, its content—and potentially novel research findings of significant originality—may spread uncontrollably, several panel members warned.

If reviewers' opinions differ regarding a manuscript's relevance and rigor, the editorial office may use AI to support its weighing of what revisions to request or whether to accept the work as is. After publication, AI can also help to create summaries and translations of the manuscript (Ahn 2024; Kousha & Thelwall 2023.) Panel members had very few comments or assumptions about the use of artificial intelligence in publishing decisions. Across panels, participants highlighted both the positive and negative uses of AI, and the threats and opportunities these pose for doing science.

The panelists regarded AI as an efficient tool and reflected on where the limits of its use lie. They asked whether AI is already now eroding existing doubts about the usefulness and necessity of science. At the same time, they believe that even as AI develops, there will remain things that cannot be handed over from humans to machines.

Ways of doing science will change, but as the panelists put it: "Everything is open. The future is uncertain." "The threats are materializing… the same applies to the opportunities." "We should find a good marriage between humans and AI."

References

Ahn, S. (2024) 'The transformative impact of large language models on medical writing and publishing: current applications, challenges and future directions', *Korean Journal of Physiology and Pharmacology*, 28(5), pp. 393–401. Available at: https://doi.org/10.4196/kjpp.2024.28.5.393

Checco, A., Bracciale, L. and Loreti, P. (2021) 'AI-assisted peer review', *Humanities and Social Sciences Communications*, 8, p. 25. Available at: https://doi.org/10.1057/s41599-020-00703-8

da Veiga, A. (2025) 'Ethical guidelines for the use of generative artificial intelligence and artificial intelligence-assisted tools in scholarly publishing: a thematic analysis', *Science Editing*, 12(1), pp. 28–34. Available at: https://doi.org/10.6087/kcse.352

Farber, S. (2024) 'Enhancing peer review efficiency: a mixed-methods analysis of artificial intelligence-assisted reviewer selection across academic disciplines', *Learned Publishing*, 37(4), e1638. Available at: https://doi.org/10.1002/leap.1638

Hosseini, M. and Resnik, D.B. (2025) 'Guidance needed for using artificial intelligence to screen journal submissions for misconduct', *Research Ethics*, 21(1), pp. 1–8. Available at: https://doi.org/10.1177/17470161241254052

Kousha, K. and Thelwall, M. (2023) 'Artificial intelligence to support publishing and peer review: a summary and review', *Learned Publishing*. Available at: https://doi.org/10.1002/leap.1570

Zhang, A., Gao, Y., Suraworachet, W., Nazaretsky, T. and Cukurova, M. (2025) 'Evaluating trust in AI, human, and co-produced feedback among undergraduate students', *arXiv*. Available at: https://arxiv.org/abs/2504.10961

NEGOTIATING INSTITUTIONAL, COMMERCIAL AND MACHINIC AUTONOMY: GENERATIVE AI, RESEARCH EVALUATION AND THE POLITICS OF DIGITAL INFRASTRUCTURES IN THE UK

JOSIE R. WEST
King's Business School, King's College London, UK

Keywords: GenAI, Scientific autonomy, Research information infrastructures, Assetisation, Generativity, Technoscientific capitalism

This paper examines the role of generative artificial intelligence (genAI) within research information infrastructures in the UK and its implications for autonomy, evaluative practices and governance under technoscientific capitalism. Drawing on preliminary interview data with academic staff, research managers and commercial and open research information providers in the UK, alongside findings from a recent qualitative study of genAI in REF 2029 preparation, the paper first situates contemporary debates on research evaluation within a longer shift from bibliometrics to data infrastructures and advanced analytics. Second, it analytically brings together STS perspectives on digital research information infrastructures, the concept of generativity and political economy accounts of assetisation to conceptualise "thinking infrastructures" that actively structure evaluation and scientific autonomy. It mobilises the concept of generativity to explore the emerging capabilities of genAI in producing new (and possibility unintended) indicators and classifications from interlinked metadata. It then situates genAI within a political economy of assetisation, highlighting how research metadata are repurposed as economic assets under the market-oriented, capitalist-imperialist governance logics of REF and global ranking pressures. Finally, by conceptualising research information systems as "thinking infrastructures," the presentation examines how commercial and open providers and university administrators embed, negotiate, or resist genAI, revealing the political nature of infrastructural power and autonomy in the contemporary UK research landscape.

Context

This contribution examines how generative AI is becoming embedded within research information infrastructures and how this reshapes autonomy, evaluative practices and governance in the UK under broader conditions of technoscientific capitalism.

Over the past two decades, debates on research evaluation have shifted in harmony with the rapid advancement of technology. Bibliometrics and indicator-based assessment have increasingly expanded to encompass new types of data and data analytics that seek to capture different dimensions of research performance, such as policy impact. In this context, one can also observe the growing role of generative artificial intelligence (genAI) in research evaluation and governance. A recent rapid qualitative study of institutional uses and perceptions of GenAI in UK REF 2029

preparation and assessment identified conflicting perspectives (Watermeyer et al., 2025). Findings indicate expectations of workload reduction; resistance due to concerns over human autonomy and the marginalisation of academic judgement; a broad consensus that the REF will become increasingly AI-infused and, inevitably, fully automated; low trust linked to limited AI literacy; and uneven institutional investment shaped by financial capacity (ibid).

We extend these findings by drawing on new insights from a major cross-disciplinary, international study spanning Canada, Germany, France and the UK: '*Between Economy and Democracy: Reorganizing Research Evaluation through Metadata in the Digital Era.*' Drawing on preliminary UK interview findings with commercial and open research information and evaluation platforms and their users (such as university administrators), this paper engages core debates on GenAI and data capitalism through three complementary analytical perspectives: STS approaches to *digital information infrastructures* (Bowker & Star, 2000; Kornberger et al., 2019; Star, 1999), the concept of *generativity* (Kornberger et al., 2017; Mennicken & Kornberger, 2021), and political economy accounts of *assetisation* (Birch et al., 2021; Birch & Muniesa, 2020).

We argue that the turn towards new forms of metadata and data analytics, including genAI, in research evaluation cannot be understood simply as a technical innovation or efficiency gain. Rather, it reflects deeper political-economic dynamics in which digital research information infrastructures function as socio-material arrangements that actively structure what counts as knowledge, excellence and impact. In the UK, these developments are occurring in parallel with intensifying competitive pressures within the Research Excellence Framework (REF), where institutions are increasingly positioned as market actors 'compelled' to strive for reputational advantage and leadership in GenAI uptake (Watermeyer et al., 2025).

Contribution

The paper makes several critical contributions. First, we explore the shifting debate from bibliometrics to data infrastructures. We suggest that rapid capitalist advancement of metadata and analytics has destabilised earlier evaluative regimes based on fixed indicators. As digital infrastructures evolve, evaluation practices increasingly rely on algorithmic systems that are imbued with the market logics of data appropriation, commodification and surveillance. This promises efficiency and productivity, while simultaneously driving governance into the hands of actors with commercial interests.

Second, we present preliminary findings from qualitative interviews with academic staff, research managers, and evaluative professionals in the UK, alongside open and commercial research information providers, focusing on perceptions and uses of genAI. These findings are compared with the recent GenAI REF report (Watermeyer et al., 2025), and interpreted through our novel theoretical lenses of generativity, assetisation and "thinking infrastructures." This will unfold in the following ways.

We mobilise the concept of *generativity* to conceptualise the distinctive effects of genAI-enabled research information infrastructures. Building on work that highlights the socio-material entanglement of digital infrastructures and evaluation practices (Bowker & Star, 2000; Star, 1999; Kornberger et al., 2019), we argue that genAI extends the generative capacities of digital infrastructures by enabling the continuous production of new indicators, classifications and evaluative categories from existing metadata (Kornberger et al., 2017; Mennicken & Kornberger, 2021). Unlike traditional bibliometrics, which rely on predefined indicators, genAI operates

on large-scale, interlinked datasets that allow classification to emerge from algorithmic processing itself (Mennicken & Espeland, 2019; Krüger, 2020). We investigate whether this shifts evaluation from reactive or performance-based dynamics towards a generative logic, with outcomes that are perhaps not yet anticipated by their human operators.

Drawing on the concept of *assetisation* (Birch & Muniesa, 2020; Birch et al., 2021), we situate the uptake of genAI within a political economy of data and infrastructures. We show how metadata on research, often produced through publicly funded academic labour, are increasingly repurposed as economic assets through analytic capabilities, pricing structures and proprietary tools. This dynamic is reinforced by what has been described as an "institutional data imperative" (Fourcade & Healy, 2017, p.9), whereby organisations are compelled to produce and retain data irrespective of immediate use (Sadowski, 2019). In the UK, this imperative is closely tied to market-oriented governance logics, where universities are encouraged to demonstrate technological leadership and innovation, including in genAI adoption, as part of their competitive positioning within REF and global rankings. We link this to broader historical debates on the capitalist-imperialist nature of both rankings and datafication (Couldry and Mejias, 2019; Nassiri-Ansari and McCoy, 2023; Sadowski, 2019).

The paper conceptualises digital research information infrastructures as *"thinking infrastructures"* that structure attention, shape cognition and distribute agency across human and non-human actors (Kornberger et al., 2019). We examine the role of users (university administrators) and providers (commercial/open platforms) in embedding genAI within infrastructures or resisting their encroachment into academic judgement (Gillespie, 2010; McCoy & Rosenbaum, 2019). By examining how genAI is interpreted and negotiated by actors within the UK landscape, the paper highlights the political nature of infrastructural power at this particularly historical juncture in time.

Empirically, the paper is situated within a wider cross-disciplinary and international research project examining the roles of open and commercial digital research information infrastructures in research evaluation across Canada, France, Germany, and the UK. While the broader project adopts a comparative perspective, this contribution focuses specifically on the UK context, using preliminary interview data to illuminate how new forms of metadata on research and genAI are reshaping evaluation practices and governance debates.

The paper speaks directly to conference themes on scientific autonomy under pressure, commercial infrastructures and dependencies in research evaluation and artificial intelligence and algorithmic evaluation. It contributes to ongoing debates on responsible metrics by extending them towards questions of 'principle-led' data infrastructures and the governance of genAI in research evaluation (Curry et al., 2022). Primarily, we argue that fostering a diverse, trustworthy and reputable scientific ecosystem requires critical engagement with the infrastructural and political-economic conditions under which genAI is deployed, ensuring that innovation in evaluation does not come at the expense of academic freedom, institutional autonomy and democratic accountability.

References

Bowker, G.C. and Star, S.L. (2000) *Sorting things out: classification and its consequences*. Cambridge, MA: MIT Press.

Birch, K., Cochrane, D. and Ward, C. (2021) 'Data as asset? The measurement, governance, and valuation of digital personal data by Big Tech', *Big Data & Society*, 8(1), 20539517211017308.

Birch, K. and Muniesa, F. (eds.) (2020) *Assetization: turning things into assets in technoscientific capitalism*. Cambridge, MA: MIT Press.

Couldry, N. and Mejias, U.A. (2019) 'Data colonialism: rethinking big data's relation to the contemporary subject', *Television & New Media*, 20(4), pp. 336–349.

Curry, S., Gadd, E. and Wilsdon, J. (2022) 'Harnessing the metric tide: indicators, infrastructures and priorities for research assessment'.

Fourcade, M. and Healy, K. (2017) 'Seeing like a market', *Socio-Economic Review*, 15(1), pp. 9–29.

Gillespie, T. (2010) 'The politics of 'platforms'', *New Media & Society*, 12(3), pp. 347–364.

Kornberger, M., Pflueger, D. and Mouritsen, J. (2017) 'Evaluative infrastructures: accounting for platform organization', *Accounting, Organizations and Society*, 60, pp. 79–95.

Kornberger, M., Bowker, G.C., Elyachar, J., Mennicken, A., Miller, P., Nucho, J.R. and Pollock, N. (2019) *Thinking infrastructures*. Bingley: Emerald.

Krüger, A.K. (2020) 'Quantification 2.0? Bibliometric infrastructures in academic evaluation', *Politics and Governance*, 8(2), pp. 58–67.

Mennicken, A. and Espeland, W.N. (2019) 'What's new with numbers? Sociological approaches to the study of quantification', *Annual Review of Sociology*, 45, pp. 223–245.

McCoy, C. and Rosenbaum, H. (2019) 'Uncovering unintended and shadow practices of users of decision support system dashboards in higher education institutions', *Journal of the Association for Information Science and Technology*, 70(4), pp. 370–384.

Mennicken, A. and Kornberger, M. (2021) 'Von Performativität zu Generativität: Bewertung und ihre Folgen im Kontext der Digitalisierung', *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 73(1), pp. 451–478.

Nassiri-Ansari, T. and McCoy, D. (2023) 'World-class universities? Interrogating the biases and coloniality of global university rankings'.

Star, S.L. (1999) 'The ethnography of infrastructure', *American Behavioral Scientist*, 43(3), pp. 377–391.

Sadowski, J. (2019) 'When data is capital: datafication, accumulation, and extraction', *Big Data & Society*, 6(1), 2053951718820549.

Watermeyer, R.P., Phipps, L., Nieves, R.B. and Crick, T. (2025) 'REF-AI: exploring the potential of generative AI for REF2029'. Available at: https://bpb-eu-w2.wpmucdn.com/blogs.bristol.ac.uk/dist/3/1073/files/2025/11/Summary-report-final.pdf (Accessed: 26 January 2026).

HUMAN JUDGEMENT IN THE AGE OF AI: PEER REVIEW PRACTICES AND TRUST CHALLENGES IN THE SSH: INSIGHTS FROM THE UNIVERSITY OF ZADAR, CROATIA

DRAHOMIRA CUPAR, NIKOLINA PEŠA PAVLOVIĆ
University of Zadar, Department of Information Sciences and Technologies, Croatia

Keywords: AI-assisted peer review, Research integrity, Trust in peer review, Peer review, Research evaluation

Introduction

The increasing use of artificial intelligence (AI) tools is transforming the scholarly communication ecosystem, including peer review practices. Peer review, an important component of scholarly communication that ensures quality control, requires further investigation in the new AI environment.

AI-driven language models can help peer reviewers articulate their assessments more efficiently (Calamur & Ghosh, 2024). Peer reviewers report both positive and negative attitudes and experiences with AI tools. Reported benefits include workload reduction and valuable assistance in peer reviewing (Nicholas et al., 2025; Ali & Shaban, 2025; Ebadi et al., 2025). According to a large-scale survey in 2023 (Van Noorden & Perkel, 2023), researchers express significant caution regarding how AI-assisted reviews could affect the reliability and trustworthiness of the peer review process, and human oversight is considered necessary (Hosseini & Horbach, 2023). Several studies have highlighted ethical and academic integrity concerns regarding the use of AI tools (Ebadi et al., 2025), as well as the need for human oversight in AI use during peer review (Ebadi et al., 2025), emphasising that AI must not replace human judgement (Ali and Shaban, 2025; Daoudi, 2025; Calamur and Ghosh, 2024). Mollaki (2024) noted a lack of publishing policies on whether AI-assisted tools should be used by reviewers, and if so, under which conditions. A recent study on journal policies regarding the use of artificial intelligence in psychiatry and mental health, published by Baminiwatta et al. (2025), stated that '*Peer review policies mostly prohibited AI use*'.

Significant guidelines on AI were issued in 2021 by COPE in a discussion paper titled "*AI in Decision Making*," which encouraged other organisations and many journals to take similar action. European Association of Science Editors[1] (EASE) issued "*Recommendations on the Use of AI in Scholarly Communication*," encouraging journals to include AI policies with a section dedicated to peer review. Other organisations and bodies addressing AI in peer review include the STM Association[2] (the International Association of Scientific, Technical and Medical Publishers), the

[1] European Science Editors (EASE): https://ease.org.uk/
[2] International Association of Scientific, Technical and Medical Publishers (STM): https://stm-assoc.org/

International Committee of Medical Journal Editors[3] (ICMJE), and the World Association of Medical Editors[4] (WAME). *Recommendations on Chatbots and Generative Artificial Intelligence in Relation to Scholarly Publications* by WAME stated, "4) *Editors and reviewers should specify, to authors and each other, any use of chatbots in evaluation of the manuscript and generation of reviews and correspondence*." All these recommendations are promoted under the regulations of the European AI Act[5]. All AI usage is closely linked to research integrity policies and is equally relevant to authors and reviewers. A preliminary study presented in 2023 by Zabjan Bogut, Bolkovac, and Krajna at the SDEWES conference in Croatia showed that more than half of the journals indexed in the WoS Core Collection at that time did not have any AI policy in place, although most planned to develop one. All consulted studies indicate a growing need to establish clear guidelines and structured policies for the role of AI in peer review (Ali and Shaban, 2025; Daoudi, 2025; Ebadi et al., 2025; Calamur and Ghosh, 2024). The integration of AI into peer review activities raises concerns about trust, academic integrity, and research credibility. There is a lack of research on peer review practices and trust challenges in the age of AI in the Republic of Croatia.

The purpose of this study is to understand how peer reviewers in the fields of Social Sciences and Humanities at the University of Zadar perceive and use AI tools in their review process, and how these practices shape their views on research integrity and trustworthiness in the review process.

The goal of this study is to gain insights into SSH reviewers' practices and to explore their experiences and attitudes towards peer review, their concerns about ethics and academic integrity, and their expectations regarding the use of AI in scientific communication, specifically in AI-assisted peer review.

Research questions proposed in this study are:

RQ1 What are the peer-review practices of SSH reviewers from the University of Zadar?

RQ2 What are peer reviewers' practices regarding the use of AI tools in peer-reviewing publications?

RQ3 What are the attitudes of SSH peer reviewers towards AI-assisted peer review?

RQ4 How do peer reviewers perceive the ethical and academic integrity implications of using AI tools in the peer-review process?

RQ5 What expectations do reviewers have regarding journal policies, documentation, and guidelines on AI?

Methodology

This study will use a qualitative research design based on semi-structured interviews with ten reviewers from the Social Sciences (N = 5) and Humanities (N = 5) at the University of Zadar, Croatia. Participants were selected through purposive sampling to ensure they had direct experience with peer review processes. As there is no reliable or unique system for recording the number of reviews in public records of researchers, this study used several methods to identify those likely to have substantial experience in peer review. It is assumed that individuals who publish frequently and are regularly promoted in academia also often serve as reviewers. While this assumption may not hold in rare cases, it is based on the academic advancement system in Croatia, which

[3] International Committee of Medical Journal Editor (ICMJ): https://www.icmje.org/
[4] World Association of Medical Editors (WAME): https://www.wame.org/
[5] European AI Act. https://artificialintelligenceact.eu/the-act/

requires scholars to periodically serve as peer reviewers of scholarly publications, which is one of the requirements for academic promotion.
The purposive sampling process involved identifying participants based on data available in CroRIS[6], a publicly available researchers' database, and cross-checking the number of listed publications within the five years from 2021 to 2025 and the year of their advancement(s). All selected participants indicated in their CroRIS profiles that they had advanced in their positions within the past five years, and all are from the Social Sciences or Humanities. All participants are recruited voluntarily and will provide informed consent before the interview. The study will comply with GDPR requirements, ensuring confidentiality and secure storage of research data. The interviews will explore reviewers' roles as authors and their publishing practices, their peer review practices – specifically regarding AI-assisted peer review and their use of AI tools – academic integrity concerns related to AI-assisted peer review, and their needs and expectations regarding journal policies and guidelines on AI. All interviews will be recorded, transcribed, and analysed using thematic analysis to identify patterns and possible differences between SSH reviewers, as well as differences in their publishing and review practices, and in how they conceptualise trust, research ethics, and research integrity in the digital era.

Expected contribution

This study offers new insights into how SSH reviewers at the University of Zadar, Croatia, conceptualise trust, research integrity, and responsible AI use in peer evaluation. To our knowledge, no previous studies in the Croatian academic context have examined AI-assisted peer review or related concerns about research integrity and trust. The results of this study may contribute to drafting guidelines and recommendations based on researchers' expectations, which can assist in preparing AI policies for journals in the Social Sciences and Humanities.

References

Ali, S. and Shaban, M. (2025) 'Nursing academic reviewers' perspectives on AI-assisted peer review: ethical challenges and acceptance', *International Nursing Review*, 72(3), e70100. Available at: https://doi.org/10.1111/inr.70100

Baminiwatta, A., Costa, C., Weerasinghe, D., Arafat, S.M.Y. and Lund, B.D. (2025) 'Policies on using artificial intelligence adopted by journals in psychiatry and mental health', *European Science Editing*, 51, e165365. Available at: https://doi.org/10.3897/ese.2025.e165365

Calamur, H. and Ghosh, R. (2024) 'Adapting peer review for the future: digital disruptions and trust in peer review', *Learned Publishing*, 37(1). Available at: https://doi.org/10.1002/leap.1594

Checco, A., Bracciale, L., Loreti, P., Pinfield, S. and Bianchi, G. (2021) 'AI-assisted peer review', *Humanities and Social Sciences Communications*, 8(1). Available at: https://doi.org/10.1057/s41599-020-00703-8

COPE Council (2021) *COPE discussion document: artificial intelligence (AI) in decision making - English*. Available at: https://doi.org/10.24318/9kvAgrnJ

[6] CroRiS. Croatian Research Information System: https://www.croris.hr/?lang=en

Daoudi, M. (2025) 'Ethical limits and suggestions for improving the use of AI in scientific research, academic publishing, and the peer review process, based on deontological and consequentialist viewpoints', *Discover Education*, 4(1). Available at: https://doi.org/10.1007/s44217-025-00696-z

OBJECTIVE VS. SUBJECTIVE EVALUATION IN THE CONTEXT OF AI: WHAT IS "OBJECTIVITY" IN EVALUATION?

MICHAEL OCHSNER
FORS, Switzerland

Keywords: Epistemology, Objectivity, Theory, Peer review, Metrics

Background

The reform of research assessment has become a pressing concern in contemporary academic discourse. While initiatives like DORA, the Leiden Manifesto, and more recently the Coalition for Advancing Research Assessment (CoARA) may appear relatively new, social sciences and humanities (SSH) scholars have long critiqued mainstream evaluation practices. For decades, SSH scholars have challenged the metric-based evaluation paradigm that privileges English-language journal articles indexed in commercial databases while neglecting societal relevance and the full spectrum of academic work. However, traditional peer evaluation methods are not without limitations. Despite academics' constant engagement in quality assessment through their everyday scholarly practices — reading, teaching, writing — they frequently find it difficult to define research quality explicitly, as illustrated by the often-cited expression "I know it when I see it". In this sense, there is a long-standing and ongoing discussion about objectivity and subjectivity in research evaluation. In this presentation, I will link this discussion to a broader theoretical and epistemological context, as the boundaries of objectivity and subjectivity become more obviously fluid with the advent of artificial intelligence (AI) and AI-informed tools.

Objectivity vs subjectivity

The Enlightenment marked a fundamental shift in European thought. Previously, Christianity provided the dominant framework for interpreting the world. Profound social, economic, and technological transformations gradually shifted this toward rationalism, characterized by a conceptual separation between humans and nature. This transition established science as the primary means of understanding the world and shaped what would become "modernity".

A key conceptual challenge lies in the human-nature separation central to modernist thought that is also fundamental to important science policy concepts, such as "natural" and "human" sciences, "hard" and "soft" sciences or "urban" and "rural" sciences (Dilthey, 1970 /1910; Storer, 1967). The distinction is often linked to methodology: hard, natural science equals quantitative methods while soft/human science equals human interpretation, even though Dilthey's argument was exactly that both were objective, but differently (see, e.g., Davenport, 1981). Similarly, in research evaluation, it is commonly put forward that metric approaches are "objective" while peer review is human interpretation and therefore "subjective" (and, implicitly, invalid or at least biased).

Recent scholarship questions the empirical validity of this distinction between nature and human, suggesting instead that humans exist as integral components of their environments rather than separate from them (e.g., Latour, 2021; Morin, 2021/2015). This perspective recognizes the complex interdependencies between human societies and ecological systems, proposing that treating them as fundamentally separate entities may not adequately reflect observed relationships (Morin, 2021/2015). Understanding these interconnections may prove important for addressing contemporary sustainability challenges – and might be relevant for research evaluation as well when it comes to societal impact evaluation.

With the fast development of Large Language Models, Natural Language Processing and, recently, generative AI, the boundaries between human and nature get even more blurry. What is objectivity in this context? Simply because a machine based on "hard" engineering spits out an interpretation, is it more objective than a human interpretation? Intrigued by the observation that the "better" (from a human perspective) genAI outputs are, the more they become "human" (in the sense of the machine is starting to cheat, invent or insult), I suggest reflecting on the impact of the objectivity vs. subjectivity distinction in evaluation.

A heuristic

Based on Latour's (2012) Modes of Existence, I suggest a heuristic that might help integrating hard and soft sciences and, especially, different perspectives and epistemologies. The approach will especially be fruitful for inter- and transdisciplinary research but also generally for the evaluation of research in context (Ochsner, 2023). It simply states that there are several (Sub-)Modes of Veridiction that define how members of a group identify the "right" from the "wrong", "true" from the "false". For example, law establishes truth through codified norms: actions are lawful if not prohibited by the law, regardless of normative judgments. Politics, on the other hand, operates through representation: decisions are legitimate when relevant parties participate in decision-making. The same process may thus be evaluated differently across systems—an action deemed legally just may appear politically unjust if the represented majority considers it normatively wrong. The same can be right or wrong – depending on the mode of veridiction of the discipline.

We can thus conceptualize distinct sub-modes of veridiction within the scientific endeavor. In science, the mode of veridiction relies on discourse among specialists, where participants engage in rational argumentation to reach consensus on established knowledge. Within this overarching mode, differences emerge in how discussions are structured and conducted. Natural sciences center discussions on empirical findings, typically formalized mathematically; humanities emphasize logical argumentation; arts focus on aesthetic considerations. These variations reveal a fundamental challenge for interdisciplinary research and evaluation: evaluators, panel members, and colleagues operating across disciplines employ different sub-modes of veridiction. While they may address similar questions, they utilize somewhat divergent conceptual frameworks and evaluative criteria, potentially complicating mutual understanding and assessment. In a context where objectivity and subjectivity blur, I argue that the concept of "diplomacy" might help in mitigating issues of verification.

Evaluation and objectivity in the age of AI

While the boundaries between objectivity and subjectivity blur more visibly in the context of AI, i.e., machines applying algorithms produce interpretations (and change easily their "opinions", as these are just probabilities), it also becomes more visible that there is a power involved in the definition of what data flows into training of

algorithms and the forms and nature of interpretations are allowed to be delivered. A very small number of companies decide how we find, and what is delivered, as answers to our questions. AI-supported tools will be used in evaluations. The two processes, loss of objectivity and power accumulation, are both, a high risk and an opportunity. I argue that the opportunity is greater. Objectivity has never existed and algorithmic power has been centralised since the metric turn in evaluation. Both have been kept under the radar but through the advent of generative AI have become impossible to ignore. However, we must seize the opportunity and include reflections on power relationships and right to interpretation systematically into evaluation procedures. The concept of "diplomacy" across subfields of veridictions might prove to become a useful tool to do so.

References

Davenport, E. (1981) 'Updating Wilhelm Dilthey: values and objectivity in literary criticism', *Mosaic: An Interdisciplinary Critical Journal*, 14(4), pp. 89–105. Available at: https://doi.org/10.2307/24780686

Dilthey, W. (1970) *Der Aufbau der geschichtlichen Welt in den Geisteswissenschaften*. Frankfurt am Main: Suhrkamp (original work published 1911).

Latour, B. (2012) *Enquête sur les modes d'existence: une anthropologie des modernes*. Paris: La Découverte.

Latour, B. (2021) *Où suis-je? Leçons du confinement à l'usage des terrestres*. Paris: La Découverte.

Morin, E. (2021) *Penser global: l'homme et son univers*. Paris: Flammarion (original work published 2015).

Ochsner, M. (2023) 'Diversity in modes of veridiction in knowledge generation for society: what STEM can learn from the SSH', *PIAS Proceedings of the Paris Institute for Advanced Study*, 17, article no. 4. Available at: https://doi.org/10.5281/zenodo.8314546

Storer, N.W. (1967) 'The hard sciences and the soft: some sociological observations', *Bulletin of the Medical Library Association*, 55(1), pp. 75–84.

Trowler, P. (2014) 'Academic tribes and territories: the theoretical trajectory', *Österreichische Zeitschrift für Geschichtswissenschaften*, 25(3), pp. 17–26. Available at: https://doi.org/10.25365/oezg-2014-25-3-2

1.2 RESEARCH EVALUATION REFORMS, GOVERNANCE AND SCIENTIFIC AUTONOMY

CONFRONTING THE REFORMS OF ITALY'S UNIVERSITIES. THE EXPERIENCE OF THE RETE DELLE SOCIETÀ SCIENTIFICHE

MARIO PIANTA
Scuola Normale Superiore, Italy

Keywords: University, Research policies, Scientific societies

Since 2023 Italy's university and research system has experienced major changes, including reduction of public financing, wide-ranging institutional reforms and a large expansion of private for-profit online universities. A large number of policy changes have been introduced, addressing the recruitment of precarious researchers, the rules for public competitions for professorships in universities, the evaluation system, the governance process, the funding system, the loose regulations favouring the activities of private for-profit online universities, etc. The fragmented nature of such initiatives has prevented so far a high-level national debate on the issues. Filling a gap in public debate, 144 Scientific societies have developed a critical view of current reforms, with analyses of the current situation and proposals for amendments (Rete delle società scientifiche italiane). This experience in Italy's policy debate is relevant for understanding the changes under way and the future of the country's higher education and research system.

References

Rete delle società scientifiche italiane (n.d.) *Rete delle società scientifiche italiane*. *Scienza in rete*. Available at: https://www.scienzainrete.it/rete-delle-societ%C3%A0-scientifiche-italiane (Accessed: 16 April 2026).

THE IMPACT OF GLOBAL AND EUROPEAN POLITICAL, TECHNOLOGICAL, AND EPISTEMIC TRANSFORMATIONS ON THE BELGIAN FEDERAL SSAH RESEARCH ECOSYSTEM

MARC VANHOLSBEECK
Université Libre de Bruxelles; Belgian Science Policy Office (BELSPO), Belgium
SHARON LECOCQ
KULeuven; Belgian Science Policy Office (BELSPO), Belgium
EVA GERLAND, BRIGITTE DECADT, JUDITH BIERNAUX, NIENKE ROELANTS, JÉRÔME POURBAIX, MADELINE SIMON, TITOUAN TEUNENS
Belgian Science Policy Office (BELSPO), Belgium

Keywords: Research policies, Impact of SSAH, Geopolitics, Artificial Intelligence, Open science, Research evaluation

Introduction and objectives

This conference paper presents a new research project examining the impact of global and European contextual transformations on Belgium's federal research ecosystem. This work in progress cross-analyzes perceptions among BELSPO (Belgian Federal Science Policy Office)[1] policymakers and researchers active in the federal scientific institutes (FSIs) that are managed by BELSPO, with a focus on social sciences, arts, and humanities (SSAH). It employs a mixed-methods approach, combining literature review, focus groups, interviews, and a survey, to address the central research question: How do federal policymakers and researchers perceive the impact of current global and European transformations (political, technological, and epistemic) on federal science policymaking and research activities in the SSAH?

Contextual changes impacting research policies and practices

The following contextual changes at global and European levels are considered, clustered into geopolitical/economic, technological, and epistemic categories:

- Geopolitical and political economy dynamics: International conflicts and global competition intensify pressures for market-oriented and dual-use research, e.g. in EU framework-programme discussions (Draghi, 2023; EASSH, 2025). Research faces heightened focus on security and strategic sovereignty (e.g., concerns over U.S. funding and access to data, China's rapid innovation growth

[1] BELSPO is the federal administration responsible for preparing, implementing, and evaluating science policy at the Belgian federal level, managing federal scientific institutes, Belnet, and various research programs, while providing the government with scientific expertise to inform public decisions in areas such as climate, biodiversity, space, and heritage. The Coordination Directorate plays a key linking and coherence role: it ensures coordination of science policy at federal, interfederal, and international levels, organizes consultation with the Belgian Communities and Regions, monitors European and multilateral programs and initiatives, and aligns Belgian scientific priorities with major international strategic frameworks.

and technology dominance), scientific sanctions (e.g., suspending R&I cooperation with Israel or Russia), geopolitical shifts from decolonization and partnership recalibrations, and threats to academic freedom and autonomy (e.g., U.S. developments or rising illiberal democracies in the EU). At the same time, initiatives like the EU multilateral dialogue on values and principles in research and innovation aim at re-establishing a common framework for global research cooperation (European Commission, 2022).

- Technology environment: Surges in data and computer power drive generative artificial intelligence (AI) and high-performance computing (HPC), alongside AI-Open Science convergence (Giglia & Tammaro, 2025; Hossein et al., 2025). Open Science advances worldwide via FAIR (Findable, Accessible, Interoperable, Reusable) data/metadata standards, DOI attribution, and dedicated repositories (UNESCO, 2023). At the EU level, the EOSC (European Open Science Cloud) fosters a federated, trusted digital environment for FAIR data and open science across disciplines, including SSAH (Taddei & Paolucci, 2025; Valdestilhas et al., 2025)
- Epistemic shifts: Interdisciplinary convergence spans fundamental, strategic, and applied research for epistemic, innovation, and policymaking reasons (OECD, 2025; Vanholsbeeck & Naji, 2025). Data science methods proliferate in SSAH (Davidson et al., 2019), notably computational social sciences and digital humanities, while big data becomes a topic for critical scrutiny in SSAH research (Becerra & Ratovicius, 2022). Generative AI integrates across disciplines, including SSAH (Morris, 2023; Pu et al., 2024), especially via agent-based simulations (Polhill et al., 2023) and LLMs for tasks like thematic analysis (Karjus, 2023). Ethical questions intensify in global Open Science contexts (Leonelli, 2023).

Impact on federal research policies

Focus groups with BELSPO policymakers, including but not limited to the authors of this paper, will identify and explore perceptions of current and anticipated impacts from the abovementioned contextual changes on federal research policies, with a specific attention to the situation of the SSAH. A list of effective transformations in research policy and ongoing transformative initiatives will be compiled.

Drawing on the authors' experience, reflexivity, and initial literature review, the focus groups will examine topics such as (without limitation):

- Competitiveness: the shift toward competitiveness amid EU/national budget constraints (e.g. European Competitiveness Fund's role in future 10th framework programme discussions)
- Research security: (1) strategic integration of research security and science/cultural diplomacy; (2) constraints and opportunities in international cooperation and partnerships (e.g., hosting of threatened researchers in FSIs)
- AI: BELSPO's engagement in infrastructures like AI Antenna or Gigafactories, risking prioritization of industry over academia and STEM over SSAH;
- Open Science: (1) rising tensions between Open Science and knowledge security/dual-use approach (as in current/future EU programmes) (2) challenges and opportunities in Open Science synergies at federal, Belgian and European levels;

- Policies and policy-related initiatives: adaptation of new or revised BELSPO mandates (e.g., in Open Science) and guidelines (e.g., relating to responsible use of AI in science).

Perceived impact on federal SSAH researchers

Semi-structured interviews with SSAH researchers working in FSIs will explore how they perceive the influence of global and European transformations - and their concrete impact on federal policies and initiatives - on key aspects of their research work: research topic diversity and orientation, access/use of scientific information and data, scholarly communication, public and policy-oriented outreach, and research evaluation approaches.

Methodology, analytical framework and research hypotheses

This mixed-methods study combines:

- Literature review of scientific and policy sources (Section 2)
- Focus groups with BELSPO policymakers (Section 3)
- Semi-structured interviews with federal SSAH researchers (Section 4)
- Quantitative survey of federal researchers.
- Qualitative data analysis will use thematic coding within a SWOT (Strengths, Weaknesses, Opportunities, Threats) framework. Interpretation will apply an ecosystemic lens, drawing on information ecology concepts (Stepp et al., 1999), adapted to academic ecosystems (Vanholsbeeck et al., 2024) and more broadly informed by perception theories.

Triangulation across stakeholders (policymakers vs. researchers) will highlight convergences as well as divergences. Survey statistics will validate qualitative patterns.

Given the primarily qualitative design, the following hypotheses are mostly inferential, exploratory, non-exhaustive, and open to refinement:

- H1: Federal SSAH researchers display a rather limited awareness not only of global and European contextual changes but even more so of how these transformations have an impact on federal research policies and initiatives. Their feeling of urgency as well as their engagement and willingness to co-design future policies with BELSPO policymakers remain modest;
- H2: BELSPO policymakers view global and European transformations as driving federal policy toward competitiveness, dual-use research, large-scale AI infrastructures, recalibrated international partnerships, and Open Science-security balances, potentially at SSAH's and societal innovation expense;
- H3: Federal SSAH researchers perceive global/European transformations and federal policies/initiatives as dual-edged: threatening research topic diversity and academic freedom (e.g., within non-dual-use SSAH fields) while creating opportunities in data access and treatment, Diamond Open Access publications and interdisciplinary outreach;
- H4: Both groups anticipate potentially expanded roles for SSAH research in evidence-informed policymaking (e.g., human-centric AI systems, innovation adoption), but more constrained dedicated budgets for SSAH research.

Conclusion and project strengths

This research project tightly links global/EU evolutions to Belgium's federal ecosystem. It bridges in a single case study policymakers and researchers in the SSAH,

including dedicated attention to research communication and evaluation practices. Presentation at RESSH 2026 will share preliminary focus group and exploratory interview results, and engage participants in constructive discussions in order to refine the research design and hypotheses.

References

Becerra, G. and Ratovicius, C. (2022) 'Social sciences and humanities on big data: a bibliometric analysis', *JISTEM – Journal of Information Systems and Technology Management*, 19, e202219011.

Davidson, E., Edwards, R., Jamieson, L. and Weller, S. (2019) 'Big data, qualitative style: a breadth-and-depth method for working with large amounts of secondary qualitative data', *Quality & Quantity*, 53(1), pp. 363–376. Available at: https://doi.org/10.1007/s11135-018-0757-y

Draghi, M. (2024) *The future of European competitiveness: report by Mario Draghi*. Luxembourg: European Commission.

EASSH (2025) *Society in FP10 proposal is for research on society*.

European Commission (2022) *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on the global approach to research and innovation: Europe's strategy for international cooperation in a changing world*. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52021DC0252

Giglia, E. and Tammaro, A.M. (2025) 'Bridging artificial intelligence (AI) and open science: an interview with Elena Giglia', *Digital Library Perspectives*, 41(2), pp. 389–393.

Hosseini, M., Horbach, S.P., Holmes, K. and Ross-Hellauer, T. (2025) 'Open science at the generative AI turn: an exploratory analysis of challenges and opportunities', *Quantitative Science Studies*, 6, pp. 22–45.

Karjus, A. (2023) 'Machine-assisted quantitizing designs: augmenting humanities and social sciences with artificial intelligence', *arXiv*. Available at: https://arxiv.org/abs/2309.14379

Leonelli, S. (2023) *Philosophy of open science*. Cambridge: Cambridge University Press.

Morris, M.R. (2023) 'Scientists' perspectives on the potential for generative AI in their fields', *arXiv*. Available at: https://arxiv.org/abs/2304.01420

OECD (2025) *OECD science, technology and innovation outlook 2025: driving change in a shifting landscape*. Paris: OECD Publishing. Available at: https://doi.org/10.1787/5fe57b90-en

Polhill, J.G., Heppenstall, A., Batty, M., Salt, D., Colasanti, R., Milton, R. and Hare, M. (2023) 'Exascale computing and 'next generation' agent-based modelling', *Review of Artificial Societies and Social Simulation*, 9.

Pu, Z., Shi, C.L., Jeon, C.O., Fu, J., Liu, S.J., Lan, C. and Jia, B. (2024) 'ChatGPT and generative AI are revolutionizing the scientific community: a Janus-faced conundrum', *iMeta*, 3(2), e178.

Stepp, J.R. (1999) 'Prospectus for information ecology', *Journal of Ecological Anthropology*, 3(1), pp. 39–73. Available at: https://doi.org/10.5038/2162-4593.3.1.4

Taddei, L. and Paolucci, M. (2025) 'The rise of social science infrastructures in Europe and Italy', in *Longitudinal data infrastructures in Europe: tools for open science in social science research*. Cham: Springer Nature Switzerland, pp. 69–79.

UNESCO (2023) *Open science outlook 1: status and trends around the world*. Available at: https://doi.org/10.54677/GIIC6829

Valdestilhas, A., Siebes, R. and van Ossenbruggen, J. (2025) 'SSHOC-NL KG: towards a knowledge graph for social sciences and humanities'.

Vanholsbeeck, M., Šinkūnienė, J., Lendák-Kabók, K. and Gekić, H. (2024) 'Information ecosystems in early academic career building: how do researchers in the social sciences and humanities learn the tricks of the trade?', *Discover Education*, 3(1), p. 130.

Vanholsbeeck, M. and Naji, A. (2025) *Improving the integration of the social sciences, arts and humanities into evidence-informed policymaking ecosystems*.

OPENING SCIENCE AND RESEARCH INFORMATION BY LAW?

DEBORAH DE ANGELIS, STEPHEN WYBER
Knowledge Rights 21; International Federation of Library Associations and Institutions

Keywords: Open science, Publishing research, Research output, Copyright law, Cross-border research, Research exceptions, Text and data mining (TDM), Data Act, ERA Act

For over two decades, the Open Access movement has both advocated for – and practically delivered on – the goal of making the published results of research available without financial barriers. Open Science advocacy has taken this further, recognising that full openness – both as a means of driving equity and performance – means not only publishing research findings, but also the full range of outputs generated throughout the research process, including data, methodologies, code and more.

As Open Access and Science have spread, challenges and barriers that previously received little attention have become priorities. First, there is the increasing understanding, however, of the need to ensure that it is not only information FROM research that is open, but also information ABOUT research. This is the focus of the Barcelona Declaration, agreed in April 2024. Without this, the monitoring of open science, the assessment of researchers and institutions, and the effectiveness and transparency of decision-making about science and research remains dependent on the goodwill of private actors and proprietary platforms.

Secondly, there are the challenges associated with inconsistent copyright laws across borders. While the will and technological capabilities to collaborate openly do exist, this can be undermined when divergent national copyright regimes create legal uncertainty, compliance burdens, and, in some cases, structural barriers to international cooperation. Empirical evidence collected by COMMUNIA in the "Nobody puts research in a cage" initiative shows how restrictive copyright frameworks hinder transparency, reproducibility, remote access and data sharing, effectively placing research "in a cage".

This presentation will look at the consequences of these challenges, and explore potential policy responses. Regarding research information, we will discuss the opportunity to include research information in definitions of the data that should be available by law, and in particular the revision of the European Union's Data Act. This is set to incorporate the Open Data Directive, which currently regulates the accessibility of data and other outputs (except publications) from publicly funded research.

Regarding copyright frameworks, we will examine how conflicts between national copyright laws affect cross-border research projects, with particular attention to access, sharing and dissemination of research outputs. We will argue that existing harmonisation tools, including sector-specific exceptions such as text and data mining, remain insufficient to ensure a truly enabling environment for international

research. In response, we will make the case for a general, mandatory research exception at the international level, designed to safeguard the right to research, promote legal certainty, and support open, collaborative, and reproducible science across borders.

Following this, the presentation will reflect on the relative role of (EU-level) legislation on the pursuit of goals such as harmonising copyright rules and improving research assessment through greater openness. This will include reflection on both the strengths and weaknesses of a legislative approach, and what type of coordinated actions may be needed in order fully to achieve desired outcomes.

This will open the way for a specific look at laws, with a focus on the possibilities through the European Research Area Act and revision of the Data Act. It will explore potential language, and encourage feedback on this, as well as reflection on how the research community can support efforts to promote its inclusion. To support action by the field, the likely counterarguments to such proposals will also be discussed, notably the relevance of provisions around the protection of legitimate commercial interests, as already highlighted in the Directive.

The presentation will close with a broader look, both to other legislation on the agenda at the level of the European Union, as well as opportunities to advance progress at the national level. There will also be a reminder of the need to integrate efforts to update legislation with wider steps to reform research assessment and invest in the necessary infrastructure.

The presentation will connect in particular with the conference themes around scientific autonomy (with a particular focus on challenging the dominance of commercial infrastructures, building transparency, and political intervention), diversity (with a focus on enabling more responsible metrics), reforming research assessment (and the contribution of legislation to this), and towards new governance models (looking at broader frameworks).

References

Barcelona Declaration (2024) *Barcelona Declaration on Open Research Information*. Available at: https://barcelona-declaration.org/ (Accessed: 16 April 2026).

Communia (2023) *Nobody puts research in a cage: researchers' perspectives on working with copyright*, edited by T. Nobre, M. Drabczyk, E. Luth, D. De Angelis and M. Bogataj Jančič. Available at: https://communia-association.org/publication/nobody-puts-research-in-a-cage-researchers-perspectives-on-working-with-copyright/ (Accessed: 16 April 2026).

Communia (2025) *New policy paper for a general scientific research exception*. Available at: https://communia-association.org/policy-paper/policy-paper-23-for-a-general-scientific-research-exception/ (Accessed: 16 April 2026).

De Angelis, D. (2024) 'Le eccezioni per scopo di ricerca: il mondo guarda all'UE', in *Crisi e resilienza del diritto d'autore: il recepimento italiano della direttiva 2019/790*. Torino: Giappichelli, pp. 21–38.

De Angelis, D. (2023) 'Steps forward for copyright exceptions and limitations for education and research at the WIPO SCCR 43 agenda', *Infojustice*. Available at: https://infojustice.org/archives/45235 (Accessed: 16 April 2026).

Flynn, S., Schirru, L., Palmedo, M. and Izquierdo, A. (2022) 'Research exceptions in comparative copyright', *PIJIP/TLS Research Paper Series*, no. 75. Available at: https://digitalcommons.wcl.american.edu/research/75

Geiger, C. and Jütte, B.J. (2023) 'Conceptualizing a "right to research" and its implications for copyright law: an international and European perspective', *American University International Law Review*, 38(1), pp. 1–86. Available at: https://ssrn.com/abstract=4414085

Geiger, C. and Jütte, B.J. (2023) 'The right to research as guarantor for sustainability, innovation and justice in EU copyright law', in Pihlajarinne, T., Mähönen, J. and Upreti, P. (eds.) *Intellectual property rights in the post pandemic world: an integrated framework of sustainability, innovation and global justice*. Cheltenham: Edward Elgar Publishing, pp. 138–170. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4140627

Wyber, S. (2026) 'Opening research information by law?'. Available at: https://knowledgerights21.org/news-story/opening-research-information-by-law/ (Accessed: 16 April 2026).

RESEARCH EVALUATION AS ACADEMIC FINANCIALIZATION: HOW BANKING LOGICS RESHAPE THE SOCIAL SCIENCES AND HUMANITIES

NIHAN SOYCAN
Middle East Technical University, Department of Area Studies, Turkey; Ted University, Department of Foreign Language Education, Turkey
ZEHRA TAŞKIN
Hacettepe University, Institute of Informatics, Turkey; Adam Mickiewicz University in Poznań, Scholarly Communication Research Group, Poland

Keywords: Research evaluation, Academic publishing, Financialization, Metrics, Scientific autonomy

Introduction and problem statement

Research evaluation has become one of the most powerful governance mechanisms shaping contemporary academia. Decisions about hiring, promotion, funding, and institutional prestige are increasingly determined through quantitative indicators such as publication counts, citation metrics, and journal-based rankings. These instruments are often presented as neutral tools for assessing quality, yet their widespread use has profoundly altered academic behaviour and knowledge production.

This paper addresses research evaluation as a political-economic regime rather than a technical assessment practice. It argues that current evaluation systems operate according to a logic that closely resembles financialization in capitalist economies. The analysis is grounded in David Harvey's discussion of financialization in The Limits to Capital (Harvey, 2006), where value increasingly detaches from material production and is generated through circulation, credit relations, and self-expanding abstractions. Drawing on this framework, the paper conceptualizes contemporary academic publishing and evaluation as a form of academic financialization.

The argument is developed within the context of an ongoing national research project funded by TUBITAK under the 1001 program (Project No. 223K471). The project investigates how metric-based evaluation systems, publishing infrastructures, and policy reforms interact to produce questionable publishing practices and structural inequalities in research systems in Turkiye.

Conceptual framework: Financialization and academic publishing

In Harvey's account, financialization describes a systemic transformation in which money becomes capable of generating more money through mechanisms that operate independently of productive activity. Credit, debt, and speculative instruments acquire a central role, producing recurring crises that are not anomalies but structural outcomes of the system itself.

This paper argues that a similar dynamic characterizes contemporary academic publishing and evaluation. Scientific publications increasingly function as units of

circulation rather than as indicators of epistemic contribution. Citations operate under specific conditions as signals that sustain metric accumulation, while evaluation scores reproduce themselves by shaping future opportunities for funding, collaboration, and visibility. Metrics thus become self-referential instruments whose primary function is to generate further metrics.

Within this regime, academic publishing is reorganized around circulation value. Commercial infrastructures mediate the relationship between universities and academic labour, translating scholarly output into standardized and comparable indicators. Evaluation shifts away from expert judgment grounded in disciplinary norms toward a system governed by abstract signals that prioritize volume, speed, and visibility.

Empirical grounding and project contribution

The TUBITAK-funded project provides the empirical foundation for this conceptualization. It combines large-scale quantitative analyses of publication data with qualitative investigations of academic motivations and policy documents. The project has produced a comprehensive dataset mapping more than two thousand journals categorized as questionable, alongside a nationally representative sample of academic profiles in Türkiye. These data reveal that questionable publishing practices are not marginal deviations but structurally linked to evaluation systems that reward output quantity and index-based visibility.

Findings from the project demonstrate that publication behaviour is closely aligned with evaluation thresholds in promotion, incentive, and funding policies. Peaks in output frequently coincide with career-critical moments, such as applications for associate professorship or performance-based incentives. These patterns support the claim that evaluation metrics function as a form of credit system, generating pressure to accumulate measurable outputs regardless of their substantive contribution.

From this perspective, phenomena such as paper mills, low-transparency journals, and rapid-publication outlets should be understood as forms of fictitious academic production. They mirror fictitious capital in financial systems by creating measurable outputs that circulate within evaluation regimes without corresponding epistemic value.

Implications for the social sciences and humanities

The consequences of academic financialization are particularly pronounced in the social sciences and humanities. These fields often produce knowledge whose societal relevance emerges over long periods and through indirect forms of engagement. Metric-driven evaluation regimes, however, privilege short-term circulation and standardized outputs, systematically disadvantaging context-sensitive, theoretical, and critical research.

Within global centre-periphery dynamics, this system reinforces existing inequalities. Researchers working outside dominant academic centres face evaluation standards defined elsewhere, while lacking the institutional power to influence them. This produces a structural form of indebtedness, where scholars must continuously convert their work into recognized metric currencies to remain employable and visible.

Conclusion and significance

This paper contributes to current debates on research assessment by proposing academic financialization as a coherent conceptual framework for understanding the persistence and effects of metric-based evaluation systems. Rather than attributing

questionable publishing practices to individual misconduct, it situates them within the structural logic of contemporary research governance.

By linking political economy theory with empirical evidence from a nationally funded research project, the paper advances a context-sensitive critique of research evaluation that is directly relevant to ongoing reform initiatives. In line with the aims of RESSH 2026, it emphasizes the need to rethink evaluation systems in ways that protect scientific autonomy, epistemic diversity, and the social value of knowledge.

Funding acknowledgment

This study was conducted as part of the project titled “Preventing Questionable Publishing Practices through Responsible Research Evaluation Policies that Consider Geopolitical Dynamics”, funded by The Scientific and Technological Research Council of Turkey (TÜBİTAK), grant number 223K471.

References

Harvey, D. (2006) *The limits to capital*. London: Verso Books.

2.1 QUALITY, VALUES AND CONCEPTUAL FRAMEWORKS

SHIELDING TRANSFORMATIVE NICHES: FROM VALUE CAPTURE TO COLLECTIVE CAPABILITIES

EIRÍKUR SMÁRI SIGURÐARSON
University of Iceland, Iceland

Keywords: Research evaluation, Transformative niches, Value capture, Collective capabilities, Academic-societal collaboration

Scientific autonomy is increasingly constrained by geopolitical agendas, digital governance, and AI-driven evaluation regimes. Yet current reform initiatives, such as the Coalition for Advancing Research Assessment (CoARA), call for more nuanced, qualitative, and context-sensitive evaluation that recognises diverse contributions to research and society. This paper makes a specific argument: autonomy is not a retreat from accountability, but a precondition for agency — the capacity of researchers and partners to define what counts as valuable, meaningful, and transformative in their own terms. Without protected space for that capacity, external accountability metrics do not supplement professional judgment; they replace it.

Drawing on the Transformative Innovation Policy framework (Molas-Gallart et al. 2021) and the Capability Approach (Robeyns 2017), the university is conceptualised as a site of transformative niches: spaces where alternative forms of knowledge, collaboration, and societal value can mature. Such niches require shielding — not isolation, but the active maintenance of conditions that sustain collective capabilities such as trust, reflexivity, shared inquiry, and mutual recognition. Without this, research communities risk value capture (Nguyen 2024), in which evaluation metrics overwrite the very judgments they were designed to inform.

CoARA's vision — grounded in qualitative judgment, peer review, and the responsible use of indicators — offers a normative horizon for such shielding. By insisting that evaluation must recognise diverse outputs and practices, CoARA implicitly endorses a conception of agency in which researchers and partners are co-authors of evaluative criteria, not merely objects of measurement. The challenge is to translate this normative vision into institutional practice in contexts where short-term impact metrics still dominate.

The paper grounds this argument in two qualitative datasets collected in Iceland. Ten semi-structured interviews with humanities scholars (archaeology, history, literature, philosophy) examined their understandings of societal impact, evaluation, and autonomy (Sigurðarson & Haraldsdóttir 2024). Three focus groups with representatives from cultural heritage institutions — museums, archives, and heritage organisations — explored their expectations and experiences of collaboration with the university (Sigurðarson & Guðmundsdóttir 2025). Together, these datasets reveal a critical misalignment: scholars' resistance to impact metrics is not a rejection of societal relevance, but a defence of the agency required to cultivate long-term relationships, public conversations, and critical reflection. Cultural-heritage stakeholders, meanwhile, express strong demand for practice-embedded research, yet report structural barriers to collaboration — including limited pathways for involving students and early-career researchers —. A national survey of cultural-heritage

practitioners, currently underway, is designed to test how widespread these patterns are across institutional settings and professional roles; preliminary findings may be available for discussion at the conference.

The paper concludes by advocating "value federalism" as a governance principle: evaluation should operate at multiple levels, allowing local actors to define what counts as valuable while higher-level frameworks set boundaries against inappropriate metrification. Rather than measuring co-creation, institutions should cultivate the conditions that make it possible. Shielding is thus recast as an enabling strategy — by protecting the agency of researchers and partners to act with integrity and imagination, universities preserve their role as laboratories of societal transformation.

References

Molas-Gallart, J., Boni, A., Giachi, S. and Schot, J. (2021) 'A formative approach to the evaluation of transformative innovation policies', *Research Evaluation*, 30(4).

Nguyen, C.T. (2024) 'Value capture', *Journal of Ethics and Social Philosophy*, 27(3).

Robeyns, I. (2017) *Wellbeing, freedom and social justice: the capability approach re-examined*. Cambridge: Open Book Publishers.

Sigurðarson, E.S. and Guðmundsdóttir, S. (2025) *Skýrsla til NSN – hugsað til gagns*. Report.

Sigurðarson, E.S. and Haraldsdóttir, E. (2024) 'Hugsað til gagns: hugvísindi og samfélagsleg áhrif rannsókna', *Skírnir*, 198 (Autumn).

WHAT WE TALK ABOUT WHEN WE TALK ABOUT QUALITY

ALDIS GEDUTIS
Klaipėda University, Lithuania

Keywords: Research quality, Definition(s) of quality, Family resemblances, Informativity, Intuitionism

Raymond Carver's short story "What We Talk About When We Talk About Love" inspired the title of this paper. Carver masterly depicted how seemingly clear and understandable word "love" can be interpreted in very different, even incompatible, ways, by the couple who have been married for a long time: love can be violence-related as well as simply being together or being able to see each other. Because of such a vagueness, talks about quality are not dissimilar to talks about love.

"Quality" language game, to use Ludwig Wittgenstein's term, is based on the unwritten premise, which presupposes relative clarity of the concept. Majority of players of this game take quality for granted and talk of it as if everyone in the game knows what it is without further explanations or definitions. Unfortunately, the clarity is there until somebody asks to define or to explain the notion. In such a case scholars experience St Augustinian sentiment: "If no one asks me, I know what it is. If I wish to explain it to him who asks, I do not know".

Alike love, quality is no less ubiquitous, it seems to be everywhere. Nevertheless, the question "what we talk about when we talk about quality" must be addressed, because "quality", just like "love", means too many different things, and some of them seem to be incommensurable. Digging for a definition of quality, one inevitably encounters proliferation of quality notions. This strongly suggests abandoning the idea of one all-encompassing definition.

The evaluation and quality researchers complicate the things even further. How can we extract something meaningful from this chaotic multitude, when even reputable publications such as the Encyclopedia of Evaluation, instead of helping to understand the mysteries of quality, only confuse matters further? This encyclopedia contains a short entry on quality, in which the author, Bob Williams (2005: 350), not only distinguishes between the technical and emotional connotations of the term quality, but also briefly describes the ambiguous cultural and linguistic uses of the term. Apparently, Americans, Australians, French, Germans, Japanese and New Zealanders see quality in their own cultural ways. Moreover, Robert Pirsig (1999) suggested "quality" being a dynamic phenomenon, which changes not only culturally and geographically but also with time.

Quality master-blender Peter Dahler-Larsen provides "quality" vocabulary, which includes: quality situation, quality script, quality perspective, quality problem, quality recipe, quality inscription, quality object, quality zone, quality infrastructure, quality agents, quality configurations, quality work and qualitization (Dahler-Larsen 2019: 19–20). Other scholars talk about quality sites, i.e. researchers themselves, knowledge communities, research organisations, funding agencies and national policy arenas

(Langfeldt et al. 2020). Yet others sharpen the quality tracking tools, such as quality sites, quality cultures, quality works and responsive quality articulations (Müller 2025). Then, it is, thus, possible to talk about overlapping but distinct quality layers as seen from metaphysical, macro, meso and micro perspectives (Gedutis and Kirtiklis 2026).

Therefore, instead of defining "quality", this paper opts to answer another question: should we define it at all? Is quality worth of defining? What will we gain by providing or by abandoning a definition?

If anyone were to propose such a definition, it would inevitably be either too narrow or too broad:

1. It would be too narrow or reductive, as it would omit a multitude of different indicators and criteria, so such a definition would inevitably be limited to specific situations that would be difficult to universalize and generalize, i.e., to transfer to other language games, contexts, or discourses.
2. It would be too broad or uninformative, as it would not make a proper distinction between what the concept of quality is appropriate for and what it is not. In this way, anything could become an object or subject (of attribution) of quality. The concept of quality can be so broad that it becomes difficult to apply in specific situations and practices where it is important to assess the "quality" of object X or phenomenon Y. So, if the definition of quality were inherently abstract, is there anything to which these descriptions would not apply?

A partial solution that helps avoid the need to precisely define quality is the provision characteristic of ethical intuitionism to declare a certain problematic concept as undefinable but intuitively understandable. However, it must be acknowledged that the effectiveness of such a decision is based on the assumption of unity of intuition, i.e., it is valid only if the majority's intuition about quality is the same or at least compatible. But what if intuitions differ? What characteristics would intuition need to have in order to gain greater interpretative power in such a case? Ultimately, how can we avoid a conflict of interpretations, and if it does arise, how can we find a solution?

Another possibility is not only to define any requirement, but also to abandon the concept itself in contexts where it is disputed and where there is disagreement about it. After all, no one (from a universal point of view) knows exactly what quality is. In this way, it becomes an empty concept, a concept without content or a concept without clear characteristics. Perhaps, instead of discussing it, it would be better to ask what we are evaluating when we say we are evaluating "quality"? After all, depending on the context, instead of quality, we can talk about novelty, originality, craftsmanship, and so on. Which, incidentally, is what is done in certain cases. Moving from the abstractness of "quality" to the concreteness of "novelty", "originality", "craftsmanship" etc., it would suffice to define the latter concepts and use them to evaluate the suitability, adequacy, appropriateness etc. of the research.

In this context Ludwig Wittgenstein might offer a (tentative) solution, which he calls "family resemblances": "for the various resemblances between members of a family: build, features, colour of eyes, gait, temperament, etc. overlap and criss-cross in the same way" (Wittgenstein 1968: 32e). Each person has her own unique set of features, which become even more apparent in an anonymous crowd. The situation changes when that person appears in an album of extended family photos or videos, where the similarities between family members become apparent. Similarly, following Wittgenstein's methodological precept of "seeing what is common" (Wittgenstein

1968: 34e), not unlike criteria and indicators may indicate their belonging to the Quality family, but only when they are considered (all) together, rather than individually.

Because its vagueness, this indirect method of detecting quality may be criticized by apologists for clarity, which Wittgenstein himself seems to anticipate: "But is a blurred concept a concept at all?" (Wittgenstein 1968: 34e). However, instead of giving a precise answer or presenting strong arguments, he merely poses three questions: "— Is an indistinct photograph a picture of a person at all? Is it even always an advantage to replace an indistinct picture by a sharp one? Isn't the indistinct one often exactly what we need?" (Wittgenstein 1968: 34e). Perhaps the vague image of quality actually satisfies everyone concerned with quality? This question is to be answered during the RESSH 2026 conference.

References

Dahler-Larsen, P. (2019) *Quality: from Plato to performance*. Cham: Palgrave Macmillan.

Gedutis, A. and Kirtiklis, K. (2026) *Apie kokybę ir kitus humanitarinių mokslų demonus [Of quality and other demons of the humanities]*. Vilnius: Phi knygos.

Langfeldt, L., Nedeva, M., Sörlin, S. and Thomas, D.A. (2020) 'Co-existing notions of research quality: a framework to study context-specific understandings of good research', *Minerva*, 58, pp. 115–137. Available at: https://doi.org/10.1007/s11024-019-09385-2

Müller, K. (2025) *Quality and the humanities: negotiating research quality in Swedish humanities, 1980s–2020s*. Doctoral thesis. Stockholm: KTH Royal Institute of Technology.

Pirsig, R. (1999) 'Subjects, objects, data and values', in Aerts, D., Broekaert, J. and Mathijs, E. (eds.) *Einstein meets Magritte: an interdisciplinary reflection*. Dordrecht: Springer Science + Business Media, B.V., pp. 79–98.

Williams, B. (2005) 'Quality', in Mathison, S. (ed.) *Encyclopedia of evaluation*. Thousand Oaks, London and New Delhi: Sage Publications, p. 350.

Wittgenstein, L. (1968) *Philosophical investigations*. 3rd edn. New York: Macmillan Publishing Co.

DEVELOPMENT OF A VALUES-BASED APPROACH TO RESEARCH EVALUATION AND IMPACT IN THE ARTS AND HUMANITIES

ANDREW PLUME, SILVIA DOBRE, ALVIN DING, PRAGYA SINGH, NOELLE GRACY
Elsevier, United Kingdom
JINGWEN MU, CHRISTOPHER CHAN, STUART CHRISTIE, AHTI-VIEKKO PIETARINEN
Hong Kong Baptist University, Hong Kong

Keywords: Research evaluation, Qualitative research, Values-based scholarship array, Novel quantitative indicators, Arts and Humanities scholarship

Introduction and motivation

The evaluation of the productivity, reach (or audience), and economic and societal impacts of arts and humanities (AH) scholarship[1] at scale has presented challenges for decades. While the evaluation of individual scholars may incorporate the richness and diversity of AH outputs and outcomes by relying on self-reporting in CVs and contextualised expert assessment, these data do not scale easily at faculty or university levels where quantitative approaches prevail. Databases on which such indicator-informed evaluation rely are traditionally oriented towards comprehensive coverage of STEM fields, contributing to a significant visibility gap for AH scholarship. As a result, such databases and the indicators derived from them (particularly publication-oriented indicators of productivity) may perpetuate reporting biases, iteratively and in aggregate, of a closed STEM-facing system which all too often has served, also, to create an ingrained scepticism for colleagues working across diverse AH disciplines.

At the same time, universities, funders and governments increasingly need to understand, measure, and demonstrate the value of AH scholarship to their stakeholders, even as they presently lack appropriate, field-tested, and community-accepted tools to do so. In the UK context, the higher education & research sector is presently facing a direct, even existential, challenge to demonstrate measurable contributions to graduate training for the workplace, future-ready learning, upskilling for re-entry learners and - ultimately - economic growth for the nation as a whole (Press, 2026). Similar calls globally often place AH disciplines firmly in the line of sight for funding cuts (Gordon & Argue, 2024; Compton, 2025).

Issues in AH evaluation for evaluators and the evaluated: findings from interviews with scholars and administrators

To create a grounding on the evaluation needs and current practices of AH scholars and administrators, we have conducted 17 semi-structured interviews with individuals in 7 different countries (with a focus on Hong Kong and the UK) and across several different AH disciplines. We found substantial evidence for the existence of systematic and persistent issues in AH evaluation that transcend national and disciplinary

1 A note on terminology: In some SSH fields, the term ‘scholarship’ is preferred over the term ‘research’; for brevity we default to the term ‘scholarship’ here.

boundaries but also encourage the development of a conceptual approach to AH evaluation that takes these disciplines' distinctive values (not data availability or integrity) as the starting point.

Overall, we found that the core problem in indicator-informed evaluation is that well-established and community-supported indicators for AH do not exist. Instead, performance indicators are largely imported from STEM disciplines in a 'scientification' of evaluative frameworks (Lewandowska et al., 2023). AH scholars accordingly, and all too routinely, struggle to be recognized for the diverse range, significance, and impacts of the research they actually publish. Almost in lockstep, AH-discipline evaluators express a similar frustration at the paucity of indicators to readily reflect, measure, and summarise value-based impact research at higher-level aggregations. Deepening this AH-STEM divide even further are structural and temporal misalignments in norm of expected output: notably, the pace of research and the duration of impact trajectories. All interviewees noted, moreover, that work in AH disciplines can be of conceptual rather than empirical nature. It follows that timeframes for the fuller incubation of such concepts may be years, or even decades, longer than in STEM; even as the periodicity of AH performance evaluations is typically much shorter (most often aligned, in fact, with STEM-based performance cycles).

Our project thus involved collecting and classifying extensive longitudinal feedback regarding specific activities, outputs and outcomes that AH scholars engage in or produce. Broadly speaking, these included intersecting areas such as knowledge production, service contributions, prestige/influence, reach/audience (especially beyond academia), and socio-economic impact. Interviewees also suggested a variety of novel or underexplored approaches that could support the development of novel, quantitative indicators for a more AH-appropriate evaluative approach. Crucially, interviewees emphasized that more and better indicators should complement, rather than replace, narrative evaluation and peer judgment (Hicks et al., 2015).

Development of a values-based array for AH research evaluation and impact

The underlying theme that emerged across the project interviews concerned fitness for purpose: evaluative systems must reflect what AH scholars actually do, the distinctive forms that their contributions take, the timeframes over which associated outcomes and impact manifest, and the audiences they serve (Oancea, 2024). Rather than opting to create a "catch-all" catalogue of all potential indicators (and their data sources) which might be applicable in AH evaluation contexts, we instead took a narrower, and more selective, values-based approach: what statements do AH scholars wish to make in evaluative contexts that reflect their disciplinary values and audiences and which can be evidenced and supported by an array of quantitative indicators?

To date our tentative, working articulation of such a statement appears as follows: *I am a researcher who advances knowledge and/or practice as an esteemed member of a community through collaboration, collegiality, leadership and education. My impact research is intended for scholars, students, publics, practitioners and policymakers*.

We have accordingly developed a values-based array mapped across a diverse set of quantitative indicators corresponding to the underlined elements of the statement above. Each indicator is presented in the context of the value(s) a scholar pursues and for which the indicator may serve as evidentiary support. Importantly, these indicators are intended to scale to higher-order units and research-inclusive spaces beyond the

individual: such as groups, research teams, departments, schools and faculties, universities and national and international research systems. Backed by specific data findings, we argue that the evaluation of individual AH scholarship is best conducted using such a values-based array engaging robustly with context-sensitive expert assessment.

In this presentation we will explore further the underlying premise of such a values-based array in detail, as well as share current progress on the development of novel quantitative indicators for AH evaluation. Ongoing work in our team aims to gather data and compute such indicators as well as to test their feasibility for and acceptability to both AH scholars and evaluators. It is our goal that such indicators do not supplant prevailing evaluative judgements but rather are embraced as an integral part of a values-driven approach.

References

Compton, E. (2025) 'Making the case for arts and humanities', *Times Higher Education Campus*. Available at: https://www.timeshighereducation.com/campus/making-case-arts-and-humanities (Accessed: 11 March 2026).

Gordon, P. and Argue, M. (2024) 'Government's Marsden Fund cuts: all humanities, social sciences research funding slashed', *RNZ (Radio New Zealand)*. Available at: https://www.rnz.co.nz/news/national/535669/government-s-marsden-fund-cuts-all-humanities-social-sciences-research-funding-slashed (Accessed: 11 March 2026).

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S. and Rafols, I. (2015) 'Bibliometrics: the Leiden manifesto for research metrics', *Nature*, 520, pp. 429–431. Available at: https://doi.org/10.1038/520429a

Lewandowska, K., Kulczycki, E. and Ochsner, M. (2023) 'Evaluation of the arts in performance-based research funding systems: an international perspective', *Research Evaluation*, 32(1), pp. 19–31. Available at: https://doi.org/10.1093/reseval/rvac017

Oancea, A. (2024) 'Hard-to-assess research-impact nexuses in the humanities, arts, and social sciences', in *Handbook of meta-research*. Cheltenham: Edward Elgar Publishing, pp. 55–61. Available at: https://doi.org/10.4337/9781839105722.00012

Press, M. (2026) 'UK universities must evolve to deliver the brighter future people want', *Times Higher Education*. Available at: https://www.timeshighereducation.com/opinion/uk-universities-must-evolve-deliver-brighter-future-people-want (Accessed: 11 March 2026).

2.2A RESEARCH EVALUATION PRACTICES ACROSS INSTITUTIONAL AND PERIPHERAL CONTEXTS

FROM COMMITMENT TO PRACTICE: IMPLEMENTING ARRA COMMITMENT 2 IN CAREER PROGRESSION PROCEDURES AT THE NATIONAL RESEARCH COUNCIL OF ITALY

LOTTIE PROVOST, FRANCESCA DI DONATO
National Research Council - Institute of computational linguistics "Antonio Zampolli" (CNR-ILC), Italy
ANDREA MANNOCCI
National Research Council - Institute of Information Science and Technologies "Alessandro Faedo" (CNR- ISTI), Italy
FABRIZIO PECORARO
National Research Council – Research Institute for Population and Social Policies (CNR-IRPPS), Italy

Keywords: Responsible research assessment, Career progression, Policy implementation, CoARA, Research performing organisation

Background and rationale

The global shift toward responsible research assessment (RRA) reflects growing concerns about the limitations of quantitative, metric-driven evaluation and is embodied in initiatives such as DORA, the Leiden Manifesto, the Coalition for Advancing Research Assessment (CoARA), and the Agreement on Reforming Research Assessment (ARRA) (Arentoft et al., 2022). Setting a shared direction for changes in research assessment practices, the ARRA places particular emphasis on the need to "base research assessment primarily on qualitative evaluation for which peer review is central, supported by the responsible use of quantitative indicators" (ARRA Commitment 2). The overarching aim is to promote a shift towards assessment criteria that prioritise qualitative judgements of research quality, while recognising that quantitative indicators may play a supportive role when used responsibly and contextually.

While the ARRA has garnered substantial international support (923 signatories as of 23 January 2026), the transition from high-level commitments to their operationalisation remains challenging. As institutions move from signing the agreement to implementing its principles, a persistent gap often emerges between institutional policies and their practical application (CoARA Working Group ACA, 2024). Recent empirical studies have explored institutional responses to this reform agenda primarily through large-scale surveys and quantitative analyses of assessment policies and attitudes (CoARA Working Group ACA, 2024; Hyrkkänen et al., 2023; Barnett et al., 2024; Lim et al., 2025). These studies indicate an increasing engagement with ARRA-aligned frameworks at the strategic level, yet also highlight enduring tensions in achieving the intended balance between qualitative judgement and quantitative indicators in assessment practices. Traditional metrics, such as Journal Impact Factor and citation counts, continue to play a dominant role, while the

qualitative recognition of diverse research outputs and activities related to Open Science, societal impact, and science communication remains limited or inconsistently applied (Pontika et al., 2022; Hyrkkänen et al., 2023; CoARA Working Group ACA, 2024).

Although this body of literature provides valuable insights into policy development and institutional intentions, it largely overlooks the critical phase of implementation, where assessment principles are translated into concrete evaluation criteria and procedures.

Case study and methods

Addressing this gap, the study presented here examines the implementation of ARRA principles within career progression procedures at the National Research Council of Italy (CNR). Conducted as a pilot activity within the GraspOS project (GA n. 101095129, January 2023 - December 2025), the study analyses four competitive calls for career progression launched by CNR in 2023, following its signature of the ARRA in 2022. These calls introduced ARRA commitments as a novel evaluation framework and, for the first time in CNR, included the use of Narrative CVs and an explicit emphasis on qualitative merit.

Based on a comprehensive documentary analysis of 4,090 evaluation criteria defined by 90 different evaluation committees, the study examines four competitive career progression calls launched by CNR in 2023 following its signature of the ARRA late 2022. Official call documents, attachments, and committee-defined evaluation criteria were systematically collected from the CNR public portal and analysed using a structured qualitative content analysis conducted through iterative discussion and shared coding among team members. The criteria were coded against a shared analytical matrix developed by the research team to assess compliance with ARRA Commitments 1–4 and the explicit recognition of Open Science practices. In line with Open Science principles, the preprint version of our manuscript is publicly available on Zenodo, together with the study's underlying data and Data Management Plan (Di Donato et al., 2026a; Di Donato et al., 2026b; Provost, 2026).

Within this framework, the present contribution focuses specifically on ARRA Commitment 2, examining how the balance between qualitative evaluation and the responsible use of quantitative indicators is operationalised in practice. Criteria were classified as predominantly qualitative or quantitative based on whether the assessment required substantive expert judgement of the content or relied primarily on mechanical, metric-based scoring schemes and dubious quality proxies. By analysing career progression procedures in relation to institutional strategies and formal policies, this work offers a detailed, practice-oriented assessment of how responsible research assessment principles are enacted within a national research performing organisation.

Findings

The analysis reveals a pattern of partial and uneven implementation of ARRA Commitment 2 across CNR career progression procedures. Three recurring patterns emerge.

- First, compliance with qualitative assessment principles is often selective: while many evaluation committees formally prioritise qualitative judgement, a substantial share continues to rely on quantitative or purely mechanical criteria for specific contributions, frequently applied independently of the intrinsic quality of the assessed product or qualification

- Second, while research outputs are more frequently subject to qualitative judgement, qualifications and other contributions, such as teaching and evaluation, are often assessed through fixed scoring schemes, indicating the persistence of metric-based logics in specific assessment domains. As a matter of fact, variations emerge across evaluation committees. While all committees operated within the same institutional framework, they differed in how they defined and applied evaluation criteria for these contributions, revealing heterogeneous interpretations of qualitative assessment principles across disciplinary clusters.
- Third, although the Narrative CV sections are generally aligned with ARRA Commitment 2, the level of detail and interpretative guidance provided to evaluation committees varies considerably, reflecting heterogeneous understandings of what qualitative assessment entails in practice.

Overall, the findings point to uneven and selective implementation of Commitment 2, highlighting the challenges institutions face in translating qualitative assessment principles into coherent and consistently applied evaluation criteria within career progression procedures. By offering a practice-oriented analysis of these dynamics, the study provides RPOs with a valuable perspective for evaluating their own assessment procedures. This echoes the E of the SCOPE Framework for Research Evaluation (Himanen et al., 2024), which calls on organisations to "Evaluate your Evaluation", encouraging critical reflection on how qualitative assessment is operationalised, the identification of areas for improvement, and the refinement of implementation strategies to better align assessment processes with intended career development outcomes. These findings may be complemented by future research employing alternative analytical and methodological approaches.

References

Arentoft, M., Berghmans, S., Borrell-Damian, L., Bottaro, S., Faure, J.-E., Gaillard, V., Glinos, K., Albacete, J.L., Morais, R., Morris, J., Schiltz, M. and Stroobants, K. (2022) *Agreement on reforming research assessment*. Zenodo. Available at: https://doi.org/10.5281/zenodo.13480728

CoARA Working Group ACA (2024) *Mapping academic career assessment reform initiatives: survey outcomes*. Zenodo. Available at: https://doi.org/10.5281/zenodo.14548157

Di Donato, F., Mannocci, A., Pecoraro, F. and Provost, L. (2026a) *Mind the gap: from policy to practice at CNR. An analysis of career advancement criteria in light of the agreement on reforming research assessment*. Zenodo. Available at: https://doi.org/10.5281/zenodo.18642800

Di Donato, F., Mannocci, A., Pecoraro, F. and Provost, L. (2026b) *Supplemental material to the article "Mind the gap: from policy to practice at CNR. An analysis of career advancement criteria in light of the agreement on reforming research assessment"* (1.0) [data set]. Zenodo. Available at: https://doi.org/10.5281/zenodo.18657087

Himanen, L., Conte, E., Gauffriau, M., Strøm, T., Wolf, B. and Gadd, E. (2024) 'The SCOPE framework: implementing ideals of responsible research assessment', *F1000Research*, 12, p. 1241. Available at: https://doi.org/10.12688/f1000research.140810.2

Hyrkkänen, A.-K., Ivanović, D., Pölönen, J., Kari, M. and Pylvänäinen, E. (2023) *GraspOS deliverable 2.1: OS-aware RRA approaches landscape report*. Zenodo. Available at: https://doi.org/10.5281/zenodo.11098095

Lim, B.H. *et al.* (2025) 'Regional and institutional trends in assessment for academic promotion', *Nature*, 638, pp. 459–468. Available at: https://doi.org/10.1038/s41586-024-08422-9

Pontika, N., Klebel, T., Correia, A., Metzler, H., Knoth, P. and Ross-Hellauer, T. (2022) 'Indicators of research quality, quantity, openness, and responsibility in institutional review, promotion, and tenure policies across seven countries', *Quantitative Science Studies*, 3(4), pp. 888–911. Available at: https://doi.org/10.1162/qss_a_00224

Provost, L. (2026) *Data management plan CNR pilot: career progression at CNR in light of the agreement on reforming research assessment* (1.0). Zenodo. Available at: https://doi.org/10.5281/zenodo.18679355

UK Reproducibility Network, Barnett, J., Clark, R., Darby, R., Desborough, K., Evans, T.R., Farran, E.K., Gowie, E. and Jacobs, N. (2024) *OR4 research assessment survey report: working paper no. 5*. Open Science Framework. Available at: https://doi.org/10.31219/osf.io/z52cn

REFORMING RESEARCH ASSESSMENT WITHIN AN ITALIAN NATIONAL RESEARCH CENTER: HURDLES AND PATH FORWARD

CHIARA MONTAGNA, GIULIANA RUBBIA, MARCO ANZIDEI, ROSARIA TONDI, SANDRO DE VITA, PIETRO BONFANTI, GRUPPO DI LAVORO COARA@INGV
National Institute of Geophysics and Volcanology, Italy

Keywords: CoARA action plan, Responsible scientific career evaluation, Diversity of scientific contributions

INGV, Istituto Nazionale di Geofisica e Vulcanologia, the largest Italian research center devoted to Solid Earth research and monitoring activities, joined the Agreement on Reforming Research Assessment (ARRA; Arentoft et al., 2022) in July 2023, formalizing a commitment to a global reform of the way we, as a research community, value science and how it is conducted. The subsequent development of the CoARA Action Plan (Anzidei et al., 2025) in September 2025 outlines an ambitious roadmap based on the core commitment to recognize the multiplicity and diversity of contributions and careers in geophysical research.

The Action Plan acknowledges that the current evaluation system at INGV exhibits some differences with what the CoARA (Coalition for Advancing Research Assessment, www.coara.org) reform proposes. To advance in the path set forth by the Agreement, the INGV CoARA Action Plan establishes clear objectives, stemming from the CoARA core commitments: i) recognizing the diversity of contributions, explicitly including a wider range of research products in assessment practices, such as software, datasets, peer reviews, research infrastructures, outreach and teaching material, patents, and geological maps; ii) reviewing and developing assessment processes, shifting towards narrative CVs to allow researchers and technologists to actively define their professional paths, and incentivising open science practices; iii) using metrics responsibly, ensuring that all products are evaluated primarily on the basis of qualitative parameters and expert judgment; iv) ensuring transparency and awareness, providing guidelines and organizing training to raise awareness about unconscious biases for hiring and promotion committees.

These clear commitments and the progressive spirit of the Action Plan can help reducing some discrepancies and potential non-conformities with the CoARA principles that have been highlighted by a critical analysis of recent INGV competitive calls for both career advancement and general recruitment. Recent evaluation criteria encourage, for example, the ineffective system of high-impact factor journals (Beigel et al., 2025) and, in some cases, fail to recognize crucial research contributions including e.g. datasets creation and curation. In the context of an institution that has monitoring and dissemination among its core missions, evaluation still lacks sufficient recognition for monitoring networks implementation and management, research infrastructure management, data storage and curation, open source software creation, communication and dissemination activities.

One of the main inconsistencies concerns the ambiguous definition of professional profiles, and the lack of any internal documentation regarding their specificities as well as of a general framework by which researchers shall be evaluated. Detailed criteria and scores are defined by each Committee after the application deadline: candidates apply without knowing the specific weights and criteria their scientific outputs and achievements will be evaluated for. Interpretation of national law that regulates access to public service jobs can indeed allow applicants to know in advance the specific evaluation criteria, as the recent CNR case confirms.

Particularly for the most recent career advancement calls, the scoring mechanism favored individuals who engage in a wide variety of activities over those who specialize in one single area. Widening the spectrum of possible career paths, indeed, shall not lead to homogeneization but instead to the recognition of excellence in different activities. The calls contain no explicit trace of rewarding open science practices, such as early sharing, open collaboration, and publishing in Diamond Open Access journals, which are strongly advocated by CoARA for reproducibility and quality. Finally, the evaluation of scientific outputs is still dominated by publications and heavily based on quantitative/bibliometric criteria (e.g., JCR impact factor and quartiles) in direct contrast with CoARA's advice against relying exclusively on publication venue or format metrics.

INGV CoARA Action Plan establishes a robust framework for reform, to overcome a still fragmented and inconsistent internal research evaluation landscape The current evaluation system is still characterized by a strong dependence on journal-based metrics, excessive variability in the weight given to crucial qualifications and active roles, and a procedural structure that does not explicitly guarantee the adoption of new qualitative and inclusive criteria in the application phase. The lack of a centralized, reformed institutional policy, as evidenced by the significant differences in how individual commissions value the same titles, poses a major challenge to the coherent implementation of the CoARA principles. For the reform to be successful, a crucial next step is to ensure that all new calls explicitly reference the CoARA Agreement and the approved INGV Action Plan, and that the evaluation structure itself is reformed to prioritize qualitative judgment and transparently recognize the full diversity of research contributions.

References

Anzidei, M., Bonfanti, P., Carosi, A., Cecere, G., Chiodetti, A.G., de Vita, S., Ippolito, A., Montagna, C., Piersanti, A., Rubbia, G., Simoncelli, S., Tondi, R. and Liotta, M. (2025) *CoARA action plan – Istituto Nazionale di Geofisica e Vulcanologia*. Zenodo. Available at: https://doi.org/10.5281/zenodo.17130865

Arentoft, M., Berghmans, S., Borrell-Damian, L., Bottaro, S., Faure, J.-E., Gaillard, V., Glinos, K., Albacete, J.L., Morais, R., Morris, J., Schiltz, M. and Stroobants, K. (2022) *Agreement on reforming research assessment*. Zenodo. Available at: https://doi.org/10.5281/zenodo.13480728

Beigel, F. *et al.* (2025) 'The drain of scientific publishing', *arXiv* (version 2, preprint). Available at: https://doi.org/10.48550/ARXIV.2511.04820

THE INTERNATIONALIZATION OF SCIENCE IN LATIN AMERICA: A COMPARATIVE ASSESSMENT AMONG SEVEN UNIVERSITIES OF THE MONTEVIDEO GROUP

AUGUSTO CLEMENTE
Federal University of Paraná, Brazil; Centre for Science and Technology Studies, The Netherlands
ANDRÉ BRASIL
Centre for Science and Technology Studies, Netherlands

Keywords: Internationalization of science, International collaboration, Responsible research assessment, Latin America, Montevideo Group

Introduction

As part of university strategic planning, the Internationalization of Higher Education (IHE) involves integrating international, intercultural, and global dimensions into universities' purposes and functions (Gao, 2019). In Latin America, this topic has sparked an intense debate about the characteristics of the IHE in the region, as well as the challenges affecting scientific cooperation among Latin American universities (Codato et al., 2020).

To examine university internationalization profiles and the intrinsic challenges to this process in Latin America, this study analyzes seven universities belonging to the Montevideo Group - a consortium of approximately 50 Higher Education Institutions (HEIs) aimed at strengthening regional scientific cooperation. The universities in our scope are three Brazilian HEIs: the Federal Universities of Paraná (UFPR), ABC (UFABC), and Santa Maria (UFSM) two Argentinian: the National Universities of Córdoba (UNC) and Rosario (UNR) the University of the Republic (UDELAR) in Uruguay; and the University of Chile (UCHILE).

These universities were selected because they are leading HEIs belonging to AUGM. Also, the comparison among them enables the identification of similarities and differences in their governance towards internationalization in distinct higher educational systems (Brazil, Argentina, Uruguay, and Chile). They combine significant research capacity and measurable international visibility, making them well-suited to examining the tensions between regional integration and global engagement.

Adopting a comparative perspective to evaluate their international standards across research and governance, this paper seeks to generate evidence to support university decision-making by asking:

1. What are the similarities and differences in the internationalization of science among these seven universities?
2. How has the internationalization process unfolded in these institutions?
3. How can international collaboration across Latin American universities be deepened?

Literature Review

Although universities have always had international dimensions and strategies seeking cross-border cooperation, this topic became prominent in the 1990s. Having its basis in globalization theories, the IHE has become a significant concern for universities worldwide (Gao, 2019). Over the past three decades, HEIs have taken into consideration goals, objectives, and strategies to develop an international dimension in science and education (Knight, 2004; 2005; 2013). Internationalization policies have been considered a means to achieving higher positions in global rankings and attracting foreign researchers (Jöns and Hoyler 2013; Olcay and Bulu 2017). The impact of this logic has not gone unnoticed in Latin American HEIs, which are struggling to manage their own internationalization under Global-North parameters (Morosini et al. 2017).

The literature on the IHE in Latin America has compiled data, concepts and analytical propositions that can be summarized in four main hypotheses. First, the decolonial critique of the persistent reliance on English language as a structural constraint on international academic engagement (Wit et al. 2005; Gao 2019). Second, the option to prioritize funding for international academic mobility over investments in "internationalization at home" policies (Dewey and Duff 2009; Wit 2020; Gao 2019). Third, there is stronger collaboration among Spanish-speaking countries than between them and Brazil (Brasil 2018; 2021). Finally, the preference for partnerships with HEIs in Europe and North America rather than with regional counterparts (Marmolejo-Leyva et al. 2015).

Therefore, examining the internationalization strategies and collaboration patterns of Latin American universities sheds light on new issues that can inform solutions more context-sensitive for policymakers in the region.

Developing countries face challenges in internationalizing their science and often "look up to" Global North in order to align their research with global standards. In this regard, our primary goal is to promote responsible evaluation if we wish to foster South-to-South cooperation. As a result, we propose the integration of scientometric methods into university management and decision-making processes. By introducing them, we aim to advance methodological innovations on institutional actors engaged in the IHE in Latin America.

The literature about IHE has expanded its canon, as it has long been limited to academic mobility. This expansion has impacted the dimensions of curriculum, research, governance, community, and culture. Nevertheless, there is not any established theory on university internationalization. The word "internationalization" is used as a theoretical "umbrella term", which would require a more systematic theory of this phenomenon, sensitive to national differences and to developing countries (Gao 2019).

Due to its theoretical challenge, other issues began to arise regarding unintended consequences of internationalization, such as brain race, diploma mills, and the commodification of research and universities (Knight 2013). For instance, extensive efforts have been made to adequately address the brain drain phenomenon. This aspect is not only about attracting and retaining scientists from developing countries, but also about how academic mobility addresses social and economic inequalities (Sugimoto et al. 2017). Furthermore, there are barriers to global collaborative science development, as the concentration of talent in a few countries creates unfair competition (Macháček et al. 2022).

Methodology

Our methodology is based on a mixed-method approach that combines qualitative and scientometric analyzes. Following Noyjons' (2002) recommendation about the validation of the bibliometric data through qualitative interpretation, we conducted in-depth interviews with policymakers and international office representatives from the seven selected universities. This qualitative strategy allowed us to identify shared concerns related to universities' governance, and to assess which indicators are both useful and feasible for them. Furthermore, the qualitative approach provided insights to understand the drivers and rationales underpinning internationalization strategies.

Complementing the qualitative analysis, the scientometric approach offers an overview of the scientific internationalization in the region. Bibliographic data were retrieved from OpenAlex repository for the period from 2020 to 2024, via the CWTS SQL server. The empirical material was analyzed using two sets of statistical techniques: (i) descriptive analyzes, including frequencies and cross-tabulations of variables such as year of publication, type of collaboration, author's institutional affiliations and countries, scientific field, and publication language; and (ii) relational analyzes, based on co-authorship maps of countries and institutions. Tableau and VOSviewer were employed as analytical tools.

Results and Discussion

The preliminary results can be checked in the following links:

1. Tableau dashboard: https://public.tableau.com/views/AUGMInternationalCollaboration_17630504512420/Inter_collabrate?:language=en-US&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link
2. Institutional co-authorship map: https://tinyurl.com/2c7enguh
3. Country co-authorship map: https://tinyurl.com/24es57sy

They indicate a steady increase in international collaboration rates, albeit with differences across countries: Brazilian universities exhibit comparatively low levels and flat growth compared with their Spanish-speaking counterparts. At the same time, Brazilian universities collaborate more nationally, which may be partially explained by Brazil's large territory. Differences across scientific domains highlight these trends: there is a high number of international collaborations in the Physical Sciences, while the Social Sciences show lower levels of international cooperation.

Institutional co-authorship maps reveal a clear division between Brazilian and Spanish-speaking universities, with a more dense cooperation network among the latter. Their shared language might appear as one possible explanation for this proximity. However, such a hypothesis is unlikely, given the dominance of English in scientific communication across all scenarios. Another piece of evidence comes from the country co-authorship map, which reveals the USA as the most central node. Despite the absence of North American HEIs in the dataset, the USA stands out as highly connected to all other countries. This fact indicates an orientation among Latin American researchers toward collaborating with Global North authors.

These preliminary findings seem to support the argument that the Global North has shaped the scientific production in Latin America. This interpretation is reinforced by the content analysis of the interviews carried out with university policymakers. By and large, their internationalization strategies are justified through performance indicators to improve institutional positions in Global university rankings.

Although this paper focuses mainly on the internationalization of science, it is part of a broader project that addresses other international dimensions, such as student and faculty profiles. Furthermore, in addition to academic deliverables, the project includes several outputs developed through a co-creation process: (1) training workshops for staff offered to each participating university; (2) tailored policy briefs and video-abstracts; (3) an interactive dashboard; and (4) a final multi-stakeholder workshop to assess the findings and settle policy recommendations.

References

Brasil, A. (2018) *Exploring differences in scientific production originated from professional and academic graduate courses in Brazil*.

Brasil, A. (2021) 'Beyond the Web of Science: an overview of Brazilian papers indexed by regionally relevant databases', in *Proceedings of the 18th International Conference on Scientometrics and Informetrics (ISSI 2021)*, Leuven, Belgium, pp. 193–204.

Codato, A., Madeira, R. and Bittencourt, M. (2020) 'Political science in Latin America: a scientometric analysis', *Brazilian Political Science Review*, 14. Available at: https://doi.org/10.1590/1981-3821202000030005

Dewey, P. and Duff, S. (2009) 'Reason before passion: faculty views on internationalization in higher education', *Higher Education*, 58(4), pp. 491–504. Available at: https://doi.org/10.1007/s10734-009-9207-z

Gao, C.Y. (2019) *Measuring university internationalization: indicators across national contexts*. Cham: Palgrave Macmillan. Available at: https://doi.org/10.1007/978-3-030-21465-4

Jöns, H. and Hoyler, M. (2013) 'Global geographies of higher education: the perspective of world university rankings', *Geoforum*, 46, pp. 45–59. Available at: https://doi.org/10.1016/j.geoforum.2012.12.014

Knight, J. (2004) 'Internationalization remodeled: definition, approaches, and rationales', *Journal of Studies in International Education*, 8, pp. 5–31. Available at: https://doi.org/10.1177/1028315303260832

Knight, J. (2005) *An internationalization model: responding to new realities and challenges*.

Knight, J. (2013) 'The changing landscape of higher education internationalisation – for better or worse?', *Perspectives: Policy and Practice in Higher Education*, 17(3), pp. 84–90. Available at: https://doi.org/10.1080/13603108.2012.753957

Macháček, V. *et al.* (2022) 'Researchers' institutional mobility: bibliometric evidence on academic inbreeding and internationalization', *Science and Public Policy*, 49(1), pp. 85–97. Available at: https://doi.org/10.1093/scipol/scab064

Marmolejo-Leyva, R., Perez-Angon, M.A. and Russell, J.M. (2015) 'Mobility and international collaboration: case of the Mexican scientific diaspora', *PLOS ONE*, 10(6), e0126720. Available at: https://doi.org/10.1371/journal.pone.0126720

Morosini, M.C., Dalla Corte, M.G. and Guilherme, A. (2017) 'Internationalization of higher education: a perspective from the great south', *Creative Education*, 8(1), article 1. Available at: https://doi.org/10.4236/ce.2017.81008

Noyons, E.C.M., Buter, R.K. and van Raan, A.F.J. (2002) 'Bibliometric mapping as a science policy tool', in *Proceedings of the Sixth International Conference on Information Visualisation*, pp. 679–684. Available at: https://doi.org/10.1109/IV.2002.1028848

Olcay, G.A. and Bulu, M. (2017) 'Is measuring the knowledge creation of universities possible?: a review of university rankings', *Technological Forecasting and Social Change*, 123, pp. 153–160. Available at: https://doi.org/10.1016/j.techfore.2016.03.029

Sugimoto, C.R. *et al.* (2017) 'Scientists have most impact when they're free to move', *Nature*, 550(7674), article 7674. Available at: https://doi.org/10.1038/550029a

Wit, H. de (2020) 'Internationalization of higher education: the need for a more ethical and qualitative approach', *Journal of International Students*, 10(1), article 1. Available at: https://doi.org/10.32674/jis.v10i1.1893

Wit, H. de *et al.* (eds.) (2005) *Higher education in Latin America: the international dimension*. Washington, DC: World Bank. Available at: http://hdl.handle.net/10986/7428

RESEARCH ACTIVITY IN PERIPHERAL UNIVERSITY CONTEXTS: THE CASE OF A SPANISH UNIVERSITY CAMPUS LOCATED IN NORTH AFRICA

OSWALDO LORENZO-QUILES, LUCÍA HERRERA-TORRES
University of Granada, Spain

Keywords: Research evaluation, Peripheral universities, Academic staff perceptions, Social sciences research, Institutional support

This study examines the research activity in peripheral university contexts through the case of the Melilla University Campus, a geographically non-central campus integrated into the University of Granada (Spain). Framed within current debates on research evaluation, the paper explores how standardized bibliometric criteria shape academic practices, motivations, and perceptions in the Social Sciences, particularly in contexts marked by territorial and structural constraints. Using a quantitative, exploratory, cross-sectional design, data were collected through a questionnaire specifically developed and validated for this research, with a sample of 116 Teaching and Research Staff members, representing 54% of the campus population.

The results show that academic staff perceive their research output as moderate to high, although this assessment is consistently tempered by a strong sense of insufficient time for research. Heavy teaching loads, fragmented work schedules, and limited access to competitive funding emerge as key constraints. From a motivational perspective, research is viewed as intellectually meaningful yet demanding, with higher satisfaction reported after completing research processes than during earlier stages, suggesting anticipatory stress linked to evaluation pressures and administrative demands. Dissemination practices are strongly oriented toward publishing in Web of Science and Scopus journals, reflecting dominant evaluation criteria, but potentially marginalizing other socially relevant academic contributions. Overall, the findings underline the need for more contextualized and responsible research evaluation models that acknowledge territorial specificities and diverse forms of knowledge production in peripheral academic environments.

Contemporary research evaluation systems have a decisive influence on academic career trajectories, institutional strategies, and the very dynamics of knowledge production. This impact is particularly visible in the field of Social Sciences, where standardized performance indicators coexist with diverse research practices that are strongly shaped by disciplinary, contextual, and territorial factors (Robinson-Garcia et al., 2023; Waltman et al., 2021). Recent literature has shown that evaluation models predominantly based on bibliometric metrics tend to generate persistent tensions in these fields, affecting both academic motivation and the plurality of forms of knowledge production and transfer (Feenstra & Delgado López-Cózar, 2023; Giménez-Toledo et al., 2024; Hammarfelt & Haddow, 2018).

Within this framework, the present study analyses how the Teaching and Research Staff (TRS) of the Melilla University Campus assess their research activity. This campus is a peripheral unit integrated into a large Spanish public university, the

University of Granada. Interest in this context is justified by the growing attention that international research devotes to territorial inequalities in evaluation systems, as well as to the structural limitations faced by non-central academic environments, which are often rendered invisible in dominant evaluative models (Chams et al., 2023; Yaqub et al., 2023).

The study is based on a quantitative, exploratory, cross-sectional design, implemented through a questionnaire specifically developed for this research and validated by expert judgment. This methodological approach is consistent with recent studies that empirically and contextually examine faculty perceptions of research evaluation in the social sciences (Dewaele et al., 2021). The instrument shows high internal consistency ($\alpha = 0.814$) and makes it possible to examine key dimensions of research activity, such as self-assessment of scientific output, motivation and emotional burden associated with research, satisfaction with research promotion programs, and evaluation of available resources and infrastructure. The sample consists of 116 members of the TRS, representing 54% of the campus's total population.

The results indicate that faculty members rate their research output at moderate to high levels; however, this positive perception is tempered by a widely shared sense of insufficient time devoted to research. This perception appears repeatedly in the international literature as one of the main factors associated with academic dissatisfaction, especially in contexts where pressure to meet external productivity standards is high (Boaz et al., 2021; Robinson-Garcia et al., 2023). In the case analyzed, heavy teaching loads, fragmentation of working time, and difficulties in accessing competitive funding emerge as central explanatory factors.

From a motivational perspective, research is described as an intellectually stimulating activity with professional meaning, though simultaneously demanding. Participants report experiencing higher levels of satisfaction after completing research processes than in the stages prior to their initiation. This finding suggests the presence of anticipatory stress linked to evaluation systems, administrative requirements, and publication pressure, a phenomenon already addressed in recent studies on academic identity and evaluation in the social sciences and humanities (Dewaele et al., 2021; Feenstra & Delgado López-Cózar, 2023).

Furthermore, the analysis of scientific dissemination practices reveals a clear orientation toward publishing articles in journals indexed in Web of Science and Scopus, in line with the criteria that currently dominate the evaluation of research activity. Nevertheless, several studies have warned that this orientation may produce unequal effects across disciplines and limit the recognition of other forms of academic contribution that are particularly relevant in the social sciences, such as social knowledge transfer or interdisciplinary research (Giménez-Toledo et al., 2024; Hicks et al., 2015; Laursen et al., 2023).

Regarding institutional support (Whitley et al., 2018), TRS members express moderate to low levels of satisfaction with the research promotion programs available at different administrative levels. While the existence of calls and support mechanisms is acknowledged, respondents point to access barriers, a considerable bureaucratic burden, and limited adaptation of these policies to the specific characteristics of peripheral campuses. These constraints have been identified in previous studies as factors that reduce the effectiveness of science policies and jeopardize the sustainability of academic careers in non-central contexts (Kampfmann et al., 2022; Chams et al., 2023).

Overall, the findings highlight the need to rethink research evaluation models from a more contextualized perspective that is sensitive to territorial and disciplinary specificities, particularly in the social sciences. Moving toward responsible evaluation approaches that incorporate the real working conditions of teaching and research staff and recognize the diversity of academic contributions constitutes an unavoidable challenge for the design of more equitable and sustainable evaluation systems, as proposed by recent currents in the field of research evaluation (Waltman et al., 2021; Yaqub et al., 2023).

Funding acknowledgment

This study has been funded by the Unit of Excellence of the University Campus of Melilla (University of Granada, Spain). Reference: UCE-PP2024-02.

References

Boaz, A., Borst, R., Kok, M. and O'Shea, A. (2021) 'How far does an emphasis on stakeholder engagement and co-production in research present a threat to academic identity and autonomy? A prospective study across five European countries', *Research Evaluation*, 30(3), pp. 361–369. Available at: https://doi.org/10.1093/reseval/rvab013

Chams, N., Guesmi, B. and Gil, J.M. (2023) 'The nexus between research impact and sustainability assessment: from stakeholders' perspective', *Research Evaluation*, 32(2), pp. 484–495. Available at: https://doi.org/10.1093/reseval/rvad001

De Jong, S., Balaban, C., Holm, J. and Spaapen, J. (2020) 'Redesigning research evaluation practices for the social sciences and humanities: perspectives from the European Network for Research Evaluation in the Social Sciences and Humanities (ENRESSH)', *Social Research*, 73(1). Available at: https://doi.org/10.7220/2335-8769.73.1

Dewaele, A., Vandael, K., Meysman, S. and Buysse, A. (2021) 'Understanding collaborative interactions in relation to research impact in social sciences and humanities: a meta-ethnography', *Research Evaluation*, 30(2), pp. 179–190. Available at: https://doi.org/10.1093/reseval/rvaa033

Feenstra, R.A. and Delgado López-Cózar, E. (2023) 'The footprint of a metrics-based research evaluation system on Spain's philosophical scholarship: an analysis of researchers' perceptions', *Research Evaluation*, 32(1), pp. 32–46. Available at: https://doi.org/10.1093/reseval/rvac020

Giménez-Toledo, E., Olmos-Peñuela, J., Castro-Martínez, E. and Perruchas, F. (2024) 'The forms of societal interaction in the social sciences, humanities and arts: below the tip of the iceberg', *Research Evaluation*, 33, rvad016. Available at: https://doi.org/10.1093/reseval/rvad016

Hammarfelt, B. and Haddow, G. (2018) 'Conflicting measures and values: how humanities scholars in Australia and Sweden use and react to bibliometric indicators', *Higher Education*, 76(6), pp. 1039–1054. Available at: https://doi.org/10.1007/s10734-018-0246-4

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S. and Rafols, I. (2015) 'Bibliometrics: the Leiden manifesto for research metrics', *Scientometrics*, 102(2), pp. 429–447. Available at: https://doi.org/10.1007/s11192-014-1434-2

Kampfmann, T., Bernert, P. and Lang, D.J. (2023) 'Toward a modular evaluation approach of real-world laboratories: findings from a literature review', *Research Evaluation*, 32(1), pp. 128–143. Available at: https://doi.org/10.1093/reseval/rvac029

Laursen, B.K., Motzer, N. and Anderson, K.J. (2023) 'Pathway profiles: learning from five main approaches to assessing interdisciplinarity', *Research Evaluation*, 32(2), pp. 213–227. Available at: https://doi.org/10.1093/reseval/rvac036

Robinson-Garcia, N., Costas, R., Nane, G.F. and van Leeuwen, T.N. (2023) 'Valuation regimes in academia: researchers' attitudes towards their diversity of activities and academic performance', *Research Evaluation*, 32(2), pp. 496–514. Available at: https://doi.org/10.1093/reseval/rvac049

Whitley, R., Gläser, J. and Laudel, G. (2018) 'The impact of changing funding and authority relationships on scientific innovations', *Research Policy*, 47(1), pp. 215–229. Available at: https://doi.org/10.1016/j.respol.2017.11.005

Yaqub, O., Malkov, D. and Siepel, J. (2023) 'How unpredictable is research impact? Evidence from the UK's Research Excellence Framework', *Research Evaluation*, 32(2), pp. 273–285. Available at: https://doi.org/10.1093/reseval/rvad019

BEYOND TRADITIONAL METRICS: KEY VALUE INDICATORS FOR VALUE-BASED RESEARCH AND INNOVATION EVALUATION

MARTA MARTORELL CAMPS, VIVIANA PEREZ CLAUSEN
i2CAT Foundation & Universitat Politècnica de Catalunya, Spain

Keywords: Value-based evaluation, Key value indicators, Public value, Responsible innovation, Research evaluation

Introduction and research problem

Current digital and technological transformations pose major ethical, territorial, and environmental challenges for contemporary societies. Yet both technological innovation processes and public policymaking continue to rely predominantly on performance-based evaluation systems — Key Performance Indicators (KPIs), Objectives and Key Results (OKRs), and related quantitative metrics. While these instruments are useful for measuring efficiency, productivity, and goal attainment (Kaplan & Norton, 1996), they show significant limitations when evaluation must account for public value, social legitimacy, territorial impact, sustainability, or equality. These limitations are particularly acute in contexts characterised by complexity, uncertainty, and contested societal expectations.

The challenge intensifies in the early stages of emerging technologies, where Technology Readiness Levels (TRLs) are low and neither impacts nor success criteria are clearly defined. In such contexts, traditional metrics may distort decision-making and prioritise short-term or easily measurable outcomes over socially meaningful ones (Bozeman, 2002; Moore, 2013). Although expanded frameworks such as Environmental, Social and Governance (ESG) indicators or Societal Readiness Levels (SRL) incorporate social and environmental dimensions (Bronson et al., 2021; Campos & Marín-González, 2023), they often remain ex post and audit-oriented, limiting their capacity to guide innovation processes proactively (Stilgoe et al., 2013). This reveals a persistent gap between performance measurement and the systematic integration of societal values into evaluation and governance.

Theoretical framework

To address this gap, this paper advances the concepts of Key Values (KVs) and Key Value Indicators (KVIs) as a complementary, value-based evaluation framework. KVIs enable analysis not only of what is being done, but also for whom and for what purpose, thereby reframing evaluation as a forward-looking and normative process. Rather than focusing on retrospective performance, KVIs support anticipatory decision-making by making social, ethical, environmental, and public values explicit and operational throughout the innovation lifecycle.

The methodology builds on three complementary theoretical traditions. Public Value Theory shifts evaluation from instrumental efficiency to democratic legitimacy, collective well-being, and societal outcomes (Moore & Khagram, 2004; Moore, 2013; Soh & Martens, 2023). Responsible Research and Innovation (RRI) promotes

anticipatory, reflexive, inclusive, and responsive innovation aligned with societal values (Stilgoe et al., 2013; Fraaije & Flipse, 2019; Fisher et al., 2024; Van de Poel, 2020). Value Sensitive Design (VSD) emphasises the embedding of human values into technological artefacts through iterative processes combining conceptual, empirical, and technical analysis (Friedman et al., 2013). Despite their strong normative foundations, these approaches face well-documented challenges in translating values into operational instruments, particularly in multi-actor environments and low-TRL contexts (Bolz & de Bruin, 2019; Fraaije & Flipse, 2019). Values are frequently addressed ex post, once technologies or policies are already deployed, which limits their capacity as structuring elements of innovation (Ballon et al., 2018; Wikström et al., 2024).

KVIs are introduced as a holistic and adaptive evaluation methodology that operationalises values as key analytical dimensions during design and implementation, rather than as external normative references. Building on Wikström et al. (2024), who propose KVIs for value-based ICT design, this paper addresses a critical gap: the lack of structured methodologies guiding the full KVI development cycle, including value identification, co-design, testing, validation, and iterative refinement in real-world, multi-actor, and context-sensitive environments.

Empirical settings

This contribution draws on two complementary empirical settings. The first is the FORGING project, funded by the European Union, which served as the primary methodological design environment. In FORGING, the KVI framework was developed and refined through co-creation processes and future-oriented scenario design at early stages of technological innovation. KVIs were conceived not as outcome indicators, but as governance tools to articulate, prioritise, and negotiate values under conditions of uncertainty.

The second setting involves the application and ongoing testing of the KVI methodology within a programme for the deployment of Living Labs in Senegal (2025–2027), promoted by the Catalan Agency for Development Cooperation. Building on a previous phase (2023–2025), this programme aims to consolidate Living Labs as infrastructures for social and digital innovation aligned with sustainable development, digital inclusion, gender equity, and community empowerment. Three Living Labs are currently active in Senegal (Kaolack, Dangalma, and Ziguinchor) and are undergoing structural consolidation, including strengthening participatory governance models, capacity building of local teams, and sustainability planning.

The KVI framework is embedded in this programme through processes of co-design, application, and iterative validation. KVIs are used to operationalise community-prioritised values — such as inclusion, sustainability, legitimacy, trust, and gender equity — and to analyse coherence, tensions, and deviations between declared values, design decisions, and emerging impacts during implementation. Supported by formal monitoring mechanisms, this setting provides a robust empirical environment to assess KVIs as anticipatory and adaptive evaluation tools in open innovation ecosystems and Quadruple Helix governance structures.

Methodology

The research follows a qualitative approach with mixed elements, based on a multiple case design and iterative co-design processes with Quadruple Helix actors (government, academia, industry, and civil society). The primary unit of analysis is the KVI methodology itself, which is tested across heterogeneous thematic, territorial, and

institutional contexts. Projects and policies serve as empirical environments for methodological validation rather than as objects of study per se. This approach enables the identification of patterns, tensions, and contextual variations, as well as an assessment of methodological robustness, adaptability, and transferability across different innovation ecosystems.

Contribution and expected outcomes

Moving evaluation beyond traditional metrics supports value-oriented and forward-looking assessment practices. KVIs complement performance-based systems and enhance institutional autonomy and social legitimacy. This paper contributes to the emerging field of responsible research assessment by offering a structured methodology for value integration across the full innovation lifecycle, grounded in multi-actor empirical settings and validated through ongoing iterative processes in both European and Global South contexts.

References

Ballon, P. *et al.* (2018) 'Living labs', in Cammaerts, B. *et al.* (eds.) *Handbook on digital media and communication*. Academic Press.

Bolz, M. and de Bruin, L. (2019) 'Value sensitive design and responsible innovation', *Journal of Responsible Innovation*, 6(3), pp. 329–348.

Bozeman, B. (2002) 'Public-value failure: when efficient markets may not do', *Public Administration Review*, 62(2), pp. 145–161. Available at: https://doi.org/10.1111/0033-3352.00165

Bronson, K., Devkota, R. and Nguyen, V. (2021) 'Moving toward generalizability? A scoping review on measuring the impact of living labs', *Sustainability*, 13(2), p. 502. Available at: https://doi.org/10.3390/su13020502

Campos, I. and Marín-González, E. (2023) 'Renewable energy living labs through the lenses of responsible innovation', *Journal of Responsible Innovation*, 10(1). Available at: https://doi.org/10.1080/23299460.2023.2213145

Fisher, E. *et al.* (2024) 'Responsible innovation scholarship: normative, empirical, theoretical, and engaged', *Journal of Responsible Innovation*, 11(1). Available at: https://doi.org/10.1080/23299460.2024.2309060

Fraaije, A. and Flipse, S.M. (2019) 'Synthesizing an implementation framework for responsible research and innovation', *Journal of Responsible Innovation*, 7(1), pp. 113–137. Available at: https://doi.org/10.1080/23299460.2019.1676685

Friedman, B., Kahn, P.H., Borning, A. and Huldtgren, A. (2013) 'Value sensitive design and information systems', in *Philosophy of engineering and technology*, pp. 55–95. Available at: https://doi.org/10.1007/978-94-007-7844-3_4

Kaplan, R.S. and Norton, D.P. (1996) *The balanced scorecard: translating strategy into action*. Boston: Harvard Business School Press.

Moore, M.H. (2013) *Recognizing public value*. Cambridge, MA: Harvard University Press.

Moore, M.H. and Khagram, S. (2004) *On creating public value: what business might learn from government about strategic management* (Working paper no. 3). Cambridge, MA: Harvard Kennedy School.

Soh, E. and Martens, K. (2023) 'Public values in the socio-technical construction of autonomous vehicle futures', *Public Management Review*, 26(5), pp. 1322–1340. Available at: https://doi.org/10.1080/14719037.2023.2171094

Stilgoe, J., Owen, R. and Macnaghten, P. (2013) 'Developing a framework for responsible innovation', *Research Policy*, 42(9), pp. 1568–1580. Available at: https://doi.org/10.1016/j.respol.2013.05.008

Van de Poel, I. (2020) 'Embedding values in artificial intelligence (AI) systems', *Minds and Machines*, 30(3), pp. 385–409. Available at: https://doi.org/10.1007/s11023-020-09537-4

Wikström, G. *et al.* (2024) 'Key value indicators: a framework for values-driven next-generation ICT solutions', *Telecommunications Policy*, 48(6), 102778. Available at: https://doi.org/10.1016/j.telpol.2024.102778

2.2B CO-CREATION, COLLABORATION AND OPEN BIBLIOMETRIC INFRASTRUCTURE TO ACCOUNT FOR DIVERSITY

JOURNAL REPUTATION AND SCIENTIFIC AUTONOMY IN OPEN ACCESS PUBLISHING

XIAOTING CHEN
School of Information Resources Management, Renmin University of China, China
ISABELLE DORSCH
Sorbonne Université, CNRS, Groupe d'Étude des Méthodes de l'Analyse Sociologique de la Sorbonne, GEMASS, France
ABDELGHANI MADDI
Sorbonne Université, CNRS, Groupe d'Étude des Méthodes de l'Analyse Sociologique de la Sorbonne, GEMASS, France

Keywords: Journal reputation, Open access publishing, Scientific autonomy, Publishing practices, Grey publishers, Research evaluation

Background and problem

Research evaluation systems are increasingly central to the governance of contemporary science. Far from operating as neutral mechanisms, evaluation tools shape publication strategies, funding allocations, and career trajectories, while redefining what counts as legitimate and valuable research. In a context marked by growing geopolitical tensions, concerns about research integrity, and intensified public scrutiny, evaluation has become a powerful instrument through which scientific autonomy is reorganized and new paths of more suitable and sustainable systems are explored.

These dynamics are particularly visible in the transformation of the open access (OA) publishing ecosystem. While OA policies are often framed as promoting openness, transparency, and equity, their implementation also deeply intertwines reputational hierarchies, market concentration, and institutional incentives (Ross-Hellauer, 2022). Journals and publishers compete not only on epistemic quality, but also on perceived prestige, indexation, and compliance with evaluation standards.

In 2020, the Chinese Academy of Sciences (CAS) introduced an "early-warning list" of journals deemed problematic in terms of quality or integrity (Teixeira da Silva et al., 2024). Although not legally binding, this list rapidly became a powerful evaluative signal, widely used by universities and funding agencies to guide publication choices and allocate resources.

This paper examines how the CAS "early-warning list", as a specific reputational signal, is associated with changes in publishing practices within the open access ecosystem, focusing on China's research system.

Conceptual approach

We conceptualize the CAS early-warning list as a form of reputational governance embedded within research evaluation systems. Rather than directly regulating publishing behavior, such instruments operate indirectly by redefining reputational risk and legitimacy. They shape researchers' strategies through anticipation of

sanctions, funding eligibility, and career consequences, thereby constraining individual autonomy while reconfiguring market dynamics.

Building on science policy scholarship, we introduce the notion of a reputation trap to describe situations in which journals or publishers, once stigmatized by reputational signal, experience persistent withdrawal of submissions, even in the absence of formal prohibitions or proven misconduct. In such contexts, lists like the CAS early-warning list may operate as reputational signals that are associated with increased concentration around incumbent publishers, with potential implications for diversity in publishing ecosystems.

Research questions

The paper addresses three interrelated questions:

1. Journal-level effects: How did inclusion in the CAS early-warning list affect publication activity in flagged journals?
2. Publisher-level spillovers: Did reputational penalties extend beyond listed journals to entire publisher portfolios?
3. System-level consequences: How did reputational evaluation reshape the structure and concentration of the OA publishing market?

Data and methods

We analyze a large-scale bibliometric dataset comprising 2,050,088 NSFC-funded journal articles published between 2016 and 2024 by 236,078 Chinese researchers, covering 10,006 journals indexed in OpenAlex. We construct an author–journal–year panel and exploit the 2020 CAS early-warning list as an exogenous reputational shock (Chen et al., 2025).

Using difference-in-differences models with author-by-year fixed effects, we estimate how researchers reallocated their publications across journals following the introduction of the list.

This design isolates within-author shifts in publishing behavior, controlling for time-varying productivity and systemic trends. We further examine heterogeneity across publisher types, distinguishing dominant international publishers, OA-oriented "grey" publishers, and domestic or regional outlets.

Key findings

Our results reveal a sharp and persistent decline in publication activity in journals included on the CAS early-warning list. This effect emerges immediately after 2020 and intensifies over time, with no sign of recovery. Event-study analyses confirm the absence of pre-trends and demonstrate that reputational penalties are durable rather than transitory.

Importantly, reputational effects extend beyond individual journals. OA-oriented and recently established publishers experience substantial spillovers, as non-listed journals within their portfolios also lose submissions. By contrast, established international publishers are largely insulated from reputational damage and absorb displaced publications, reinforcing their dominant market positions.

At the system level, these dynamics lead to a reconsolidation of the publishing market. While OA policies initially facilitated diversification through fast-growing OA publishers, reputational governance reverses this trend by pushing researchers back toward "safe" and prestigious venues. Alternative OA models, such as Diamond or Green OA, remain marginal and fail to absorb the shock.

Implications for research evaluation and scientific autonomy

These findings highlight how research evaluation tools can unintentionally undermine scientific autonomy and pluralism. By amplifying reputational risk, evaluation systems encourage risk-averse publishing strategies and reinforce structural inequalities between publishers. Rather than fostering a diverse and resilient OA ecosystem, reputational governance contributes to market concentration and dependency on incumbent actors. The Chinese case illustrates broader global tensions in research evaluation. As evaluation systems increasingly rely on lists, rankings, and indicators, they reshape scientific behavior in ways that may conflict with policy goals of openness, innovation, and integrity. Understanding these dynamics is essential for designing evaluation frameworks that preserve institutional autonomy, support diverse publishing models, and avoid unintended concentration effects.

References

Chen, X., Dorsch, I. and Maddi, A. (2025) 'Trapped by prestige: systemic vulnerabilities of open access publishing'. Available at: https://doi.org/10.5281/zenodo.17331084

Ross-Hellauer, T. (2022) 'Open science, done wrong, will compound inequities', *Nature*, 603(7901), p. 363. Available at: https://doi.org/10.1038/d41586-022-00724-0

Teixeira da Silva, J.A., Nazarovets, S., Daly, T. and Kendall, G. (2024) 'The Chinese early warning journal list: strengths, weaknesses and solutions in the light of China's global scientific rise', *The Journal of Academic Librarianship*, 50(4), 102898. Available at: https://doi.org/10.1016/j.acalib.2024.102898

CO-CREATING NARRATIVE RESEARCH ASSESSMENT: THE COMON EXPERIMENT AT KYOTO UNIVERSITY

YU SASAKI, ASA NAKANO, HIROSHI HASHIZUME
Office of Research Acceleration, Kyoto University, Japan

Keywords: Responsible research assessment, Narrative assessment, Co-creation, Academic autonomy, Reviewer training

The international debate on research assessment reform has increasingly emphasized the need to move beyond narrow, metric-driven evaluations toward more responsible, context-sensitive approaches. The Research on Research Institute's recent typology of national research assessment systems represents a significant step in systematizing these developments, offering a comparative lens through which to understand continuity and change across countries. At the same time, the typology leaves scope for further engagement with research assessment debates and institutional experiments unfolding in contexts such as Japan, which are not yet reflected in the cases analyzed.

This paper responds to this opportunity by examining Japan's research assessment landscape through the lens of an institutional case study: COMON, a dialogical and narrative-based research assessment framework developed at Kyoto University. It approaches Japan as a context in which multiple forms of research assessment coexist at the national level, creating distinctive opportunities for experimentation and reform at the institutional level.

Kyoto University is internationally recognized for its strong tradition of academic autonomy and long-term scholarly excellence, often referenced by its highest number of Nobel laureates in Asia. Since the corporatization of national universities in Japan in 2003, universities have faced growing accountability demands while striving to preserve conditions for curiosity-driven and fundamental research. In the absence of a single, centralized national research assessment framework, these tensions have been felt in diverse and often subtle ways at the institutional level. Against this backdrop, COMON was developed at Kyoto University as part of a university-wide organizational reform to address these tensions.

COMON is a formative and dialogical research assessment system that emphasizes departmental-level narratives rather than standardized indicators. COMON recognizes five perspectives of research value—Cumulative foundations, Originality, Maturing excellence, Outreach impact, and Network centrality—allowing departments to articulate how research quality and diversity are realized in relation to their intellectual missions. In particular, the perspective of Maturing excellence seeks to make visible forms of fundamental research whose significance may take years or decades to emerge and which are often poorly captured by conventional metrics.

By situating COMON in relation to RoRI's typology, this paper attempts to show how institutional-level initiatives can both complement existing national research assessment paradigms and serve as sites of experimentation that may, over time,

inform the evolution of national-level arrangements—even in the absence of explicit centralized frameworks. It also highlights the practical challenges inherent in narrative and qualitative research assessment, particularly the training and calibration of reviewers capable of evaluating research across disciplinary boundaries.

The development of COMON itself has been shaped by sustained and democratic dialogue among researchers and staff across the university. This process underscores that responsible assessment is not merely a technical design problem but a social and organizational one, requiring trust, shared understanding, and capacity building within academic communities.

In conclusion, this paper contributes to ongoing discussions on how Responsible Research Assessment can be meaningfully designed, implemented, and sustained across diverse research environments by drawing on the experience of institutional experimentation in Japan. The case of COMON at Kyoto University can be understood not only as an institutional response to national policy pressures but also as part of a broader effort to re-articulate the value of academic knowledge in ways that help cultivate the evolving conditions for academic autonomy in contemporary universities. The experience suggests how dialogical and community-based approaches to research assessment can function as building blocks for research environments in which scholarly communities themselves continually articulate and recognize the diverse values of their work, thereby sustaining the conditions in which innovative discoveries and profound intellectual contributions are able to emerge and mature over time. More broadly, the COMON initiative points to the importance of co-creating research assessment practices with scholarly communities and engaging in dialogue with researchers and practitioners developing narrative and qualitative approaches to evaluation across different contexts. In this sense, COMON is not presented as a finished model, but as an invitation to collaboratively explore how research assessment and research environments can be shaped together to sustain dynamic conditions for academic autonomy across institutional and national settings.

References

COMON (n.d.) *COMON: Kyoto University's dialogue-based research assessment system for co-creating new academic commons*. Available at: https://www.research.kyoto-u.ac.jp/en/comon/ (Accessed: 16 April 2026).

Kawano, M. and Poole, G.S. (2021) 'National universities: autonomy in their governance; ideology and practice', in Snowden, P. (ed.) *Handbook of higher education in Japan*. Amsterdam: Amsterdam University Press.

Rushforth, A. and Hammarfelt, B. (2023) 'The rise of responsible metrics as a professional reform movement: a collective action frames account', *Quantitative Science Studies*, 4(4), pp. 879–897. Available at: https://doi.org/10.1162/qss_a_00280

Rushforth, A., Sivertsen, G., Wilsdon, J. *et al.* (2025) *A new typology of national research assessment systems: continuity and change in 13 countries* (RoRI Working Paper no. 15). Research on Research Institute. Available at: https://doi.org/10.6084/m9.figshare.29041787.v4

Sato, I. (2024) 'Performance indicators and policy cycles: on paradoxical longevity of management buzzword in Japan's higher education reforms', *The Doshisha Business Review*, 75(4), pp. 409–429.

THEY COLLABORATE, BUT DO THEY "CO-PRODUCE?" EXAMINING ACADEMIC SCIENTISTS' COLLABORATIVE PATTERNS WITH GOVERNMENT

LUYU DU, JULIA MELKERS
Center for Organization Research and Design (CORD), Arizona State University, United States
NICOLAS ROBINSON-GARCIA
Universidad de Granada, Spain

Keywords: Knowledge co-production, Research evaluation, Societal impact, University-government collaboration, Bibliometrics

Introduction

The funding and conduct of science is aimed at knowledge advancement but also in improving the human condition (Bozeman & Youtie, 2017). Policymakers and the general public expect that new knowledge discoveries derived from scientific and engineering research, particularly publicly-funded research, will diffuse and contribute to society writ large (Sarewitz, 2007). In this context, aligning academic research with societal needs has become an increasingly pressing concern, especially when considering the substantial public funding in science (Crow & Dabars, 2015; Gerber et al., 2023). These issues are not limited to the U.S. context; globally, questions of how to enhance societal benefits for research (de Jong et al., 2016), and how to measure these broader impacts (Bornmann, 2013), are growing in the research policy community. What is less clear, however, is the capacity of academic researchers to be in the position to contribute to broader impacts, in particular to diffuse knowledge to non-academic communities, and how this varies across the scientific community.

While there are a number of mechanisms to accomplish broader impacts, we examine faculty engagement in knowledge co-production with non-academic partners. Given the persistent knowledge-action gap, co-production of knowledge emerges as one of the most important ideas among scholarly discourses (Miller & Wyborn, 2020), which typically involves meaningful collaboration integrating diverse perspectives of academics and non-academics throughout the research process (Yua et al., 2022). In light of increasing pressure to demonstrate societal impacts of research, understanding what drives academics to engage with non-academic partners is critical to improving evaluation processes relevant faculty advancement and other organizational reward systems.

Much of the existing work in academic collaboration with other sectors has focused on university-industry relations (important for commercialization) or public research agencies/laboratories (involving collaboration among scientists in both sectors). Yet, co-production can also involve engagement with a range of nonacademic organizations, both research-oriented and non-research oriented. As such, a variety of innovative research approaches have gained significant visibility in scientific practice,

such as use-inspired basic research and knowledge co-production (Howarth et al., 2022). Regardless of variations in terminologies, these approaches often involve collaboration between academics and non-academics to generate and/or translate scientific knowledge (Nguyen et al., 2020). These collaborative efforts are perceived to be effective means to align research with societal needs, thereby enhancing the use and impact of knowledge (Miller & Wyborn, 2020).

Despite the potential benefits, research collaborations with non-academics, outside of collaboration with industry oftentimes for the purposes of commercialization, have not been widely acknowledged or institutionalized in today's scientific enterprises (Doberneck, 2016; Hart & Silka, 2020). In the evaluation of faculty activities and productivity, these forms of collaboration can often be viewed as outreach and community engagement, rather than other forms of collaboration, specifically knowledge-co-production. In this model, engagement with non-academics can provide meaningful insight to research questions and societal challenges that may in fact shape the research design and trajectory. These efforts are often hindered by conventional disciplinary cultures and academic evaluation and reward systems, which prioritize research excellence over the additional efforts required for external collaborations (Klein & Falk-Krzesinski, 2017). Some critics further view these collaborative practices as a departure from the traditional focus on pure scientific inquiry (Henkel, 2005). This ambivalence reflects the tension between the established academic norms and the growing expectation for academics to collaborate with non-academics. Faculty advancement typically hinges on scholarly production and impacts, which may not fully capture the value of research collaborations with non-academics.

In this paper, we ask: how do faculty engage with non-academics for knowledge co-production?

We are especially interested in collaborations with colleagues in non-traditional organizations – those that do not have a research function or engage in commercialization. By engagement, we focus on the roles that non-traditional collaborators play when working with researchers from research extensive universities. While there is a tremendous diversity of institutions across all sectors, we focus primarily on collaborations with governmental organizations in the United States. Further, because many organizations in the federal government have research units, we focus primarily on collaboration with state and local governments, which typically neither fund nor conduct research. Further, these organization have a strong mission in strengthening societal impacts within their jurisdictions.

Investigating this question will help disentangle the relationships between the academic researchers and public servants, shedding light on the different configurations and intended impacts of this cross-sectoral research. Furthermore, findings will offer valuable insight for university policymakers, identifying practices that could inform the evaluation of faculty research activities, productivity, and intended societal impacts via these knowledge co-production activities.

Data and methods

The data for this study are drawn from two major sources: the 2025 national "State of the Professoriate" survey (NETWISE III) of U.S. academic scientists and engineers, and matched bibliometric data of the survey respondents. The NETWISE III respondent groups includes over 3,000 faculty respondents in nine disciplines from 158 universities in the U.S. A strength of our study is that we have detailed survey and demographic data that is matched with lifetime bibliometric records drawn from the

Scopus Database. For this study, combining survey and bibliometric data creates a unique dataset that allows for in-depth analyses of publication and collaborative behaviors and roles.

To systematically distinguish academic and non-academic collaborators in publication records, we developed a Python-based affiliation classifier for US organizations. The classifier combines a ROR (Research Organization Registry) knowledge base, rule-based heuristics and a local LLM (preliminary findings were produced using Qwen-7B Instruct). It focuses on US affiliations and operates in two stages. First, organizations are classified by sectoral type (supranational organization, government, university, research institute, company, NGO, hospital or other). Then, those categorized under government are further classified into federal, state, local or unknown. In addition, government entities are tagged by the presence or absence of a research-oriented mission.

Preliminary findings

Our analysis involves both descriptive summaries of engagement and co-authorship with non-academic partners, including the variation in those types of organizations across levels of government. We are also conducting a series of regressions (OLS and Logit) to examine the different roles that non-academic collaborators play, as well as the topical areas on which they focus. While these models are still being adjusted (and not shown here), our preliminary results show that there are distinctions in the ways in which academic researchers collaborate with these non-traditional actors.

References

Bornmann, L. (2013) 'What is societal impact of research and how can it be assessed? A literature survey', *Journal of the American Society for Information Science and Technology*, 64(2), pp. 217–233. Available at: https://doi.org/10.1002/asi.22803

Bozeman, B. and Youtie, J. (2017a) 'Socio-economic impacts and public value of government-funded research: lessons from four US National Science Foundation initiatives', *Research Policy*, 46(8), pp. 1387–1398. Available at: https://doi.org/10.1016/j.respol.2017.06.003

Crow, M.M. and Dabars, W.B. (2020) *The fifth wave: the evolution of American higher education*. Baltimore: Johns Hopkins University Press.

Doberneck, D. (2016) 'Are we there yet?: outreach and engagement in the Consortium for Institutional Cooperation promotion and tenure policies', *Journal of Community Engagement and Scholarship*, 9(1). Available at: https://doi.org/10.54656/RNQD4308

Gerber, L.R., Barton, C.J. and Anderson, D.M. (2023) 'Aligning the logics of inquiry and action to address the biodiversity crisis', *Conservation Biology*, 37(5), e14128. Available at: https://doi.org/10.1111/cobi.14128

Hart, D.D. and Silka, L. (2020) 'Rebuilding the ivory tower: a bottom-up experiment in aligning research with societal needs', *Issues in Science and Technology*, 36(3), pp. 64–70.

Henkel, M. (2005) 'Academic identity and autonomy in a changing policy environment', *Higher Education*, 49(1–2), pp. 155–176. Available at: https://doi.org/10.1007/s10734-004-2919-1

Howarth, C., Lane, M., Morse-Jones, S., Brooks, K. and Viner, D. (2022) 'The "co" in co-production of climate action: challenging boundaries within and between science, policy and practice', *Global Environmental Change*, 72, 102445. Available at: https://doi.org/10.1016/j.gloenvcha.2021.102445

Klein, J.T. and Falk-Krzesinski, H.J. (2017) 'Interdisciplinary and collaborative work: framing promotion and tenure practices and policies', *Research Policy*, 46(6), pp. 1055–1061. Available at: https://doi.org/10.1016/j.respol.2017.03.001

Miller, C.A. and Wyborn, C. (2020) 'Co-production in global sustainability: histories and theories', *Environmental Science & Policy*, 113, pp. 88–95. Available at: https://doi.org/10.1016/j.envsci.2018.01.016

Nguyen, T. *et al.* (2020) 'How does integrated knowledge translation (IKT) compare to other collaborative research approaches to generating and translating knowledge? Learning from experts in the field', *Health Research Policy and Systems*, 18(1), p. 35. Available at: https://doi.org/10.1186/s12961-020-0539-6

Sarewitz, D. (2007) 'Does science policy matter?', *Issues in Science and Technology*, 23(4), pp. 31–38. Available at: https://www.jstor.org/stable/43314512

Yua, E., Raymond-Yakoubian, J., Daniel, R.A. and Behe, C. (2022) 'A framework for co-production of knowledge in the context of Arctic research', *Ecology and Society*, 27(1), art. 34. Available at: https://doi.org/10.5751/ES-12960-270134

DISCOVERABILITY AND OPEN ACCESS: LIBRARY STRATEGIES AND IMPACTS ON RESEARCH EVALUATION

BENEDETTA CALONACI
University of Florence, Italy
IRENE PIERGENTILI
Istituto Nazionale di Fisica Nucleare. Laboratori di Frascati, Italy
MARIA ADELAIDE RANCHINO
National Research Council - Planning, Programming and Central Library Unit, Italy

Keywords: Discoverability, Open access, Research evaluation, Academic libraries, Altmetrics

Open-access scientific literature has seen enormous growth over the last five years, supported by huge investment from universities and research institutions. Nevertheless, ensuring that research outputs are openly available is no longer sufficient if they cannot be effectively found, accessed, and evaluated. Discoverability thus represents a key factor that directly influences not only the visibility of research, but also its potential to be assessed, recognized, and reused across different contexts.

The Working Group on Open Access Discoverability was established in 2025 with the aim of investigating how the issue of the discoverability of open access resources is perceived and addressed by two key communities: academic librarians and researchers. Following the approach of a similar pilot-project conducted in the Netherlands (Faniel et al., 2024), the Group adopted a survey-based methodology articulated in three phases.

The first phase consisted of an exploratory survey administered in real time during the Stelline Conference 2025, designed to capture immediate perceptions and professional experiences (Armocida et al., 2025).

The second phase has involved a broader number of Open Access academic librarians, with a more structured survey, which was distributed via email at the national level, with the support of the Italian Library Association (AIB). The results of this survey will be presented in May 2026, during the Stelline Conference.

The third phase (2026-2027), supported by the National Research Council (CNR), will investigate the perception and use of Open Access publication among Italian researchers, with another survey which will target universities of different sizes and research institutions across Italy.

Both surveys investigate several dimensions of OA discoverability, including awareness and use of discovery tools, perceptions of the visibility of OA resources, the role of library infrastructures and services, and the perceived relationship between discoverability and research evaluation processes.

Preliminary findings from the first phase highlight several emerging trends: the importance of integrating selected OA Resources within the electronic collections of

academic libraries, as well as raising users' awareness and understanding of open access; the need for collaboration between researchers, publishers, librarians and service providers to improve interoperability between systems; and most of all, the need for providing OA resources with a comprehensive set of metadata and persistent identifiers.

The ongoing phase and the next one, through surveying librarians and researchers, will enable a more robust and representative overview of institutional strategies for OA discoverability, common challenges, and emerging best practices at a national level.

Building on this perspective, discoverability is conceptualized not merely as a technical feature of information systems, but as a strategic dimension of scholarly communication infrastructures. In a context where Artificial Intelligence and automated discovery tools increasingly influence what is retrieved, read, and ultimately evaluated, discoverability plays a decisive role in shaping the visibility and perceived value of research outputs. It emerges as a critical lever in determining what is found, counted, and valued within contemporary research evaluation systems.

In this context, academic and research libraries are undergoing a profound transformation: they are no longer merely intermediaries for access to resources, but strategic stakeholders in ensuring the quality, visibility, and reuse of scientific research outputs (Henderson, 2017; Tarragó, Oliveira, 2025). Their core activities now include metadata curation, support for Open Access publishing, management of institutional repositories, user training, and the promotion of Open Science practices (Henderson, 2017).

These activities are essential to support the reform of research assessment systems, in line with initiatives such as the Declaration on Research Assessment (DORA) and the Coalition for Advancing Research Assessment (CoARA), which advocate for the adoption of holistic metrics capable of capturing the real impact of scientific production (Inchcoombe, 2025).

In order for search engines and discovery systems to return accurate information, research outputs must be described through high-quality metadata structured according to the FAIR principles (Findable, Accessible, Interoperable, Reusable), thereby ensuring machine-actionability, i.e., the ability of computational systems to autonomously locate and integrate data (Wilkinson et al., 2016). The integration of persistent identifiers (PIDs) - such as DOIs for documents, ORCID iDs for researchers, and ROR identifiers for organizations - together with the use of controlled vocabularies and the application of open licences, constitutes the necessary infrastructure to ensure transparency and reproducibility in scientific research (Knudsen, 2025).

However, the effectiveness of these tools is often undermined by issues related to their use by researchers, such as the proliferation of "orphan records" (i.e., incomplete ORCID profiles or profiles with non-accessible information) and ambiguities arising from homonymy or name variants (El Shamly, Subaveerapandiyan, 2026). Furthermore, the lack of standardization in institutional affiliations leads to fragmentation and distortions in evaluative metrics. It is therefore crucial that institutions provide clear guidelines and enforceable policies on affiliation practices, requiring the use of PIDs. Effective institutional communication ensures that publications are correctly attributed to the organization, thereby enhancing global visibility and guaranteeing the accurate assignment of research outputs.

These requirements acquire even greater importance in light of the growing pervasiveness of Artificial Intelligence (AI) in academic and research environments (Deshmukh, Kadam, 2026). AI-driven systems rely on large-scale information corpora that must be reliable, openly accessible, well-structured, and easily discoverable (Zhu J. (2025). Inadequate metadata, limited accessibility, or fragmented dissemination can therefore affect not only human users, but also the performance, biases, and outputs of automated systems.

In this context, the integration of generative AI into scholarly publishing introduces additional layers of complexity. On the one hand, AI can help streamline workflows, reduce costs, and enhance discoverability, potentially making the Open Access model more sustainable. On the other hand, it raises concerns related to the emergence of proprietary ecosystems and new forms of exclusion, which may limit access to knowledge and exacerbate informational inequalities (Leo S. Lo, 2025; White, 2025). Publishers are therefore faced with strategic choices between open models and models based on advanced services and proprietary content.

Strengthening discoverability is essential to ensure that Open Access can fully support more transparent, inclusive, and responsible research evaluation practices. In this process, academic libraries play a strategic role, not only by developing and managing infrastructures, but also by fostering awareness, supporting researchers, and contributing to the development of new competencies required to navigate an increasingly complex and AI-mediated information ecosystem (Association of College and Research Libraries, 2025).

The issue of discoverability of Open Access resources should therefore be framed within the broader ecosystem of research evaluation. This ecosystem involves a complex interplay of actors (researchers, institutions, funders, policymakers), processes (production, peer review, dissemination, impact measurement), and contexts (cultural, social, economic), all contributing to the assessment of the quality, effectiveness, and value of research. Increasingly, this assessment is not based solely on bibliometric indicators, but also on qualitative dimensions that take into account the societal impact and the broader visibility of scholarly production. These perspectives support more responsible and inclusive approaches to research evaluation (Giglia, E. 2024).

In this scenario, libraries play a key role not only as access infrastructures, but also as active agents in promoting the ethical and responsible use of AI, developing relevant competencies, and contributing to policy-making processes that ensure equity, transparency, and inclusion (Association of College and Research Libraries, 2025).

References

Armocida, D., Boccone, A. and Calonaci, B. (2025) 'La "discoverability" delle risorse open access', *Biblioteche Oggi*, 4. Available at: https://doi.org/10.3302/0392-8586-202504-010-1

Association of College and Research Libraries (2025) *AI competencies for academic library workers*. ACRL. Available at: https://www.ala.org/acrl/standards/ai (Accessed: 16 April 2026).

Baessa, M., Lery, T., Grenz, D. and Vijayakumar, J. (2015) 'Connecting the pieces: using ORCIDs to improve research impact and repositories', *F1000Research*, 4, p. 195. Available at: https://doi.org/10.12688/f1000research.6502.1

Bulock, C. (2021) 'Finding open content in the library is surprisingly hard', *Serials Review*, 47(2), pp. 68–70. Available at: https://doi.org/10.1080/00987913.2021.1936416

Cox, J. (2018) 'Positioning the academic library within the institution: a literature review', *New Review of Academic Librarianship*, 24(3–4), pp. 217–241. Available at: https://doi.org/10.1080/13614533.2018.1466342

Deshmukh, R.K. and Kadam, S. (2026) 'Next-gen repositories: AI-enhanced tools for open access and scholarly communication', *Journal of Advance and Future Research*, 4(1), pp. 744–751. Available at: https://www.rjwave.org/jaafr/papers/JAAFR2601340.pdf

Faniel, I., Brannon, B., Langa, L.A., Doyle, B. and van der Werf, T. (2024) *Improving open access discovery for academic library users*. OCLC Research. Available at: https://doi.org/10.25333/37QC-5845

Farr, U. (2025) 'The state of research assessment: insights from a survey of 6,600+ researchers', *Springer Nature Blog*. Available at: https://www.springernature.com/gp/librarians/the-link/research-management-blogpost/research-assessment-insights-6600-researchers/27769018 (Accessed: 16 April 2026).

Giglia, E. (2024) 'Open science e valutazione: non solo una questione di "alternative"', *Quaderni di Sociologia*, 95(LXVIII), pp. 125–153. Available at: https://doi.org/10.4000/13k0j

Goodier, S. (2024) 'Open science for evaluation use: how applying open science principles can further open up evaluation', *New Directions for Evaluation*, 2024(184), pp. 41–45. Available at: https://doi.org/10.1002/ev.20625

Haak, L.L., Meadows, A. and Brown, J. (2018) 'Using ORCID, DOI, and other open identifiers in research evaluation', *Frontiers in Research Metrics and Analytics*, 3, p. 28. Available at: https://doi.org/10.3389/frma.2018.00028

Henderson, M.E. (2017) *Data management: a practical guide for librarians*. London: Bloomsbury Academic.

Inchcoombe, S. (2025) *The state of research assessment: researcher perspectives on evaluation practices*. Springer Nature. Available at: https://stories.springernature.com/state-of-research-assessment/ (Accessed: 16 April 2026).

Knudsen, M. (2025) *Pure support for research ecosystems through national persistent identifier strategies: interim review and discussion*. Available at: https://dspacecris.eurocris.org/handle/11366/2713 (Accessed: 16 April 2026).

Leo, S.L. (2025) 'Generative AI and open access publishing: a new economic paradigm', *Library Trends*, 73(3), pp. 160–176.

Mi, J. (2024) 'Making open resources discoverable: collaborative approaches for enhanced access', *International Journal of Librarianship*, 8(4), pp. 17–29. Available at: https://doi.org/10.23974/ijol.2024.vol8.4.350

Pianos, T., Moore, H. and Krüger, N. (2025) *Call to action for libraries: improve the discoverability of open textbooks. A step-by-step guide*. Zenodo. Available at: https://doi.org/10.5281/zenodo.15301461

Quarati, A. (2021) 'Open government data: usage trends and metadata quality', *Journal of Information Science*, 49, pp. 1–24. Available at: https://doi.org/10.1177/01655515211027775

Shamly, H.A.H.E. and A, S. (2026) 'Author name disambiguation in scholarly research: a bibliometric perspective', *Open Information Science*, 10(1). Available at: https://doi.org/10.1515/opis-2025-0035

Steltenpohl, C.N., Standiford Reyes, L., Pei, M.A. and Lee, L.E. (2024) 'Open science in qualitative evaluation: considerations and opportunities', *New Directions for Evaluation*, 2024(184), pp. 35–40. Available at: https://doi.org/10.1002/ev.20627

Tarragó, N.S. and Oliveira, T. de A. (2025) 'Librarians, scientific publishing and open science: competencies and challenges in an evolving ecosystem', *RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação*, 23, e025024. Available at: https://doi.org/10.20396/rdbci.v23i00.8680042/en

White, R.R. (2025) 'Generative artificial intelligence tools in journal article preparation: a preliminary catalog of ethical considerations, opportunities, and pitfalls', *JDS Communications*, 6(3), pp. 452–457. Available at: https://doi.org/10.3168/jdsc.2024-0707

Zhu, J. (2025) 'Generative AI for content discovery in academic publishing: a content provider's perspective', *Information Services and Use*, 45(1–2), pp. 48–55. Available at: https://doi.org/10.1177/18758789251336075

FROM THIRD MISSION TO KNOWLEDGE VALORISATION: SOCIETAL IMPACT AND SSH CASE STUDIES IN THE ITALIAN VQR 2020–2024

BRIGIDA BLASI, MARCO MALGARINI, SANDRA ROMAGNOSI
National agency for the evaluation of universities and research (ANVUR), Italy

Keywords: Research evaluation, Knowledge valorisation, Social Sciences and Humanities, Societal impact, Community engagement

In recent years, research assessment systems have increasingly acknowledged the need to move beyond evaluation models based predominantly on scientific outputs and quantitative indicators. These issues are at the core of international debates on responsible research assessment, as reflected in initiatives such as the Coalition for Advancing Research Assessment (CoARA), which calls for evaluation practices that recognise diverse contributions and respect disciplinary specificities.

Within this context, Social Sciences and Humanities (SSH) occupy a particularly critical position, as their societal value often unfolds through cultural, social, educational, and institutional processes that are not easily captured by traditional metrics (Blasi, 2024; Bonaccorsi et al. 2018). The excessive focus of evaluation procedures on commercial and technology transfer dimensions and relations with business have been criticized as representing only a small proportion of activities that have a social and economic impact. Rather informal, flexible, and less visible interactions, the so-called hidden connections (Perkmann et al., 2015) are prevalent, especially in SSH where the economic impact is not granted (Olmos-Peñuela et al., 2014; Hughes et al., 2011; Godin and Doré, 2005). On the other hand, SSH communities have traditionally been less engaged in enhancing engagement in these fields. Ochner et al. (2016) note how cooperation and outward orientation are not always seen as positive elements in these disciplines and are rather seen as a waste of time, taking resources away from research. Finally, several authors have supported the fundamental role of humanities to address and solve some long term global issues, such as aging of society, climate change, or immigration, as well as the impact of new technologies such as Artificial Intelligence (Bonaccorsi, 2026).

The Italian Research Quality Assessement (VQR) 2020–2024 represents a significant development in this evolving landscape. Compared to previous national assessment exercises, the VQR 2020–2024 introduces a conceptual shift from the notion of Third Mission to that of knowledge valorization (valorizzazione delle conoscenze). This transition reflects a broader and more inclusive understanding of how research generates value for society, aligning with international calls and policies to recognise a plurality of research contributions beyond academic excellence narrowly defined.

In the VQR 2020–2024 framework, knowledge valorization is conceived as a multidimensional and process-oriented activity encompassing cultural, social, institutional, and economic dimensions. Rather than framing societal impact as a linear or downstream effect of research, the evaluation model acknowledges the importance of long-term interactions, context of reference, relational dynamics, and

participatory practices involving non-academic actors. This approach resonates strongly with CoARA principles, particularly the emphasis on qualitative judgement, narrative evidence, and context-sensitive evaluation.

A further innovation of the VQR 2020–2024 lies in the thematic organization of knowledge valorization activities. Submitted case studies were classified into predefined thematic areas, such as public engagement, production and management of public goods, and sustainability and Agenda 2030 goals, beyond the areas of technology transfer and life sciences. This thematic framework was designed to enhance comparability while allowing sufficient flexibility to accommodate disciplinary diversity. For SSH research, this approach offers a particularly suitable framework for recognizing heterogeneous pathways to societal impact that do not conform to standardized indicators or linear models of knowledge transfer.

Within this framework, a substantial corpus of knowledge valorization case studies can be clearly attributed to SSH disciplines or to interdisciplinary projects with a strong SSH component. Based on an analysis of titles, primary and secondary thematic areas, and associated keywords, several hundred SSH-related cases were identified. These cases are predominantly distributed across three thematic areas: public engagement, production and management of public and collective goods; sustainability and social inclusion; Agenda 2030 goals. SSH case studies are comparatively less represented in areas focused on technological transfer or commercial exploitation, confirming the distinct nature of SSH knowledge valorization processes.

Across the dataset, SSH case studies tend to articulate impact through qualitative narratives that foreground processes, relationships, and social change rather than measurable outputs. These narratives often highlight sustained engagement with communities, public institutions, schools, cultural organizations, and civil society actors. In line with CoARA recommendations, such narrative-based approaches enable evaluators to capture contributions that would otherwise remain invisible within metric-driven frameworks, particularly those related to co-creation of knowledge and mutual learning.

Beyond the formal thematic categories defined by the VQR, a qualitative analysis of titles and keywords reveals a set of recurrent thematic clusters that cut across disciplinary boundaries and institutional contexts. One of the most prominent clusters concerns public engagement and science–society dialogue. Case studies in this cluster focus on making research accessible and meaningful to broader publics through educational initiatives, participatory communication formats, and interactive cultural activities. These cases reflect a move away from unidirectional dissemination towards dialogical and engaged forms of interaction, in which societal actors actively contribute to shaping the meaning and use of research knowledge.

A second major cluster revolves around social inclusion, communities, and collective well-being. Frequently recurring keywords include inclusion, participation, inequality, accessibility, and community. SSH case studies in this cluster address issues such as education, migration, citizenship, social cohesion, and the empowerment of vulnerable groups. Societal impact is often described in terms of enhanced capacities, increased awareness, and strengthened social networks, underscoring the relational and transformative dimensions of SSH research.

A third cluster centers on cultural heritage, memory, and public goods. SSH research plays a central role in the preservation, interpretation, and valorization of cultural and educational assets, often through collaborations with museums, archives, local administrations, and cultural associations. In these cases, knowledge valorization is

closely linked to the idea of knowledge as a common good, co-produced and sustained through long-term partnerships and shared responsibility. Co-creation emerges as a defining feature, with non-academic actors involved in research design, interpretation, and dissemination.

A fourth, transversal cluster relates to sustainability and societal challenges, frequently framed within the Agenda 2030 framework. SSH case studies in this area address environmental awareness, social responsibility, and policy-relevant knowledge for sustainable development. These projects often adopt

interdisciplinary approaches and engage with public institutions and local communities, contributing to long-term societal objectives rather than immediate or instrumental outcomes.

An exploratory examination of the geographical distribution of SSH case studies suggests that these thematic clusters are broadly homogeneous across the Italian territory. Public engagement, social inclusion, cultural heritage, and sustainability emerge as recurrent themes in all macro-regions, indicating that the VQR framework supports a consistent recognition of SSH knowledge valorization nationwide. At the same time, some context-sensitive variations can be observed. Cultural heritage-related cases appear particularly prominent in areas characterized by strong historical identities, while social inclusion and sustainability-oriented projects often reflect region-specific socio-economic challenges. These differences do not point to rigid geographical specialization but rather illustrate the adaptability of SSH knowledge valorization to local contexts and societal needs.

Overall, the analysis of SSH-related case studies within the VQR 2020–2024 highlights the potential of knowledge valorization as an evaluative concept aligned with international efforts to reform research assessment. By combining thematic classification with narrative-based evaluation and qualitative judgement, the VQR provides an example of how national assessment systems can recognise societal impact, community engagement, and co-creation of knowledge without reducing SSH contributions to simplified metrics.

References

Blasi, B. (2024). 'Social transformation and impact in HASS: Evolution in the debate and current assessment challenges', in Spigarelli, F., Kempton, L. and Compagnucci, L. (eds.), *Entrepreneurship and Digital Humanities*. Edward Elgar Publishing, Cheltenham, pp. 55–74. Available at: https://doi.org/10.4337/9781035331864.00013

Bonaccorsi, A. (2026). *The Knowledge of Humanities*. Brepols. Available at: https://doi.org/10.1484/M.RORAHSS-EB.5.149812

Bonaccorsi, A., Blasi, B. and Romagnosi, S. (2018). 'Do SSH researchers have a third mission (and should they have)?', in Bonaccorsi, A. (ed.), *The Evaluation of Research in Social Sciences and Humanities*. Springer International Publishing, Cham, pp. 361–392. Available at: https://doi.org/10.1007/978-3-319-68554-0_16

CoARA (2022). *Agreement on Reforming Research Assessment*. Zenodo. Available at: https://doi.org/10.5281/zenodo.13480728

Godin, B. and Doré, C. (2005). *Measuring the impacts of science: Beyond the economic dimension*. HIST Lecture. Helsinki Institute for Science and Technology Studies.

Hughes, A., Kitson, M., Probert, J., Turner, R., Bullock, A. and Milner, I. (2011). *Hidden connections: Knowledge exchange between the arts and humanities and the private, public and third sectors*. Centre for Business Research, Cambridge.

Ochsner, M., Hug, S.E. and Daniel, H.D. (2016). 'Humanities scholars' conceptions of research quality', in Ochsner, M., Hug, S.E. and Daniel, H.D. (eds.), *Research Assessment in the Humanities: Towards Criteria and Procedures*. Springer, Cham. Available at: https://doi.org/10.1007/978-3-319-29016-4

Olmos-Peñuela, J., Molas-Gallart, J. and Castro-Martínez, E. (2014). 'Informal collaborations between social sciences and humanities researchers and non-academic partners'. *Science and Public Policy*, 41(4), pp. 493–506. Available at: https://doi.org/10.1093/scipol/sct075

Perkmann, M., Fini, R., Ross, J.M., Salter, A., Silvestri, C. and Tartari, V. (2015). 'Accounting for universities' impact: Using augmented data to measure academic engagement and commercialization by academic scientists'. *Research Evaluation*, 24(4), pp. 380–391. Available at: https://doi.org/10.1093/reseval/rvv020

2.3A MULTILINGUALISM ACROSS DISCIPLINES AND RESEARCH EVALUATION

THE INFLUENCE OF MULTILINGUALISM AND OUTPUT FORMATS ON THE DEVELOPMENT OF RESEARCH TOPICS IN THE SOCIAL SCIENCES AND HUMANITIES

TAKAYUKI HAYASHI
National Graduate Institute for Policy Studies (GRIPS), Japan
HONAMI NUMAJIRI
Kyoto University, Japan
GO YOSHIZAWA, SHOHEI OKAMOTO
Ernst & Young ShinNihon LLC, Japan

Keywords: Multilingualism, Publication formats, Research evaluation, Influence on research topics, Social Sciences and Humanities

Introduction

A central challenge in evaluating social sciences and humanities (SSH) research concerns multilingualism and diverse research output formats. International frameworks like DORA, the Leiden Manifesto, and the Helsinki Initiative warn against overreliance on bibliometric indicators favoring English-language journal articles. While these concerns are established in Europe, they have only recently gained importance in Japan (Science Council of Japan, 2021, 2025), a non-English-speaking research system with strong domestic scholarly tradition.

Historically, research evaluation has exerted limited influence in Japan. Since 2004, a part of national university evaluations have primarily relied on qualitative review of representative research outputs by discipline, rather than on quantitative indicators, and their linkage to funding allocation has remained marginal (typically below 1%). In recent years, however, Japan's declining global rankings in academic publication volume, as measured by Scopus and Web of Science, have prompted the introduction of competitive funding schemes at the institutional level that require the specification of key performance indicators (KPIs). Indicators such as the share of Top 10% Highly Cited Papers have become increasingly common, generating strong pressure on SSH researchers to publish in English-language international journals.

This pressure cannot be understood merely as a problem of measurement bias. A growing body of literature demonstrates that evaluation indicators actively shape researcher behavior and organizational strategies (Whitley & Gläser, 2007; Rijcke et al., 2016). In SSH fields, where research topics are closely tied to language, culture, and social context, evaluation regimes that over-privilege English-language journal articles risk redirecting research agendas toward problem framings dominant in the Anglophone world. As a result, research addressing domestically salient historical, social, and policy issues may be systematically under-produced.

Despite this risk, SSH scholars in Japan have struggled to provide empirical evidence demonstrating why non-English outputs and books matter for knowledge production.

While choices of publication media — particularly books — are often explained in terms of audience reach, visibility, disciplinary conventions, and career-related incentives (Williams et al., 2018), these accounts rarely clarify whether and how research content itself differs across languages and output formats. Differences in publication practices have often been asserted normatively, without showing how distinct research themes depend on specific languages and output formats. This study addresses that gap by empirically examining how research topics vary across languages and publication media.

Data and methods: Challenges in data construction under the structural features of publishing in Japanese SSH

In Japan, no comprehensive and systematically curated database of Japanese-language "academic books" exists. The boundary between academic and general books is unclear, with extensive "hybrid publishing" (Williams et al., 2018; Honingh et al., 2019)—that is, books written simultaneously for scholarly audiences and the general public—and the frequent absence of formal peer review for academically oriented books, making it difficult to define scholarly books institutionally. In addition, books often lack abstracts and citation metadata, which constrains quantitative analysis.

Similarly, there is no exhaustive database for Japanese-language journal articles. Publishing platforms such as J-STAGE provide partial coverage, while the article database CiNii—whose records are also incorporated into OpenAlex—includes a mixture of academic and non-academic materials and lacks consistent disciplinary classifications. These infrastructural limitations further reduce the visibility of SSH research produced in Japanese.

This study analyzes three output types produced by Japanese researchers: Japanese-language academic books, Japanese-language journal articles, and English-language journal articles authored by researchers affiliated with Japanese institutions. The analysis focuses on three SSH fields: history, sociology, and management studies.

Book data were constructed using the "BOOK database" provided by Nichigai Associates, a Japanese bibliographic information service. SSH academic books published between 2019 and 2024 were extracted through a stepwise procedure based on disciplinary classification (NDC10), author affiliation with universities, and content filtering, resulting in a final dataset of 39,673 books. Japanese-language journal articles were identified by compiling discipline-specific journal lists from academic society directories and J-STAGE, and then retrieving articles from CiNii. English-language journal articles were retrieved from Scopus using relevant ASJC codes and Japanese institutional affiliations.

Textual data consisted of book summaries and article titles and abstracts. English-language articles were machine-translated into Japanese to ensure comparability. Sentence embeddings were generated using a Japanese Sentence-BERT model, followed by K-means clustering and UMAP-based two-dimensional visualization. This approach enabled comparison of thematic distributions across output types and languages, as well as analysis of their relative composition within topic clusters.

Results

Figure 1 shows that while research topics in history overlap across Japanese-language books (blue), Japanese-language journal articles (green), and English-language journal articles(orange), certain thematic domains are disproportionately associated with specific output types. This indicates that some topics are shared across media, whereas others are systematically skewed by publication format and language.

However, comparing multiple databases simultaneously introduces methodological challenges, as interdisciplinary research—particularly at the interface between the humanities and the natural sciences, such as agricultural history, which draws on both historical analysis and agricultural or environmental sciences—is strongly shaped by the disciplinary classification structures embedded in each database.

To address this issue, Figure 2 introduces a two-dimensional mapping in which topics are positioned by their distance from the disciplinary centroid (central versus peripheral) and by the relative concentration of Japanese-language books. Under this framework, the upper-left region identifies topics that are central to the discipline yet predominantly expressed through Japanese-language books.

In history, Japanese-language books are strongly associated with some central disciplinary themes requiring extended narrative and contextualization, including ancient traditions, Japan's medieval era of civil warfare, and the Edo period and its late phase. These topics also attract broad public interest and are poorly suited to short journal-article formats. Peripheral yet book-centered topics include socially salient themes such as life and the self, gender, public finance, military bases, and cultural figures. These topics are often located at the boundary between the humanities and social sciences and tend to be articulated through narrative, context-rich forms of scholarship. By contrast, article-dominated clusters—both Japanese and English—tend to focus on theory-driven, comparative, or empirically structured topics such as philosophy, foreign history, diplomacy, and scientific thought.

Similar patterns emerge in sociology (Fig.3) and management studies (Fig.4). Japanese-language books concentrate on socially embedded and practice-oriented knowledge, including care, discrimination, community, business practice, and professional skills. By contrast, English-language journal articles emphasize abstract, theoretical, and quantitative research aligned with international publication standards, focusing on themes such as institutional analysis, causal inference, evaluation, labor and employment, innovation, investment, and accounting frameworks. Japanese-language journal articles occupy an intermediate position, presenting applied research rooted in domestic contexts within conventional academic formats.

Discussion

These findings demonstrate that language and publication media are not neutral containers of research outputs; rather, they are constitutive of research agendas themselves. When evaluation systems privilege particular output forms, they influence not only individual publication strategies but also the thematic distribution and long-term development of entire disciplines. This study thus provides empirical support for concerns that evaluation regimes can reshape disciplinary centres and peripheries in knowledge production.

Finally, the analysis highlights structural weaknesses in Japan's research information infrastructure. The lack of comprehensive and well-classified databases for books and Japanese-language journals makes SSH research less visible and more difficult to assess on its own terms. Addressing these infrastructural gaps is essential for developing evaluation frameworks that recognize linguistic diversity, epistemic plurality, and scientific autonomy.

Figure 1. Topic Distribution across Languages and Publication Formats in History

Figure 1 Topic Distribution across Languages and Publication Formats in History

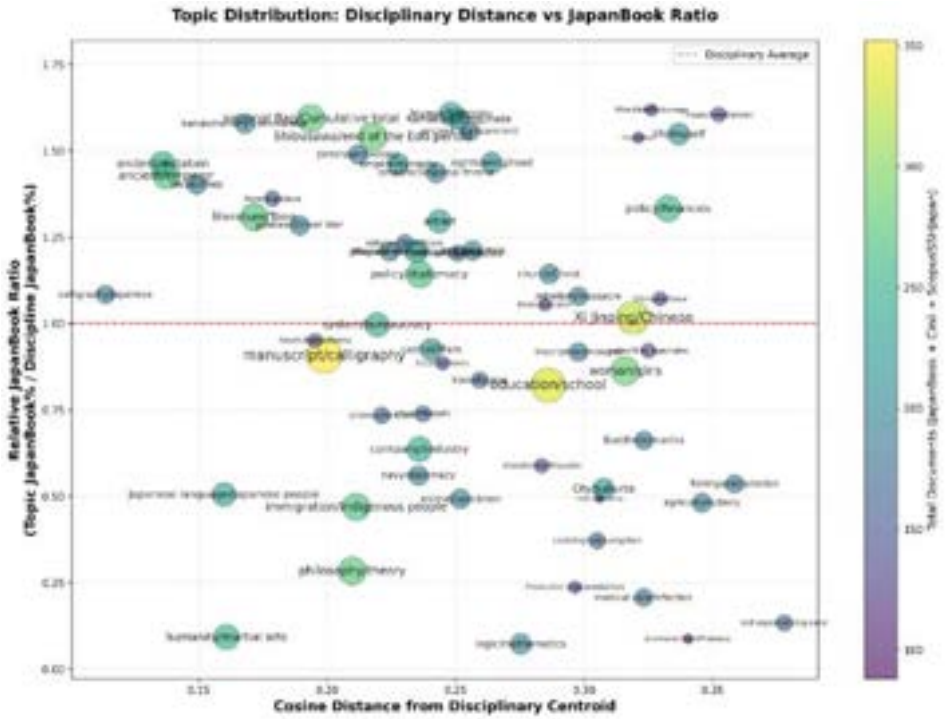

Figure 2 Research Topics in History by Disciplinary Centrality and Book Orientation

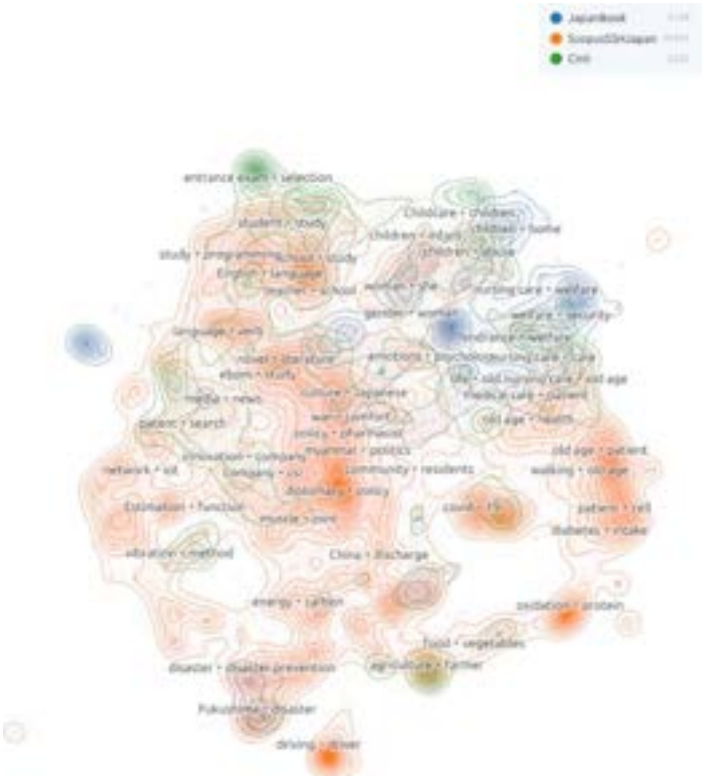

Figure 3 Topic Distribution in Sociology

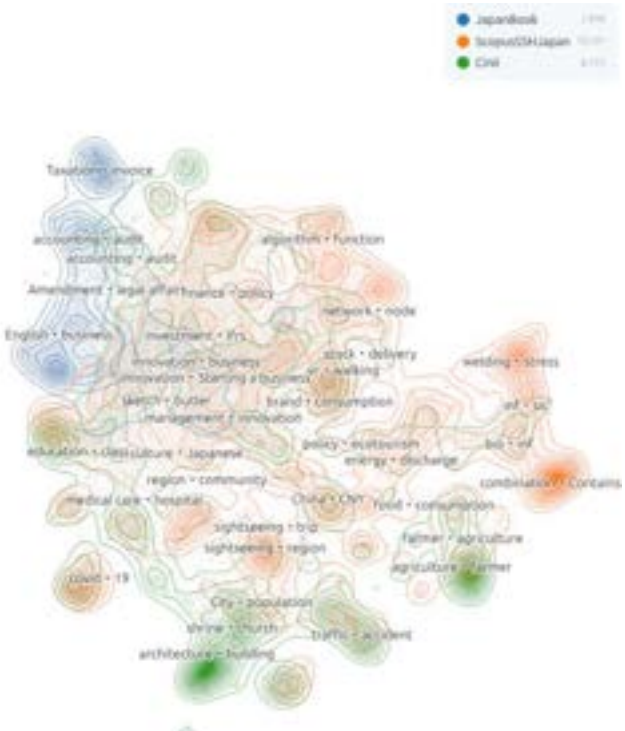

Figure 4 Topic Distribution in Management

References

Honingh, N.V.D.B., Spaapen, J. and Prins, A. (2019) *Evaluating academic research at the crossroads of academic and societal demands*. RESSH 2019.

Rijcke, S. de, Wouters, P.F., Rushforth, A.D., Franssen, T.P. and Hammarfelt, B. (2016) 'Evaluation practices and effects of indicator use: a literature review', *Research Evaluation*, 25(2), pp. 161–169. Available at: https://doi.org/10.1093/reseval/rvv038

Science Council of Japan (2021) *Toward research evaluation for the advancement of science: challenges and prospects for desirable research evaluation*. Available at: https://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-25-t312-1en.pdf (Accessed: 16 April 2026).

Science Council of Japan (2025) *Specific measures to improve research assessment to revitalize research*. Available at: https://www.scj.go.jp/ja/info/kohyo/pdf2/kohyo-26-t394-1en.pdf (Accessed: 16 April 2026).

Whitley, R. and Gläser, J. (2007) *The changing governance of the sciences: the advent of research evaluation systems*. Dordrecht: Springer.

Williams, G., Basso, A., Galleron, I. and Lippiello, T. (2018) 'More, less or better: the problem of evaluating books in SSH research', in Bonaccorsi, A. (ed.) *The evaluation of research in social sciences and humanities*. Cham: Springer International Publishing, pp. 133–158. Available at: https://doi.org/10.1007/978-3-319-68554-0_6

ASSESSING THE VISIBILITY OF NON-ENGLISH SCHOLARLY LITERATURES IN OPEN RESEARCH INFORMATION INFRASTRUCTURES: THE CASE OF JAPANESE-LANGUAGE SSH PUBLICATIONS

HONAMI NUMAJIRI
Research and Development Laboratory, Kyoto University Library, Japan
TAKAYUKI HAYASHI
National Graduate Institute for Policy Studies, Japan

Keywords: Open research information, OpenAlex, Social Sciences and Humanities, Barcelona Declaration

Introduction

The openness and circulation of scholarly knowledge have long been fundamental to scientific progress. From academic journals to Open Science movements, science has advanced alongside infrastructures for knowledge sharing—a principle articulated in Merton's norm of communism (Merton, 1942; Stracke, 2020). However, contemporary evaluation systems rely on commercial databases (Web of Science, Scopus) exhibiting severe linguistic biases, as English overwhelmingly dominates (Numajiri et al., 2025). Such biases pose serious challenges for Social Sciences and Humanities (SSH), where scholarly contributions are often communicated through books, regional journals, and local languages deeply embedded in specific cultural contexts. Recent global initiatives, including the UNESCO Recommendation on Open Science (UNESCO, 2021) and the Barcelona Declaration on Open Research Information (2024), have explicitly recognized these challenges. The 2013 "Facing the Future" conference identified cross-disciplinary synergies as central challenges for SSH research infrastructures (ALLEA, 2014). The Barcelona Declaration responds by positioning open research information infrastructures as foundational to reforming research assessment practices (DORA and CoARA), with emphasis on visibility of non-English scholarship.

OpenAlex, an open scholarly database developed by OurResearch, has attracted growing international attention as a potential alternative. Preliminary analyses indicate substantially greater linguistic diversity than conventional databases, with Japanese representing one of the major non-English scholarly publications. However, little is known about the temporal stability, qualitative composition, and thematic coverage of Japanese-language publications—particularly regarding SSH evaluation suitability.

This study systematically examines Japanese-language publications in OpenAlex, addressing three questions:

(1) Is temporal coverage sufficiently stable for evaluation?

(2) Does document-type composition reflect SSH output diversity?

(3) How do OpenAlex's classification systems capture Japanese SSH themes and societal connections?

Data and methods

This study analyzes OpenAlex data (March 31, 2025, dump; ~190 million records) for publications from 2000–2024. Japanese-language publications were identified using the language metadata field ("ja").

Disciplinary and thematic analysis relied on two classification systems provided by OpenAlex. The legacy Concept system assigns multiple Wikipedia-derived concepts, while the newer Topic system assigns a single primary topic via algorithmic integration of metadata, organized hierarchically (Domain–Field–Subfield–Topic). The Topic system comprises approximately 4,500 topics. Each publication receives scores for all topics, with the highest-scoring topic assigned as the primary topic.

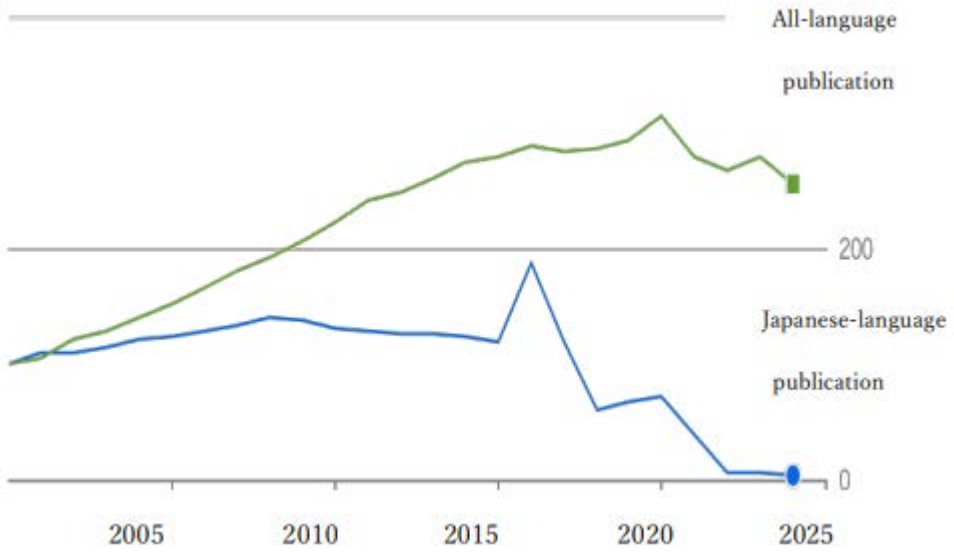

Figure 5 Comparison of Annual Changes in the Number of Japanese Documents and Total Documents Included (the green line represents publications in all languages, while the blue line represents Japanese-language publications. The y-axis indicates the number of publications (in thousands) indexed in OpenAlex)

Three analytical steps were conducted. First, annual publication counts were calculated to examine longitudinal trends and identify structural discontinuities, particularly those associated with changes in data ingestion policies, such as the discontinuation of Microsoft Academic Graph (MAG) sources. Second, document types and publisher distributions were analyzed to assess the balance between data quality and output diversity. Third, Topic-based extraction of HSS publications was combined with Concept-level analysis to compare their effectiveness in identifying socially relevant research themes.

Results

Temporal stability (RQ1)

Time-series analysis reveals substantial instability. Counts increased gradually through the mid-2010s, surged in 2016 (396,373, 1.6-fold increase), then declined sharply from 2018 (124,666, half the 2017 level), with pronounced contraction after 2021(80,061 in 2021; 11,592 in 2022; 12,021 in 2023; 6,067 in 2024).

This volatility corresponds to OpenAlex's data ingestion policy changes. The 2016 surge reflects Microsoft Academic Graph (MAG) integration; the post-2021 collapse stems from MAG's discontinuation (December 2021) and shift toward DOI-centric sourcing, severely diminishing visibility of domestic journals and repositories lacking international identifiers. Overall OpenAlex coverage remained stable at 9-10 million publications annually after 2021, indicating Japanese-language instability reflects sourcing policy rather than research output trends, fundamentally undermining standalone evaluation suitability.

Document-Type composition and output diversity (RQ2)

Japanese-language records are overwhelmingly dominated by journal articles (97.95%). During the MAG period (2000-2021), articles comprised 97.99%, books 1.18%, chapters 0.41%. Post-MAG (2022-2024), articles decreased to 92.27%—not from increased book coverage, but because grants rose from 0.15% to 5.81%. Critically, books declined to 0.05%, chapters to 0.16%.

MAG removal improved quality by eliminating non-academic magazines but also excluded legitimate publishers (Igaku-Shoin: 15,298 records) and university repositories (Hiroshima, Tsukuba, Kyoto), drastically reducing domestic visibility. This quality-diversity trade-off is particularly problematic for SSH, where books constitute primary outputs (Hicks, 2004).

Publisher analysis reveals additional complexity. International commercial publishers (Elsevier: 14.51%, Springer: 3.40%, Wiley: 3.31%) dominate, but URL analysis shows these largely represent machine-translated J-Global metadata (Elsevier: 87,141 records with machine translation tags), creating confusion between genuine Japanese-language originals and translated metadata. URL analysis reveals that during the MAG period, approximately 770,000 Japanese-language publications were ingested via J-Global, with titles containing machine translation indicators such as "JST-Kyoto University machine translation" or "Powered by NICT."

Moreover, non-academic publishers appeared prominently: Bungeishunjū (0.97%), Keizaikai (0.74%), Shōgyōkai (0.53%). These represented general magazine articles (e.g., "Do Citizens Want the Tokyo Olympics?", "The Man Who Awakened Shohei Ohtani's Two-Way Play") rather than scholarly publications, indicating that CiNii's broad coverage of magazine articles had been ingested into OpenAlex.

Thematic coverage and classification systems (RQ3)

Topic analysis identified 21,484 Japanese SSH publications (0.4%), distributed across Social Sciences (10,355), Psychology (4,128), Arts and Humanities (2,266), Business (2,279), Economics (1,566). Top topics reveal Japan-specific societal challenges: Earthquake and Disaster Impact Studies (1,065), Japanese History and Culture (882), Urbanization and City Planning (732), Urban Transport (614), Migration, Aging, and Tourism (368).

Arts and Humanities show severe disciplinary imbalance. Linguistics dominates: Translation Studies (169), EFL/ESL Teaching (151), Language and Discourse (96). Traditional disciplines remain nearly invisible: Literature (87), Philosophy (42), History (29)—reflecting structural bias toward DOI-bearing articles; Japanese humanities predominantly publish in books and domestic journals lacking DOIs.

Concept analysis (level 2) produced substantial noise. While meaningful concepts appeared (Politics: 483, Tourism: 139), cross-disciplinary ambiguity was pervasive (Process, Gene, Scale, Current—natural science terms misapplied). Keyword searches yielded minimal results: Natural disaster (26), Healthy aging (3)—dramatically lower

than Topic equivalents (Earthquake: 1,065; Migration/Aging: 368), demonstrating Topic's superior contextual sensitivity through integrating titles, abstracts, journals, and citations. Publisher distribution (Urban Planning Institute: 2,195; Civil Engineers: 1,205) suggests higher proportions of genuine Japanese-language originals in Social Sciences.

Discussion and implications

This study examined Japanese-language publications in OpenAlex across three dimensions. First, temporal instability stems from data sourcing transitions rather than research output changes. The shift from MAG-based coverage to DOI-centric ingestion fundamentally altered Japanese-language records, severely diminishing visibility of domestic journals and repositories—undermining standalone evaluation viability. Second, MAG removal improved quality by eliminating non-academic magazines but drastically reduced books (1.18%→0.05%) and regional publishers. This quality-diversity trade-off is problematic for SSH, where books constitute primary outputs (Hicks, 2004). The challenge of distinguishing machine-translated metadata from authentic Japanese-language originals persists. Third, Topic analysis effectively identified Japan-specific societal themes (disasters, aging, urban planning), while Concept analysis produced cross-disciplinary noise. Topic classification should serve as the primary framework, with Concept as supplementary. However, traditional humanities remain severely underrepresented (dozens of records), and book coverage remains insufficient despite SSH's book-centric culture. Realizing Barcelona Declaration goals requires deliberate ingestion strategies for disciplinary/linguistic diversity, including improved book coverage, regional publication indexing, and original-language/translation differentiation. This study contributes evidence-based insights to multilingual evaluation debates and guidance for inclusive research information infrastructures.

References

Eisenstein, E.L. (1980) *The printing press as an agent of change*. Cambridge: Cambridge University Press. Available at: https://doi.org/10.1017/CBO9781107049963

Facing the Future: European Research Infrastructures for the Humanities and Social Sciences – ALLEA (2014) *Facing the future: European research infrastructures for the humanities and social sciences*. Available at: https://allea.org/portfolio-item/facing-the-future-european-researchinfrastructures-for-the-humanities-and-social-sciences/ (Accessed: 28 March 2025).

Hicks, D. (2004) 'The four literatures of social science', in Moed, H.F., Glänzel, W. and Schmoch, U. (eds.) *Handbook of quantitative science and technology research: the use of publication and patent statistics in studies of S&T systems*. Dordrecht: Springer, pp. 473–496. Available at: https://doi.org/10.1007/1-4020-2755-9_22

Numajiri, H., Rafols, I., Brasil, A., Hayashi, T. and Hayashi, K. (2025) 'The Barcelona Declaration: building an open research information systems in Europe and Japan', *STI Horizon*. NISTEP. Available at: https://doi.org/10.15108/stih.00403

Stracke, C.M. (2020) 'Open science and radical solutions for diversity, equity and quality in research: a literature review of different research schools, philosophies and frameworks and their potential impact on science and education', in *Lecture*

Notes in Educational Technology, pp. 17–37. Available at: https://doi.org/10.1007/978-981-15-4276-3_2

UNESCO (2021) *UNESCO recommendation on open science*. Available at: https://doi.org/10.54677/MNMH8546

MULTILINGUALISM IN RESEARCH ASSESSMENT: INSIGHTS FROM COMPARATIVE SURVEYS IN ITALY AND HUNGARY

GINEVRA PERUGINELLI, SEBASTIANO FARO, TOMMASO AGNOLONI
National Research Council – Institute of legal Informatics and Judicial Systems (CNR-IGSG), Italy
VIRAG VAS
Hungarian Research Network (HUN-REN) - Research Centre for Natural Sciences, Hungary
ZEHRA TAŞKIN
Hacettepe University, Institute of Informatics, Turkey; Adam Mickiewicz University in Poznań, Scholarly Communication Research Group, Poland

Keywords: Multilingualism, Scientific communication, English dominance, Reducing linguistic biases

Introduction

Multilingualism is widely recognized as a fundamental dimension of global scientific collaboration, shaping the production, dissemination and evaluation of research across disciplines and national contexts (Pölönen et al., 2021). The report "Languages Matter: Global Guidance on Multilingual Education" (2025), released on the occasion of UNESCO's International Mother Language Day, reaffirms the need to embed linguistic diversity within educational systems. In parallel, the UNESCO Recommendation on Open Science (2021) identifies multilingualism as a core principle of scientific practice, promoting inclusive scholarly communication and equitable access to knowledge. Similarly, the Helsinki Initiative on Multilingualism in Scholarly Communication (2019) advocates for publication practices that value research produced in multiple languages and address language biases in research metrics and rankings.

Within this context, the language of scientific communication is not neutral but conveys values, structures epistemic hierarchies and shapes access to visibility, resources and career advancement. While the consolidation of English as the lingua franca of global science has enhanced international dissemination and comparability, it has also introduced new forms of inequality, as research published in other languages is often undervalued within dominant evaluation systems. These dynamics are particularly evident in national contexts where multilingual publication traditions coexist with assessment frameworks strongly oriented toward English-language outputs (Kulczycki et al., 2020).

This contribution examines multilingualism in scientific communication from a research assessment perspective, drawing on empirical evidence from two survey-based studies conducted in Italy and Hungary. The analysis is framed within ongoing reforms such as the Agreement on Reforming Research Assessment (ARRA), promoted by the Coalition for Advancing Research Assessment (CoARA, 2022), which

explicitly recognizes multilingualism as a dimension to be valued rather than discouraged, calling for the assessment of research contributions regardless of the language of communication and for the reduction of inappropriate metric-based practices that introduce language-related biases.

The two surveys on multilingualism and research assessment

The research originates from a collaboration between the National Research Council of Italy (CNR)[1] and the Hungarian Research Network (HUN-REN)[2], developed within the CoARA Working Group "Multilingualism and Language Biases in Research Assessment", where initial exchanges highlighted both shared concerns and structural differences in how linguistic practices are perceived and strategically adopted.

In Italy, the study was first developed through a structured questionnaire designed and implemented by the Institute of Legal Informatics and Judicial Systems of the CNR (IGSG-CNR), with the support of the Institute for Research on Population and Social Policies (IRPPS-CNR). Conducted in late 2024, the survey consisted of 33 questions and explored publication languages, perceived advantages and disadvantages of English dominance, attitudes toward multilingualism and expectations regarding research evaluation. It collected 910 responses from a total population of 6,833 researchers (response rate: 13.3%), providing a detailed account of linguistic practices within Italy's largest public research institution.

In 2025, the Italian survey was translated to Hungarian and conducted to study perceptions of language diversity and biases among the HUN-REN scientific community. Hungarian Research Network (HUN-REN) is Hungary's largest public research institution, with a mission centred on high-level interdisciplinary research. HUN-REN encompasses 11 research centres, 7 research institutes, and 116 supported research groups conducting basic and applied research, the research body is composed of approximately 3,500 researchers. The survey engaged 211 respondents, representing 6% of the employees, and provides valuable insights into language use, attitudes, and evolving linguistic challenges within the academic community.

Overview on results

The Italian findings (Peruginelli and Faro, 2025) reveal a complex linguistic landscape in which the CNR research community regularly operates using multiple languages, with English remaining predominant in formal scientific communication; notably, a substantial number of respondents with Italian as their mother tongue report publishing exclusively in English. At the same time, the language practices vary across disciplinary areas, particularly between Social Sciences and Humanities (SH) and other domains, including Science, Technology, Engineering, Mathematics (STEM).

The data further show that 92.6% of respondents routinely prepare grant applications in languages other than Italian, demonstrating a high degree of engagement with international research contexts. However, approximately 23% report experiencing disadvantages in such situations, pointing to persistent challenges related to linguistic proficiency, communicative precision, and the additional cognitive effort required when working in a non-native language.

The Italian survey also highlights the growing role of language technologies in mitigating linguistic barriers, though unevenly across disciplines and career stages, alongside persistent institutional challenges such as language requirements and limited support services. While most researchers adapt to international linguistic

[1] https://www.cnr.it/
[2] https://hun-ren.hu/research_network

demands, some report difficulties in expressing complex ideas in non-native languages. Disciplinary differences, particularly between SH and other fields, suggest that more context-sensitive language policies may be needed. Despite these challenges, multilingualism is widely perceived as a positive and valuable dimension of scientific work

The results of the Hungarian survey show that 83.5% of respondents indicated that the use of language other than Hungarian does not hinder their participation in funding applications, and they do participate. However, it is necessary to note that 14.5% do feel disadvantaged when using a non-Hungarian language and see this requirement as an obstacle to funding applications.

Although the two surveys have not yet been comparatively analysed and have only been examined individually in detail, both differences and common patterns have already emerged. A very first comparative analysis across the two countries reveals that attitudes toward multilingualism remain largely positive: in Italy, 85% of respondents consider it an essential and enriching component of their professional activity, while in Hungary this proportion is even higher (90.1%). In both cases, respondents tend to view the use of languages other than their own, despite its challenges, as an opportunity for intellectual growth and global collaboration rather than a threat to academic freedom.

Researchers in STEM in both Italy and Hungary tend to adopt English as an almost exclusive language for publications and conference presentations. By contrast, in the SH, a more hybrid use of languages is observed, with national languages playing a significant role depending on the intended audience and contextual factors.

These trends are also reflected in respondents' attitudes toward the role of English in scientific communication and research assessment, particularly in relation to citation databases. As illustrated in Figure 1, respondents in both countries highlight the dual role of English in citation databases, recognizing it as both a key enabler of international visibility and a potential source of marginalization for non-English-language research. This pattern is consistent across the different dimensions explored (a–d), although variations emerge in the intensity of agreement. Italian responses are more dispersed across agreement, neutrality and disagreement, whereas Hungarian responses are more concentrated, with a clear majority expressing concern about the marginalizing effects of English dominance.

Conclusion

The data collected in these two surveys provide a first empirical basis for reflecting on how multilingualism is experienced and negotiated within contemporary research systems, highlighting both shared structural patterns and nationally specific configurations. Building on these findings, the contribution discusses the implications of language practices for research assessment, with particular attention to issues of equity, visibility and the recognition of diverse forms of scientific output. It also explores how current evaluation frameworks may inadvertently reinforce linguistic hierarchies and how ongoing reform processes can incorporate more inclusive and context-sensitive approaches.

Finally, the study points to future research directions, including a comparative analysis of the Italian and Hungarian datasets and the extension of the survey to other national contexts, to better understand how linguistic diversity interacts with institutional and disciplinary factors.

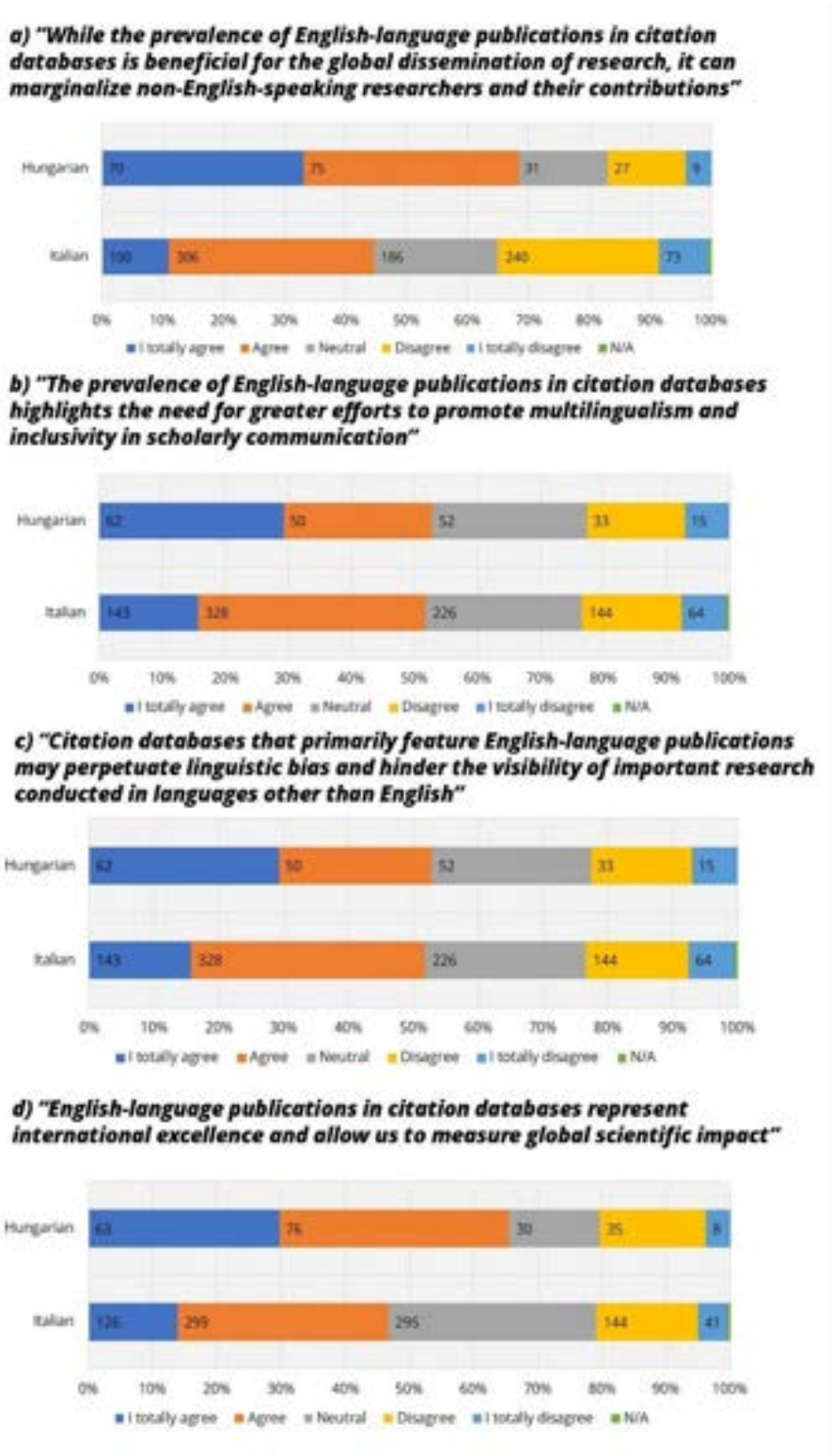

Figure 1. Comparative analysis of perceptions of English dominance in citation databases

References

Amano, T., Ramírez-Castañeda, V., Berdejo-Espinola, V., Borokini, I., Chowdhury, S., Golivets, M. *et al.* (2025) 'Language, economic and gender disparities widen the

scientific productivity gap', *PLOS Biology*, 23(9), e3003372. Available at: https://doi.org/10.1371/journal.pbio.3003372

Coalition for Advancing Research Assessment (CoARA) (2022) *Agreement on reforming research assessment*. Available at: https://coara.eu/agreement/the-agreement-full-text/ (Accessed: 16 April 2026).

Coalition for Advancing Research Assessment (CoARA) (n.d.) *Working group on multilingualism and biases in research assessment*. Available at: https://coara.eu/working-groups/working-groups/wg-multilingualism-and-language-biases-in-research-assessment (Accessed: 16 April 2026).

Helsinki Initiative (2019) *The Helsinki initiative for multilingualism in scholarly communication*. Available at: https://www.helsinki-initiative.org (Accessed: 16 April 2026).

Kulczycki, E., Guns, R., Pölönen, J., Engels, T.C.E., Rozkosz, E.A., Zuccala, A.A., Bruun, K., Eskola, O., Istenič Starčič, A., Petr, M. and Sivertsen, G. (2020) 'Multilingual publishing in the social sciences and humanities: a seven-country European study', *Journal of the Association for Information Science and Technology*, 71(11), pp. 1371–1385. Available at: https://doi.org/10.1002/asi.24336

Peruginelli, G. and Faro, S. (2025) *Exploring multilingualism in research: a survey to researchers at the National Research Council of Italy*. Available at: https://zenodo.org/records/15262952 (Accessed: 16 April 2026).

Pölönen, J., Kulczycki, E., Mustajoki, H. and Røeggen, V. (2021) 'Multilingualism is integral to accessibility and should be part of European research assessment reform', *LSE Impact Blog*, 7 December. Available at: https://blogs.lse.ac.uk/impactofsocialsciences/2021/12/07/multilingualism-is-integral-to-accessibility-and-should-be-part-of-european-research-assessment-reform/ (Accessed: 16 April 2026).

UNESCO (2021) *Recommendation on open science*. Paris: UNESCO. Available at: https://doi.org/10.54677/MNMH8546

UNESCO (2025) *Languages matter: global guidance on multilingual education*. Paris: UNESCO. Available at: https://doi.org/10.54675/MLIO7101

MULTILINGUALISM AND METADATA: HOW MINORITY-LANGUAGE PROTECTIONS INFLUENCE METRIC AND TECHNOLOGY SKEPTICISM IN CANADIAN RESEARCH EVALUATION

QUINTIN KRETH, KEAN BIRCH
York University, Canada

Keywords: Barcelona declaration, Open science, Multilingual research assessment, Research assessment reform, Research information systems

Advancing digitalization is leading to a massive increase in the quantity and interconnectedness of metadata about research – on who produced which research, where, with what money, and to what impact. Internationally, large companies (e.g., Clarivate, Elsevier, Digital Science) have entered the steadily growing business of 'research analytics,' which uses metadata they own and control to track and evaluate research at a completely new scale. These changes represent the transformation of metadata on research into an economic asset for publishers (Birch, 2020; Birch et al., 2021; Birch & Muniesa, 2020).

At the same time, global research rankings fueled by these metadata are continually increasing international competition for research excellence and prestige (Brankovic et al., 2018), termed by Brankovic and colleagues (2018) as the "scarcification" of research reputation. One notable effect of scarcification and the pursuit of global research rankings is the broad-based discouragement of non-English academic publishing (Hazelkorn, 2009; Shchemeleva, 2021). It is well-known that most bibliometric measures of research impact created from metadata are profoundly biased toward English-language scholarship (Rovina et al., 2021; Safón, 2013). This bias relates to both the greater quantity of research published in English than in other languages (Di Bitetti & Ferreras, 2017) and the greater coverage of English-language publications in online sources of metadata (Donathan II, 2025). Thus, universities encourage publication in English to ensure that the scholarship a university generates is included in English-biased metadata systems (Snijder & Kingsley, 2024).

Supporters of the Open Science movement have warned that the rise and spread of paid commercial digital research information systems and associated data analytical products can hamper open and equal access to knowledge about research and, thereby, contribute to the generation of new research inequalities (Aspesi & Brand, 2020; Becerril et al., 2021; Goldenfein et al., 2019; Posada & Chen, 2018). They argue that the development of new digital research information infrastructures should have a democratizing and linguistically inclusive ambition, promoting free and open access and encouraging research topic and language diversity. Advocacy groups for Open Science and responsible metrics, including DORA, CoARA, and the Barcelona Declaration, have grown significantly in recent years, becoming meaningful influences on science policy and governance.

Yet, there are few, if any, systematic, in-depth empirical studies of the influence, affordances, and constraints of the information infrastructures underlying research

evaluation, metrics, and rankings, and how they intersect with issues of language diversity and non-English publishing. This presentation examines these issues in one country, Canada, which, despite being an Anglophone-majority country with a prominent university sector, has been particularly reluctant to embrace the use of metadata and research information management systems in the context of research evaluation. Unlike many European nations, Canada has no national system of research evaluation and impact assessment (Phipps et al., 2025). There is, in fact, minimal ex-post evaluation of publicly funded research in Canada (Advisory Panel on Federal Support for Fundamental Science (Canada) & ISED Canada, 2017). In addition, a significant portion of Canadian university research funding is provided by provincial governments through budgetary allocations, instead of the federal science agencies (Advisory Panel on Federal Support for Fundamental Science (Canada) & ISED Canada, 2017).

Canada also has linguistic differences from most Anglosphere countries. Owing to its mixed colonial history, although most Canadians (94%) speak English, 18% are bilingual in English and French, and a further 4% speak French but not English (Department of Canadian Heritage, 2024; Statistics Canada, 2021). In the last forty years, Francophone Canadians have seen strengthened minority-language rights, particularly in educational settings (Chouinard, 2014; Ricento & Cervatiuc, 2010). Yet, the Canadian scientific community remains subject to many of the same pressures of internationalization and scarcification. Even at Francophone universities, there is significant variation in the prominence of French-language publishing across fields (Gentil & Séror, 2014). Francophone publishing is much more common in the humanities and social sciences than in natural sciences and engineering in Canada (Stockemer & Wigginton, 2019; van Bellen & Larivière, 2024). In this context, we ask: How do both legal and voluntary protections for minority languages in Canada influence research evaluation practices and the adoption and use of research information systems by research universities and funders?

This presentation will provide early empirical insights from an ongoing qualitative study into how datafication and digitalization shape Canadian research evaluation, based on 25 semi-structured interviews with administrators at Canadian universities and research funding agencies, including both bilingual and monolingual organizations. Interview transcripts were analyzed using inductive analysis techniques, identifying emergent themes. This study was conducted under the ethics oversight of the York University Research Ethics Board (REB).

Early findings suggest that the legal mandate to support multilingual research is a major factor shaping the use and disuse of metadata for research evaluation by research funding agencies in Canada. In particular, it appears to place significant technical limitations on the design and content of research evaluation practices. These include limiting the use of bibliometric indicators of research impact, due to evidence of English-language bias. Further, the spread of software and analytical tools, including research information management systems created to support the UK's Research Evaluation Framework (REF), has been discouraged or prevented within bilingual agencies and universities due to their inability to operate in French. These technical limitations, along with Canada's excellence in knowledge-mobilization and co-production (Phipps et al., 2025), appear to be important drivers of the popularity of narrative research evaluation and skepticism towards bibliometric indicators in Canada. We conclude with a comparative discussion of how the cultural and legal contexts of minority-language protection in Canada may plausibly provide greater

protections and more substantive support for non-English scholarship than exist in non-Anglophone countries and science systems.

This work is part of the ongoing multinational research project, "Between Economy and Democracy: Reorganizing Research Evaluation through Metadata in the Digital Era," funded by four national research agencies. This presentation draws upon research supported by the Social Sciences and Humanities Research Council of Canada (SSHRC).

References

Advisory Panel on Federal Support for Fundamental Science (Canada) and Innovation, Science and Economic Development Canada (2017) *Investing in Canada's future: strengthening the foundations of Canadian research*. Ottawa: Advisory Panel on Federal Support for Fundamental Science. Available at: https://publications.gc.ca/pub?id=9.839290&sl=0 (Accessed: 16 April 2026).

Aspesi, C. and Brand, A. (2020) 'In pursuit of open science, open access is not enough', *Science*, 368(6491), pp. 574–577. Available at: https://doi.org/10.1126/science.aba3763

Becerril, A., Bjørnshauge, L., Bosman, J., Frantsvåg, J.E., Kramer, B., Langlais, P.-C., Proudman, V., Redhead, C., Torny, D. and Mounier, P. (2021) *The OA diamond journals study, part 2: recommendations*.

Birch, K. (2020) 'Technoscience rent: toward a theory of rentiership for technoscientific capitalism', *Science, Technology & Human Values*, 45(1), pp. 3–33. Available at: https://doi.org/10.1177/0162243919829567

Birch, K., Cochrane, D. and Ward, C. (2021) 'Data as asset? The measurement, governance, and valuation of digital personal data by Big Tech', *Big Data & Society*, 8(1). Available at: https://doi.org/10.1177/20539517211017308

Birch, K. and Muniesa, F. (2020) *Assetization: turning things into assets in technoscientific capitalism*. Cambridge, MA: MIT Press.

Brankovic, J., Ringel, L. and Werron, T. (2018) 'How rankings produce competition: the case of global university rankings', *Zeitschrift für Soziologie*, 47(4), pp. 270–288. Available at: https://doi.org/10.1515/zfsoz-2018-0118

Chouinard, S. (2014) 'The rise of non-territorial autonomy in Canada: towards a doctrine of institutional completeness in the domain of minority language rights', *Ethnopolitics*, 13(2), pp. 141–158. Available at: https://doi.org/10.1080/17449057.2013.794495

Department of Canadian Heritage (2024) *Statistics on official languages in Canada*. Government of Canada. Available at: https://www.canada.ca/en/canadian-heritage/services/official-languages-bilingualism/publications/statistics.html (Accessed: 16 April 2026).

Di Bitetti, M.S. and Ferreras, J.A. (2017) 'Publish (in English) or perish: the effect on citation rate of using languages other than English in scientific publications', *Ambio*, 46(1), pp. 121–127.

Donathan, D. II, Nason, M., Tullney, M., Shi, J. and Alperin, J.P. (2025) 'Evaluating multilingual metadata quality in Crossref', *arXiv*. Available at: https://arxiv.org/abs/2503.11853

Gentil, G. and Séror, J. (2014) 'Canada has two official languages—or does it? Case studies of Canadian scholars' language choices and practices in disseminating knowledge', *Journal of English for Academic Purposes*, 13, pp. 17–30. Available at: https://doi.org/10.1016/j.jeap.2013.10.005

Goldenfein, J., Benthall, S., Griffin, D.S. and Toch, E. (2019) 'Private companies and scholarly infrastructure—Google Scholar and academic autonomy'. Available at: https://doi.org/10.2139/ssrn.3476911

Hazelkorn, E. (2009) *Impact of global rankings on higher education research and the production of knowledge* (UNESCO Forum on Higher Education, Research and Knowledge Occasional Paper no. 18). Dublin: Technological University Dublin. Available at: https://doi.org/10.21427/D7BG9J

Phipps, D., Graham, K. and Nason, E. (2025) 'Canadian approaches to research impact and its assessment', *Research Evaluation*, 34, rvaf003. Available at: https://doi.org/10.1093/reseval/rvaf003

Posada, A. and Chen, G. (2018) 'Inequality in knowledge production: the integration of academic infrastructure by big publishers', in *Proceedings of the ElPub Conference*.

Ricento, T. and Cervatiuc, A. (2010) 'Language minority rights and educational policy in Canada', in *International perspectives on bilingual education: policy, practice, and controversy*. Charlotte, NC: Information Age Publishing, pp. 21–42.

Rovira, C., Codina, L. and Lopezosa, C. (2021) 'Language bias in the Google Scholar ranking algorithm', *Future Internet*, 13(2), p. 31.

Safón, V. (2013) 'What do global university rankings really measure? The search for the X factor and the X entity', *Scientometrics*, 97(2), pp. 223–244. Available at: https://doi.org/10.1007/s11192-013-0986-8

Shchemeleva, I. (2021) '"There's no discrimination, these are just the rules of the game": Russian scholars' perception of the research writing and publication process in English', *Publications*, 9(1), p. 8. Available at: https://doi.org/10.3390/publications9010008

Snijder, R. and Kingsley, D. (2024) *Research assessment systems and the effects of publication language: manifestations in the directory of open access books*.

Statistics Canada (2021) *2021 census*. Government of Canada.

Stockemer, D. and Wigginton, M.J. (2019) 'Publishing in English or another language: an inclusive study of scholars' language publication preferences in the natural, social and interdisciplinary sciences', *Scientometrics*, 118(2), pp. 645–652. Available at: https://doi.org/10.1007/s11192-018-2987-0

van Bellen, S. and Larivière, V. (2024) 'Les revues canadiennes en sciences sociales et humaines: entre diffusion nationale et internationalisation', *Recherches sociographiques*, 65(1), pp. 15–35. Available at: https://doi.org/10.7202/1113756ar

EXPLORATION OF THE PASTEUR'S QUADRANT IN THE SOCIAL SCIENCES

TIM ENGELS, EUGÉNIE DELZENNE, ELINE VANDEWALLE
Centre for R&D Monitoring (ECOOM), University of Antwerp, Belgium

Keywords: Responsible research assessment, Research evaluation, Effects, Social Sciences and Humanities, Academic work

In her famous chapter 'The four literatures of social sciences' Diana Hicks (Hicks 2004) distinguishes between international journal literature, books, national journals, and non-scholarly press. These genres, although sometimes overlapping, are complementary and serve mostly different purposes. Whereas the international journal literature focuses on advancing basic understanding of social phenomena in a specific time frame or context, scholarly books tend to address social phenomena from a broader temporal and comparative perspective. Books with local publishers and journal articles in national journals are important to disseminate, inform and enlighten more local audiences, including the local scholarly community, while non-scholarly press mainly serves the purpose of communicating with the broader public. Many social scientists combine these genres, e.g. publishing in several languages (Kulczycki et al. 2020) and, as shown by (Verleysen en Weeren 2016), researchers focusing on co-authored English language articles in international journals, are active alongside peers in the same disciplines who are oriented towards (single-authored) book publications and articles in national or regional journals, make use of other languages besides English, and target a mixed audience of academia and interested lay readers. In this contribution we explore these profiles further through the lens of Pasteur's quadrant, a well- known typology developed by (Stokes 1997).

According to Stokes, Pasteur's quadrant is defined by two dimensions: a quest for basic understanding (or not) and an aim for application (or not). Without implying any hierarchy, Stokes thus distinguishes Bohr-like scholars who focus on basic understanding without considering applications, Edison-like scholars who focus on application without paying much attention to basic understanding, as well as Pasteur-like scholars who aim for both basic understanding and application. When nor basic understanding, nor application is the goal, tinkering is going on, e.g. when describing a case.

For sciences and medicine, being active in Pasteur's quadrant is approximated as publishing both journal articles and patents. A recent study (Scharfmann, Marx and Fleming 2025) finds that Pasteur's quadrant researchers produce more novel and more highly cited research over their careers, inspiring us to explore the profile of Pasteur-like scholars beyond the sciences and medicine. More specifically, we here explore the question "what characterizes a Pasteur's quadrant researcher in social sciences"?

First, a Pasteur's quadrant social scientist should have a track record of publishing in international journals and/or publishing books or book chapters with internationally recognized publishers. These as testimony to the quest for basic understanding. The second dimension, paying attention to application, is, however, more difficult to

pinpoint, because application typically does not occur through patents but through dissemination and/or collaboration with non-academic stakeholders (Cunningham, Miller, en Perea-Vicente 2024; Giménez-Toledo et al. 2024). We consider two possible manifestations in social sciences' publishing in Flanders:

- Publishing in Dutch (the official language of Flanders, and one of the official languages of Belgium) or French (another official language of Belgium)
- Collaborating and co-authoring with employees from organisations outside academia.

Note that these aspects can also overlap, possibly indicating a stronger quest for application. Moreover, also publications in German (the third official language of Belgium) or English (a language that is commonly spoken among educated elites, e.g. in Brussels) may serve dissemination purposes, e.g. when published in local journals and/or with local publishers. We intend to explore this further, building on the work by Cristina Arhiliuc (2026) to identify local topics and extending the methodology developed by (Pölönen et al. 2025).

At the occasion of RESSH 2026 we will present first observations on a subset of senior social science researchers, all active at Flemish universities, who combine international publishing with publishing in the regional language (Dutch) as well as co-authoring with organisations outside academia. We will also discuss the possible benefits and limitations of a Pasteur-quadrant approach in the context of assessment and evaluation of social sciences' researchers.

References

Arhiliuc, C. (2026) *Classification and characterization of research output from the social sciences and humanities based on textual metadata*. PhD thesis. Antwerp: University of Antwerp. Available at: https://doi.org/10.63028/10067/2200240151162165141

Cunningham, J.A., Miller, K. and Perea-Vicente, J.-L. (2024) 'Academic entrepreneurship in the humanities and social sciences: a systematic literature review and research agenda', *The Journal of Technology Transfer*, 49(5), pp. 1880–1913. Available at: https://doi.org/10.1007/s10961-024-10136-z

Giménez-Toledo, E., Olmos-Peñuela, J., Castro-Martínez, E. and Perruchas, F. (2024) 'The forms of societal interaction in the social sciences, humanities and arts: below the tip of the iceberg', *Research Evaluation*, 33, rvad016. Available at: https://doi.org/10.1093/reseval/rvad016

Hicks, D. (2004) 'The four literatures of social science', in Moed, H.F., Glänzel, W. and Schmoch, U. (eds.) *Handbook of quantitative science and technology research: the use of publication and patent statistics in studies of S&T systems*. Dordrecht: Kluwer Academic, pp. 473–496.

Kulczycki, E., Guns, R., Pölönen, J., Engels, T.C.E., Rozkosz, E.A., Zuccala, A.A., Bruun, K., Eskola, O., Istenič Starčič, A., Petr, M. and Sivertsen, G. (2020) 'Multilingual publishing in the social sciences and humanities: a seven-country European study', *Journal of the Association for Information Science and Technology*, 71(11), pp. 1371–1385. Available at: https://doi.org/10.1002/asi.24336

Pölönen, J., Alonso Gamboa, J.O., Beigel, F., Digiampietri, L., Laakso, M., Taşkın, Z., Vélez-Cuartas, G. and Kulczycki, E. (2025) 'Beyond the oligopoly: national journal landscapes in Latin America and Europe'.

Scharfmann, E., Marx, M. and Fleming, L. (2025) 'Pasteur's quadrant researchers bring novelty, impact to publishing, and patenting', *Science*, 390(6776), pp. 891–893. Available at: https://doi.org/10.1126/science.adx3736

Stokes, D.E. (1997) *Pasteur's quadrant: basic science and technological innovation*. Washington, DC: Brookings Institution Press.

Verleysen, F.T. and Weeren, A. (2016) 'Clustering by publication patterns of senior authors in the social sciences and humanities', *Journal of Informetrics*, 10(1), pp. 254–272. Available at: https://doi.org/10.1016/j.joi.2016.01.004

2.3B CAREER, GENDER AND EFFECTS OF EVALUATION

REIMAGINING ACADEMIC CAREER ASSESSMENT: INSIGHTS FROM GLOBAL REFORM INITIATIVES TO BROADEN EVALUATION OF ACADEMIC TALENT

RITA MORAIS
European University Association (EUA), Belgium

Keywords: Academic career assessment, Research assessment reform, Diversification of contributions, Diversification of career paths

The global landscape of academic career assessment is undergoing a critical transformation as many institutions seek to move beyond traditional, research-centric metrics toward more comprehensive evaluation frameworks. Conventional systems prioritising publication counts and bibliometrics have been widely criticised for narrowing definitions of academic excellence and undervaluing contributions in teaching, leadership, innovation, and societal engagement. Within this context, the Working Group on Reforming Academic Career Assessment (WG ACA) from the Coalition for Advancing Research Assessment (CoARA), co-chaired by the European University Association (EUA) and the Young Academy of Europe, has mapped existing reform initiatives and developed an interactive framework to support universities and other research performing organisations in their reform efforts.

This presentation aims to showcase the lessons learned from a large-scale international survey and a curated set of case studies on academic career assessment reform. It aims to offer actionable insights for higher education institutions seeking to rethink their assessment practices and better align them with broader institutional missions.

The Working Group conducted a comprehensive mapping exercise comprising two interconnected empirical strands. First, an international survey of higher education and research organisations collected data from more than 200 universities and research performing organisations on current or planned reforms of academic career assessment. The survey explored motivations for change, existing practices, and perceived challenges in revising assessment criteria and mechanisms.

Second, the WG ACA compiled a collection of 11 case studies documenting established national and international reform initiatives (e.g. Spain's ANECA reform, the Netherlands' Recognition & Rewards Programme, DORA).

Data from both sources were analysed to identify recurring themes and patterns, allowing the WG to connect practical evidence from institutions with broader lessons that can support systemic and sustainable change.

The analysis of the survey and case studies revealed a clear international momentum toward reform. A significant proportion of institutions reported actively considering or implementing changes to broaden evaluation criteria, explicitly recognising teaching, leadership, teamwork, societal impact, and other activities beyond research publications.

Across the mapping exercise, six key lessons emerged:

- Reform is a global movement: Institutions worldwide share concerns about traditional assessment models and are collectively moving toward more holistic frameworks, supported by international initiatives such as DORA and CoARA.
- Reform is possible: Institutions with diverse profiles and levels of regulatory autonomy demonstrate that reform can progress step by step, supported by mutual learning and growing communities of practice.
- Academics are central to success: Reforms are most effective when academics from all disciplines and career stages are meaningfully involved from the outset.
- Broader academic roles are gaining recognition: Institutions increasingly aim to merit a wider range of academic activities, including teaching, leadership, teamwork, and engagement.
- Bibliometrics tell a story, but not the whole story: While bibliometrics may provide useful information, they cannot fully capture the breadth of academic contributions; balanced frameworks that combine quantitative and qualitative approaches are essential.
- Reform requires reflection on excellence: Institutions need open and inclusive conversations about how excellence is defined, evaluated, and aligned with disciplinary and organisational missions.

These lessons are supported in evidence from both the survey and case studies, and are reflected in institutional practices, highlighting varied strategies for contextual adaptation.

Building on this analysis, the Working Group created an interactive conceptual framework designed to help universities and other organisations navigate relevant tools, examples, and resources aimed at supporting institutions in creating fairer, more inclusive, and more transparent evaluation processes.

Drawing on empirical evidence and real-world experiences, the WG ACA outputs provide a detailed illustration of the motivations, challenges, and activities that many institutions across Europe are navigating as they undertake academic career assessment reform. The interactive framework offers further support by helping organisations identify and select practical tools and resources adapted to their specific contexts and goals.

Reforming assessment systems is both a catalyst for and a reflection of broader shifts in academic culture. By embracing evidence-informed, participatory, and values-driven approaches, institutions can create environments that recognise diverse academic contributions and foster a more inclusive and balanced understanding of academic excellence. Ultimately, these efforts contribute to strengthening collaboration, innovation, and the societal relevance of higher education.

References

CoARA Working Group ACA (2024). *Rethinking academic career assessment: Lessons and tools for reform*. CoARA Working Group ACA. Available at: https://doi.org/10.5281/zenodo.15473311

CoARA Working Group ACA (2024). *Mapping academic career assessment reform initiatives: Survey outcomes*. CoARA Working Group ACA. Available at: https://doi.org/10.5281/zenodo.14548157

CoARA Working Group ACA (2025). *Collection of case studies on academic career assessment*. CoARA Working Group ACA. Available at: https://doi.org/10.5281/zenodo.14810977

PROMOTION AS A TEMPORAL REGULATOR OF SCHOLARLY WORK IN THE SOCIAL SCIENCES AND HUMANITIES

GÜLEDA DOĞAN
Hacettepe University, Department of Information Management, Turkey
ZEHRA TAŞKIN
Hacettepe University, Institute of Informatics, Turkey; Adam Mickiewicz University in Poznań, Scholarly Communication Research Group, Poland

Keywords: Social Sciences and Humanities, Research evaluation, Academic promotion, Publication dynamics, Metric-based governance

Introduction

Across contemporary higher education systems, academic promotion has become a central mechanism through which research evaluation is operationalized. Promotion criteria translate abstract notions of merit and excellence into concrete thresholds that structure academic careers. In doing so, they do more than assess past performance. They actively shape how, when, and in what form scholarly work is produced.

An expanding body of research documents changes in publication activity around academic promotion, often pointing to intensified output in the years preceding evaluation thresholds (e.g. Azoulay vd., 2011; Espeland & Sauder, 2007; Hlatshwayo & Ngcobo, 2023; Lee, 2014). These studies provide valuable insights into how promotion criteria interact with scholarly productivity and career trajectories. At the same time, differences across disciplinary fields are not always the primary focus of such analyses, and publication dynamics are frequently examined through indicators that are assumed to be broadly comparable.

Building on this literature, this paper approaches academic promotion as a temporal regulator of scholarly work and explores how its effects become visible in the social sciences and humanities. In these fields, dominant evaluation metrics intersect in complex ways with field-specific epistemic practices, publication formats, and research temporalities. Rather than positioning the analysis against existing studies, the paper extends current discussions by foregrounding how promotion-related dynamics may be experienced differently in SSH contexts.

Why social sciences and humanities require a distinct lens

Research in the social sciences and humanities is often characterized by long research cycles, cumulative theoretical development, and publication formats such as books and book chapters that resist short-term quantification (Doğan & Taşkın, 2020; Ochsner et al., 2017). Knowledge production in these fields frequently unfolds through extended engagement with concepts, archives, or social contexts, and its societal relevance may become visible only over time.

Promotion systems, however, increasingly rely on journal-based indicators that privilege rapid output, standardized formats, and international index visibility. When such criteria are applied uniformly across fields, they create a temporal misalignment

for SSH scholars. Evaluation does not simply measure SSH research; it reshapes it by privileging outputs that fit promotion calendars and marginalizing those that do not.

Data and analytical approach

The analysis draws on findings from an ongoing TUBITAK 1001-funded research project examining publication behaviour under metric-based evaluation regimes. Using a nationally representative sample of academics in Turkiye, the study traces publication trajectories around promotion to assistant professor, associate professor, and full professor.

Event-study visualizations, fixed-effects panel models, and publication burst analyses will be used to examine how journal articles, books, and book chapters evolve around promotion events. While the empirical analysis spans multiple disciplinary fields, this paper will interpret the results through a social sciences and humanities perspective, with particular attention to publication forms central to SSH scholarship.

SSH-relevant findings

The results show that journal article output increases systematically in the years preceding promotion, especially at the associate professor stage. Burst analysis indicates that these increases are driven largely by short-term concentration of publication activity among subsets of academics rather than by broadly shared productivity growth.

Books and book chapters display markedly different patterns. Their production shows weaker alignment with promotion timing and lacks sharp pre-promotion peaks. Post-promotion adjustments are limited or absent. These patterns reflect longer production cycles and the comparatively marginal role these outputs play in metric-based promotion frameworks.

For SSH fields, this divergence is consequential. Evaluation systems cantered on journal metrics implicitly redefine scholarly effort in ways that favour short-cycle outputs. Research that depends on sustained inquiry, theoretical depth, or local engagement becomes structurally disadvantaged, regardless of its epistemic or societal value.

Implications for research evaluation reform

The Turkish higher education system offers a revealing context due to its centralized and nationally standardized promotion criteria, which intensify the alignment between evaluation thresholds and publication behaviour. However, the implications extend beyond a single national setting. Any evaluation regime that applies uniform, metric-based promotion criteria across heterogeneous fields risks producing similar distortions.

As part of a TUBITAK-funded project on responsible research evaluation and questionable publishing practices, this paper contributes to ongoing debates on reforming research assessment. It underscores the need for evaluation frameworks that recognize disciplinary temporalities, publication cultures, and the distinctive societal roles of social sciences and humanities research.

Conclusion

By approaching promotion as a temporal regulator rather than a neutral assessment tool, this paper shows how evaluation systems shape the rhythm and orientation of SSH scholarship. Metric-based promotion regimes privilege short-cycle, easily countable outputs while constraining research that depends on long-term engagement and contextual depth.

Addressing these tensions requires evaluation systems that are sensitive to disciplinary diversity and temporal structure. Without such sensitivity, social sciences and humanities research risks being assessed through criteria that systematically misrepresent its contributions and undermine its scientific autonomy.

Funding acknowledgment

This study was conducted as part of the project titled “Preventing Questionable Publishing Practices through Responsible Research Evaluation Policies that Consider Geopolitical Dynamics”, funded by The Scientific and Technological Research Council of Turkey (TÜBİTAK), grant number 223K471.

References

Azoulay, P., Graff Zivin, J.S. and Manso, G. (2011). ‘Incentives and creativity: Evidence from the academic life sciences’. *RAND Journal of Economics*, 42(3), pp. 527–554. Available at: https://doi.org/10.1111/j.1756-2171.2011.00140.x

Doğan, G. and Taşkın, Z. (2020). ‘Humanities: The outlier of research assessments’. *Information*, 11(11), 540. Available at: https://doi.org/10.3390/info11110540

Espeland, W.N. and Sauder, M. (2007). ‘Rankings and reactivity: How public measures recreate social worlds’. *American Journal of Sociology*, 113(1), pp. 1–40. Available at: https://doi.org/10.1086/517897

Hlatshwayo, M.N. and Ngcobo, B. (2023). ‘"Doing just enough to get by": Voices of Black women early career academics on navigating the publish or perish discourse in South Africa’. *Education as Change*, 27, 21 pages. Available at: https://doi.org/10.25159/1947-9417/13352

Lee, I. (2014). ‘Publish or perish: The myth and reality of academic publishing’. *Language Teaching*, 47(2), pp. 250–261. Available at: https://doi.org/10.1017/S0261444811000504

Ochsner, M., Hug, S. and Galleron, I. (2017). ‘The future of research assessment in the humanities: Bottom-up assessment procedures’. *Palgrave Communications*, 3(1), 17020. Available at: https://doi.org/10.1057/palcomms.2017.20

HOW DO HUMANITIES RESEARCHERS EVALUATE THE IMPORTANCE OF RESEARCH OUTPUTS? EVIDENCE FROM A QUESTIONNAIRE SURVEY IN JAPAN

KEIICHI OSHIUMI
National Institute of Japanese Literature; National Graduate Institute for Policy Studies, Japan
TAKAYUKI HAYASHI
National Graduate Institute for Policy Studies, Japan

Keywords: Humanities, Research evaluation, Responsible research assessment, Diversity of research outputs, Language diversity, Career stage and evaluation

Introduction

Research evaluation systems increasingly emphasize accountability, performance measurement, and international comparability. While bibliometric indicators based on citation databases are widely used in the natural sciences, their applicability to the social sciences and humanities remains limited due to diverse publication practices, language use, and research purposes. As a result, social sciences and humanities research evaluation often relies heavily on peer review, yet this raises concerns regarding transparency, cost, and accountability. Furthermore, to correct the situation of over-reliance on bibliometric indicators, a movement toward responsible research assessment/evaluation is emerging through initiatives such as the Leiden Manifesto (2015), DORA(2018), and CoARA(2022). Despite these ongoing debates, relatively little empirical evidence exists on how social sciences and humanities researchers themselves perceive the importance of different types of research outputs.

Previous studies have examined publication patterns in the social sciences and humanities primarily by analyzing bibliographic databases. Hicks (2004), based on a literature review and an analysis of the Social Sciences Citation Index, demonstrated that social science research relies on multiple literature systems — including journal articles, books, domestic publications, and non-academic literature — while citation databases cover only a limited portion of these outputs. Kulczycki et al. (2018), using national research output databases from eight European countries, further revealed substantial cross-national and disciplinary variation in publication types, language use, and coverage by Web of Science. Their findings showed that while Western and Nordic countries tend to emphasize journal articles and English-language publications, other countries rely more heavily on local-language outputs, and that more than half of social sciences and humanities publications are not indexed in international citation databases. While these studies provide important insights into how research outputs are disseminated, they focus on observed publication patterns rather than researchers' own evaluative perspectives. In contrast, relatively little empirical research has examined which types of research outputs social sciences and

humanities researchers themselves consider important, or how such evaluations differ across researchers' attributes.

Regarding societal value, surveys conducted in the United Kingdom indicate that social sciences and Humanities researchers engage extensively in public engagement activities and are more likely than natural scientists to regard such activities as central to their professional roles (Hamlyn 2015)). Sivertsen (2019) similarly argues that, although internationalization is crucial for research quality and specialization in the humanities, interaction with society is equally important for fulfilling the broader purpose of knowledge creation. The Science Council of Japan(2017) has also characterized one distinctive feature of social sciences and humanities research as the critical examination of "values" and the cultivation of "values shared by humanity." Nevertheless, no comprehensive empirical study has systematically examined both academic and societal dimensions of research output importance from the perspective of social sciences and humanities researchers themselves.

This study aims to address this gap by examining how humanities researchers evaluate the importance of various research outputs and whether these evaluations differ according to researchers' attributes. By focusing on researchers' own perspectives, the study seeks to provide empirical foundations for more responsible and context-sensitive research assessment/evaluation in the humanities, particularly in contexts where standardized and indicator-based evaluation frameworks are increasingly influential. The survey was deliberately limited to humanities researchers, as the broader social sciences and humanities category covers a wide range of disciplines, making systematic questionnaire design more difficult. Focusing on the humanities was therefore considered an appropriate starting point for this analysis.

Methods

A questionnaire survey was conducted between April and May 2024 targeting researchers affiliated with six institutes of the National Institutes for the Humanities (NIHU) in Japan. A total of 138 valid responses were obtained (response rate: 49.8%). Respondents were asked to assess the importance of 36 types of research outputs using a five-point Likert scale. These outputs covered a broad range of scholarly and societal contributions, including peer-reviewed foreign-language journal articles, Japanese-language academic books, edited volumes, public lectures, exhibitions, and outreach activities. Demographic and professional variables—such as gender, age group, academic rank, employment status, and research field—were also collected.

The analysis proceeded in three stages. First, descriptive statistics were calculated to identify overall patterns in the perceived importance of different research outputs. Second, mean differences by gender were examined using Welch's t-tests. Third, differences by age group, academic rank, employment status, and research field were analyzed using one-way ANOVA or Welch ANOVA where assumptions of homogeneity of variance were violated. Effect sizes (Cohen's d and η^2) were reported to assess the magnitude of observed differences. When significant main effects were detected, post-hoc comparisons were conducted using pairwise t-tests with Bonferroni correction.

Results

The results reveal several notable patterns. Overall, research outputs written in Japanese—particularly Japanese-language academic books and peer-reviewed journal articles—received the highest average importance ratings, reflecting the continued centrality of academic evaluation criteria in the humanities in Japan. At the same time, this overall pattern masks substantial variation across researcher attributes.

Gender-based comparisons showed that female researchers consistently rated several types of foreign-language research outputs, including peer-reviewed co-authored journal articles ($p<0.006$, Cohen's $d=0.46$), higher than male researchers, with moderate effect sizes observed across multiple items. These findings suggest gender-related differences in how international and collaborative research activities are valued within the humanities.

Analyses by age group and academic rank revealed clear career-stage effects. Early-career researchers and those in fixed-term or tenure-track positions placed significantly greater importance on peer-reviewed foreign-language journal articles($p<0.05$, $\eta^2 \approx 0.09–0.10$), likely reflecting strategic considerations related to career advancement and prevailing evaluation requirements. In contrast, senior researchers and those without fixed-term appointments assigned higher importance to Japanese-language academic books, edited volumes ($p<0.049$, $\eta^2=0.10$), and editorial work ($p<0.0004$, $\eta^2=0.16$), indicating a broader valuation of scholarly contributions beyond journal publications.

Field-specific analyses further highlighted the diversity of humanities research practices. Differences in language orientation, preferred publication formats, and engagement with society were observed across fields, underscoring that a single, uniform evaluation framework cannot adequately capture the value of humanities research outputs.

Conclusion

These findings demonstrate that humanities researchers' evaluations of research output importance are shaped by multiple intersecting factors, including gender, career stage, employment conditions, and disciplinary context. The results suggest that evaluation systems relying narrowly on a limited set of outputs — particularly foreign-language journal articles — risk misrepresenting the diversity of scholarly contributions in the humanities and reinforcing uneven evaluative pressures across career stages.

From the perspective of responsible research assessment, this study provides empirical evidence supporting calls to move beyond one-size-fits-all metrics and to recognize a broader range of research outputs and career trajectories. By systematically documenting how humanities researchers themselves value different forms of scholarly and societal contributions, the study contributes to ongoing international discussions on research evaluation reform, including initiatives such as DORA and CoARA.

Future research should expand the survey to a broader national and international sample and incorporate qualitative methods, such as interviews, to explore the motivations underlying researchers' evaluation preferences. Such mixed-method approaches will be essential for developing research assessment frameworks that are both evidence-based and sensitive to the diversity of humanities scholarship.

Acknowledgement

This study was supported by JSPS KAKENHI Grant Number 23K00271.

References

Coalition for Advancing Research Assessment (CoARA) (2022). *The Agreement on Reforming Research Assessment*. Available at: https://coara.eu/app/uploads/2022/09/2022_07_19_rra_agreement_final.pdf (Accessed: 16 April 2026).

Declaration on Research Assessment (DORA) (2018). *Declaration on Research Assessment*. Available at: https://sfdora.org/read/ (Accessed: 5 March 2026).

Hamlyn, B., Shanahan, M., Lewis, H., O'Donoghue, E., Hanson, T. and Burchell, K. (2015). *Factors affecting public engagement by researchers*. TNS-BMRB.

Hicks, D. (2004). 'The four literatures of social science', in Moed, H.F., Glänzel, W. and Schmoch, U. (eds.), *Handbook of Quantitative Science and Technology Research*. Springer, Dordrecht, pp. 473–496. Available at: https://doi.org/10.1007/1-4020-2755-9_22

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

Kulczycki, E., Engels, T.C.E., Pölönen, J. et al. (2018). 'Publication patterns in the social sciences and humanities: Evidence from eight European countries'. *Scientometrics*, 116, pp. 463–486. Available at: https://doi.org/10.1007/s11192-018-2711-0

Science Council of Japan, Division I, Subcommittee on the Role and Promotion of the Humanities and Social Sciences (2017). *Recommendations for the comprehensive development of scholarship: Perspectives from the social sciences and humanities*. Science Council of Japan [in Japanese].

Sivertsen, G. (2019). 'Understanding and evaluating research and scholarly publishing in the social sciences and humanities (SSH)'. *Data and Information Management*, 3(2), pp. 61–71.

RESEARCH INACTIVITY IN SOCIAL SCIENCES: AN INCOMPLETE AND BIASED METRIC

MASSIMILIANO MAZZANTI, DAVIDE ANTONIOLI, FEDERICO FRATTINI
University of Ferrara, Department of Economics and Management, Italy
LORETTA MISSIROLI
University of Ferrara, Bibliometrics and Databases Office, Italy

Keywords: Research evaluation, Social Sciences and Humanities, Inactive researcher, Responsible metric, CoARA

Scientific research is increasingly subject to pressures stemming from several external factors. These include societal changes fostering a more diverse and inclusive global research community, shifts in priorities driven by global challenges, disruptive governmental initiatives that rapidly reshape research opportunities worldwide, and emerging technologies that profoundly impact the practice, assessment, and integrity of research. As highlighted by RESSH2026, these pressures can interact, potentially weakening the research ecosystem.

Assessment and the application of standardized metrics are vital practices for maintaining the acknowledgment, resilience, and advancement of science. However, poorly designed or inappropriately applied metrics can compromise the fundamental mission of research assessment. Consequently, assessment practices must be continuously updated and adapted to reflect changes in the research ecosystem.

In Italy, a key metric for assessing research quality since the inaugural Research Quality Assessment (VQR) exercise in 2011 (Ancaiani et al., 2015) is the count of "inactive" researchers. A researcher is classified as inactive if they fail to be credited with a non-null record of relevant research products (e.g., publications or cognate outputs) within a specified timeframe, which is currently five years (MIUR, 2013). This binary metric (1 for active, 0 for inactive) is incorporated into the comparative assessment and funding of research institutions, functioning as a penalty on other research productivity indicators.

This Italian practice is not unique. For instance, research assessment in Spain employs "Sexenios", which are six-year blocks where researchers must present five high-quality contributions to receive productivity certification (Relaciones con las Cortes, 1989; Jefatura del Estado, 2023). In contrast, the United Kingdom's Research Excellence Framework (REF) has evolved from an individual selectivity model, which excluded inactive researchers to inflate departmental scores, to a more inclusive system that assesses all staff with significant responsibility for research (Stern, 2016; REF 2020). This system is currently moving towards rewarding the abandonment of binary indicators (REF, 2023). Conversely, Australia abolished its Excellence in Research for Australia (ERA) in 2023, citing excessive bureaucratic burden and the failure of binary indicators to adequately reflect modern research impact (Sheil et al., 2023; Clare, 2023; Prior & Brennan 2024).

Universities and research institutions, within their autonomy, may also refine this metric or extend its application to their internal research assessment, resource

allocation, and salary progression. However, for these "local" applications, the observation window often narrows from five years to two years (e.g., University of Ferrara). This increases the risk of misalignment with the longer inherent cycles of scholarly research.

As the active/inactive status is intended to document researchers' engagement and effort, it holds significance for steering, consolidating, and enhancing the research ecosystem. Nevertheless, the binary nature of this metric can occasionally misrepresent research effort, potentially leading to a paradox: researchers who are demonstrably scholarly active and impactful through their involvement in research projects and programs may be bibliometrically assessed as inactive. Despite its seemingly straightforward definition, calculating this metric critically depends on two factors: first, the types of research outputs included in a record; and second, the length of the time span considered for that record. This issue is particularly pronounced within the Social Sciences and Humanities (SSH).

On the one hand, research is inherently unobservable in both quantitative and qualitative terms. This intrinsic challenge of knowledge production necessitates indirect assessment through proxies such as input, output, and impact (Keller, 2004). Given the qualitative heterogeneity of outputs, research assessment commonly relies on bibliometrics or surrogate metrics, depending on domain-specific research practices, which often involves restricting the types of relevant research outputs to a limited range of publication outlets. However, global transformations, the diversification of funding sources, the urgent need to collect new and reliable data, and an increasing demand for transdisciplinarity and documenting societal impact over merely scientific productivity, tend to destandardize research outputs. This boosts the likelihood of misalignment between research efforts and the monitored outputs.

On the other hand, this likelihood is potentially stronger in the SSH due to a longer publication lag, that is, the editorial delay between submission and actual publication, compared to STEM fields. Björk & Solomon (2013) documented that the average publication lag in Business and Economics (18 months) is twice as long as in Chemistry (9 months). This disparity is at least partially attributable to differing editorial practices across fields: while STEM fields favor rapid, short articles with peer-review cycles typically lasting 3–6 months, SSH research frequently results in complex long-form articles or monographs that are expected to convincingly "tell a story" by contextualizing conjectures and evidence. Furthermore, the citation half-life of SSH research is significantly longer (10–20 years) compared to the rapid obsolescence of STEM results (2–4 years), suggesting that a focus on short-term productivity overlooks the enduring impact of SSH scholarship (Larivière et al. 2006; Nederhof, 2006).

To address these inherent distortions, we propose a three-level strategy aligned with the principles of responsible research assessment. These principles include depowering journal-based metrics (DORA, 2012), valuing expert judgment (Hicks et al., 2015), acknowledging the diversity of research practices, contributions, and outputs, basing assessment primarily on qualitative judgment supported by responsible quantitative indicators, avoiding the trickle-down of simple institutional metrics and rankings into individual evaluations, and relying on more narrative evidence to provide context for interpreting research (CoARA, 2022).

The first level involves rigorously isolating the measure of research activity from complementary measures of productivity, output volume, and quality, which restrict outputs to a limited variety. The second level advocates for constructing the measure of activity based on observable research inputs to capture effort beyond publications.

Examples include submitted research proposals, official timesheets for research grants, conference presentations, and editorial and peer-review activities. The third level requires setting and maintaining updated, evidence-based, field-specific assessment windows.

To substantiate our proposal, we present a case study utilizing the dataset of research information compiled at the University of Ferrara for the VQR 2020–2024. Early results confirm a relevant editorial asymmetry between the SSH and STEM that could be mitigated by harmonizing assessment windows with field-specific practices. Furthermore, the findings demonstrate that the measure of research activity can indeed be detached from complementary measures, allowing "hidden" effort to emerge and providing insightful elements for a more aware and responsible assessment of research productivity and quality. As suggested by CoARA (2022), research institutions should leverage their autonomy to experiment with new or amended metrics and processes for potential upscaling to the broader research community if successful, thereby aligning with more advanced international practices such as the Norwegian Career Assessment Matrix (UHR, 2021), the system of recognition and rewards in the Netherlands (VSNU et al., 2019), and the drive for implementing DORA principles in the United States (e.g., Wong, 2026).

References

Ancaiani, A. et al. (2015). 'Evaluating scientific research in Italy: The 2004–10 research evaluation exercise'. *Research Evaluation*, 24, pp. 242–255. Available at: https://doi.org/10.1093/reseval/rvv008

Björk, B.-C. and Solomon, D. (2013). 'The publishing delay in scholarly peer-reviewed journals'. *Journal of Informetrics*, 7, pp. 914–923. Available at: https://doi.org/10.1016/j.joi.2013.09.001

Clare, J. (2023). 'Response to Australian Research Council Review'. Minister's Media Centre, Canberra. Available at: https://ministers.education.gov.au/clare/response-australian-research-council-review

DORA (2012). 'San Francisco Declaration on Research Assessment'. San Francisco, CA. Available at: https://sfdora.org/read/

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

Jefatura del Estado (2023). 'Ley Orgánica 2/2023, de 22 de marzo, del Sistema Universitario'. *BOE* núm. 70, de 23/03/2023. Madrid.

Keller, W. (2004). 'International technology diffusion'. *Journal of Economic Literature*, 42, pp. 752–782. Available at: https://doi.org/10.1257/0022051042177685

Larivière, V., Archambault, É., Gingras, Y. and Vignola-Gagné, Y. (2006). 'The place of serials in referencing practices: Comparing natural sciences and engineering with social sciences and humanities'. *Journal of the American Society for Information Science and Technology*, 57, pp. 997–1004. Available at: https://doi.org/10.1002/asi.20349

MIUR (2013). *Decreto Autovalutazione, Accreditamento Iniziale e Periodico delle Sedi e dei Corsi di Studio e Valutazione Periodica*. D.M. 47/2013, aggiornato con D.M. 1059/2013. Ministry of Education, University and Research, Rome.

Nederhof, A.J. (2006). 'Bibliometric monitoring of research performance in the social sciences and the humanities: A review'. *Scientometrics*, 66, pp. 81–100. Available at: https://doi.org/10.1007/s11192-006-0007-2

Prior, R. and Brennan, T. (2024). *Australian Research Council Amendment (Review Response) Bill 2023*. Parliament of Australia, Canberra.

REF (2020). *Guidance on submissions (2019/01) – REF 2021*. Research Excellence Framework, London.

REF (2023). *Initial decisions next steps December 2023 – REF 2029*. Research Excellence Framework, London. Available at: https://2029.ref.ac.uk/news/update-on-initial-decisions/

Relaciones con las Cortes (1989). 'Real Decreto 1086/1989, de 28 de agosto, sobre retribuciones del profesorado universitario'. *BOE* núm. 216, de 09/09/1989. Madrid.

Sheil, M., Dodds, S. and Hutchinson, M. (2023). *Trusting Australia's ability: Review of the Australian Research Council Act 2001. Final Report*. Australian Government, Canberra.

Stern, N. (2016). *Research Excellence Framework review*. Government of the United Kingdom, London.

UHR (2021). *NOR-CAM: A toolbox for recognition and rewards in academic careers*. Universities Norway, Oslo.

VSNU et al. (2019). *Room for everyone's talent: Towards a new balance in the recognition and rewards of academics*. VSNU, NFU, KNAW, NWO and ZonMw, The Hague.

Wong, M.L. (2026). 'Reckoning with my "ghost years": Why a high publication rate doesn't always reflect success'. *Nature*, d41586-025-04062-9. Available at: https://doi.org/10.1038/d41586-025-04062-9

2.4A CRITERIA, EXCELLENCE AND DYNAMICS IN RESEARCH EVALUATION

WHAT IS INCLUSIVE EXCELLENCE? REFLECTIONS ON EXCELLENCE AND EQUALITY IN RESEARCH EVALUATION

MAURA CONIGLIONE
University of Milan Bicocca, Italy
LUCA FALZEA
Scuola Normale Superiore, Italy
GIULIANA RUBBIA RINALDI
National Institute of Geophysics and Volcanology, Italy
ARIANNA MONTORSI
Politecnico of Turin, Italy
HELENE SCHIFFBÄNKER
Joanneum Research, Austria

Keywords: Research evaluation, Inclusive excellence, Gender, Academia, Equality

This contribution was developed in the context of the CoARA Working Group TIER - Towards an Inclusive Evaluation of Research. It investigates how the notions of excellence and gender are interlinked and explores how excellence can be de-gendered. Finally, we try to develop an idea of a more (gender) inclusive excellence concept in the context of CoARA by bringing in various perspectives.

European-level data confirms the structural patterns that motivate this study. According to the She Figures 2024 report (European Commission), gender balance has been achieved at the doctoral level (women represent ~48–50% of PhD graduates across the EU), yet this parity does not translate into senior academic positions: only 30% of Grade A positions (full professors, senior research leaders) are held by women, dropping to 19% in engineering and 24% in natural sciences. The Glass Ceiling Index quantifies this progressive attrition, with women more represented at early career stages but leaving academic careers at higher rates than men before reaching senior positions. Additional structural barriers compound this trajectory: women are more likely to hold precarious or part-time contracts, have lower success rates in securing competitive research funding, and move less internationally — a key factor for career advancement.

Our contribution focuses on research assessment and what defines the boundaries of excellent research work. We argue that the current definition of excellence is, first, based on formal indicators such as the number of citations of one's works and the number of research outputs produced by researchers and research groups. Second, excellence is constructed when these indicators are applied in practice, introducing potential bias factors such as (gendered) assumptions, power relations, and field differences.

Researchers are thus evaluated on their ability to adhere to a standard ideal model of worker (Acker, 1990) that entails complete dedication to one's job and the capacity to

build meaningful networks that increase one's institution's prestige and funding: all of these elements are connected to a certain mode of knowledge production, focused on resources and prestige, related to what Slaughter and Leslie (1997) define as academic capitalism. This system is gendered in nature: research has revealed that, so far, white men are the most favored group within research institutions due to the way knowledge is produced and assessed (Ward & Rodriguez, 2025). Indeed, this system can only function if, while some researchers dedicate themselves entirely to remunerative research, others carry out what can be defined as "academic care work" — all those activities that do not necessarily produce patents, papers, or financially driven outcomes but are vital to institutions, such as teaching, group care, organization, and dissemination. These activities are largely carried out by women in most cases (van den Brink & Benschop, 2012; Gaiaschi, 2025; Järvinen & Mik-Meyer, 2025). Dissemination and outreach activities, which are crucial for legitimizing public funding spent on research, are also barely covered by this definition of excellence.

This leads to an imbalance between women's and men's career opportunities within academia and research institutions: a research world that was historically built for a certain type of man, one that inevitably places women and other under-represented groups at the margins. This disparity in evaluation systems and career opportunities for men and women contributes to the Matthew Effect, as resources tend to go to already established, prestigious researchers and networks. Moreover, women often have to navigate a "chilly climate" (Britton, 2017), marked by microaggressions and gender-based violence, especially in male-dominated fields (Nowrouzi-Kia et al., 2025; Kocsis, 2025).

Analyzing the perceptions of 1,682 researchers (from various institutions, countries, and fields of study) on evaluation systems, using a scale from 1 to 5 (where 1 = not important and 5 = very important), we found that most respondents believe their career was assessed mainly through the following criteria: high-impact publications ($M=4.16$), research grants ($M=4.04$), citations ($M=3.77–3.85$), and H-index ($M=3.72–3.81$), while teaching, peer review, and open access were ranked lowest ($M=2.08–2.81$). Overall, traditional metrics outscored non-traditional contributions by more than 1.2 points ($d>1.2$, $p<.001$).

Women rated all criteria significantly higher than men (18 out of 20 criteria), indicating that academic excellence is perceived as more challenging for women. The largest gender gaps were found for open access, peer review, and data sharing ($ds\approx0.30$), all of which relate to scientific infrastructure and are also the least rewarded by institutions and funding agencies.

Additionally, 31.2% of respondents agreed or strongly agreed that the definition of excellence is gender-biased. Specifically, regression analysis using a Cumulative Link Mixed Model (CLMM; $N=1,277$) revealed gender as the main predictor of the perception of gender bias in definitions of excellence on a Likert scale (1=Free from bias, 5=Totally biased), with women 3.45 times more likely than men to perceive bias in the current definition of excellence. Age, field of research, and career stage were not significant predictors of the perception of bias in evaluation, indicating a stable gender gap irrespective of these variables.

Turning to strategies for redefining excellence, the acceptance of using criteria other than traditional bibliometrics was moderately high ($M=3.37$) and significantly higher for women ($M=3.56$) than for men ($M=3.13$; $p<.001$). Women also reported a greater perceived negative effect of career interruptions on evaluation ($d=0.51$). Finally, bias training for evaluation committees and institutional changes to address career gaps and bias detection tools were identified as the most effective strategies for change.

In addition to the quantitative data collected through the survey, we investigated best practices implemented in European countries to mitigate gender bias in research assessment, including, among others, narrative CVs, gender-blind recruitment and assessment, and greater recognition of teaching, dissemination, and administrative work.

Through the analysis of these data, we conclude that an inclusive excellence framework could incorporate a more qualitative approach to research evaluation, a more nuanced use of bibliometric indicators to avoid their misuse, and the promotion of a broader range of research activities, fully recognizing them as part of what excellent research entails. More broadly, this would also mean acknowledging the current neoliberal turn experienced by universities and research institutions and taking steps to counter it, in order to establish impactful research and its dissemination (through formal teaching or otherwise) as the primary focus of research activities, rather than money and prestige.

References

Acker, J. (1990). 'Hierarchies, jobs, bodies: A theory of gendered organizations'. *Gender & Society*, 4(2), pp. 139–158. Available at: https://doi.org/10.1177/089124390004002002

Britton, D.M. (2017). 'Beyond the chilly climate: The salience of gender in women's academic careers'. *Gender & Society*, 31(1), pp. 5–27. Available at: https://doi.org/10.1177/0891243216681494

Gaiaschi, C. (2025). 'Breaking the glass door in academia? Looking at the role of scientific fields and contextual factors in moderating the gender gap in recruitment: Evidence from Italy'. *Higher Education*.

Järvinen, M. and Mik-Meyer, N. (2025). 'Giving and receiving: Gendered service work in academia'. *Current Sociology*, 73(3), pp. 302–320. Available at: https://doi.org/10.1177/00113921231224754

Kocsis, J. (2025). 'Research, rigour, and rape: Facing the reality of gender-based violence in academic fieldwork'. *Gender, Place & Culture*, 32(3), pp. 480–488.

Nowrouzi-Kia, B., Yi-Chan, H., Zhu, S., Nandan, S., Bani-Fatemi, A., Howe, A., Gross, D.P., Gohar, B., Yazdani, A. and Kumar Chattu, V. (2025). 'Examining the prevalence and effects of gender-based violence in academic settings: A systematic review and meta-analyses'. *Trauma, Violence & Abuse*, 26(4), pp. 755–768.

Slaughter, S. and Leslie, L.L. (1997). *Academic capitalism: Politics, policies, and the entrepreneurial university*. Johns Hopkins University Press, Baltimore.

van den Brink, M.C.K. and Benschop, Y. (2012). 'Slaying the seven-headed dragon: The quest for gender change'. *Gender, Work & Organization*, 19(1), pp. 71–92. Available at: https://doi.org/10.1111/j.1468-0432.2011.00566.x

Ward, L.W. and Rodriguez, D. (2025). '"I am for diversity...": How a victimhood legal formula weaponizes faculty academic freedom against diversity, equity, inclusion, and social justice'. *Education Sciences*, 15(10), Article 1364. Available at: https://doi.org/10.3390/educsci15101364

SEE YOU IN COURT! JUDICIAL DYNAMICS IN RESPONSIBLE RESEARCH EVALUATION

JANNE PÖLÖNEN
Federation of Finnish Learned Societies (TSV), Finland
GINEVRA PERUGINELLI, SEBASTIANO FARO
National Research Council – Institute of Legal Informatics and Judicial Systems (CNR-IGSG), Italy
MARIA PIETILÄ
University of Eastern Finland, Finland
SAKARI MELANDER
University of Helsinki, Finland

Keywords: Research evaluation, Judicial review, University autonomy, Academic recruitment, Comparative legal analysis

Research assessment is a central part of academic careers, because it affects recruitment, promotion, and resource allocation across European universities and research institutions. The Coalition for Advancing Research Assessment (CoARA) has established international principles emphasizing qualitative peer review, transparency, and institutional autonomy. However, the implementation of responsible research assessment (RRA) does not occur in a policy vacuum – it unfolds within specific legal frameworks that determine how evaluation disputes are resolved and how principles translate into enforceable practices. The European Research Area Policy Agenda 2025-2027 explicitly recognizes legal and administrative frameworks as essential components for achieving inclusive and accountable assessment systems. However, despite this strong policy emphasis, existing scholarship has put limited attention to how legal mechanisms, and in particular judicial remedies and litigation, actually influence researcher assessment in practice. As a result, the role of courts and legal dispute resolution in shaping the implementation of responsible research assessment (RRA) remains insufficiently explored.

This study addresses this gap through a comparative analysis of Italy and Finland, two countries with different legal cultures and governance systems. Italy represents a highly regulated environment where researchers hold civil servant status and recruitment involves dense procedural architecture including national qualification mechanisms and multi-layered institutional decisions. Finland represents a system with stronger institutional autonomy where universities are independent employers of staff with considerable discretion in recruitment and career structures. The study asks: What kinds of disputes emerge in researcher recruitment? To what extent does judicial review intervene in assessment processes? How do legal frameworks and judicial remedies contribute to defining the boundaries of responsible research assessment?

Methodologically, the study combines desk research and case law analysis. For Italy, the analysis draws on systematic documentation of administrative jurisprudence examining over 1500 appeals from administrative courts (Tribunali Amministrativi

Regionali and Consiglio di Stato) concerning university recruitment, supplemented by qualitative analysis of fifteen labour court decisions (2020-2024) involving National Research Council of Italy (CNR) career progression disputes. For Finland, the study analyzes statements from two ombudsman institutions: the Equality Ombudsman and Parliamentary Ombudsman concerning academic recruitment at Finnish universities between 2014 and 2022.

The data about Italian universities' recruitment processes reveals extensive litigation concentrated around several recurring issues: lack of reasoning in evaluation decisions, excess of discretionary power, procedural irregularities, committee composition conflicts, transparency violations, and false candidate declarations. Italian courts cannot substitute their judgment for academic experts on scientific merit, but they exercise substantive oversight ensuring evaluations comply with principles of legality, transparency, and reasoned decision-making. The analysis demonstrates how administrative jurisprudence has progressively shaped practical interpretations of merit, with courts demanding that committees articulate decisions through externally intelligible reasoning. This creates a bridge between scientific judgment grounded in peer expertise and administrative legality grounded in procedural transparency. However, this system also generates a sort of "judicial path" to academic positions, where career advancement may depend as much on litigation success as on peer recognition.

The CNR labour court cases reveal structured patterns in how judicial review operates within predetermined boundaries. While courts consistently refuse to re-evaluate scientific merit directly, they exercise full review of compliance with selection criteria, equal treatment, transparency, and logical coherence. The most common remedy is ordering committees to re-evaluate candidates according to correct criteria, preserving technical discretion while removing procedural defects. Direct score attribution by courts occurs only in exceptional cases where criteria are entirely predetermined and errors purely objective. Compensation for loss of opportunity is rarely granted, requiring demonstration that candidates would have reached minimum thresholds for advancement.

The Finnish context presents a contrasting model where judicial intervention remains minimal. Finnish universities design academic career structures independently, with only professorships subject to legislative requirements for public announcement and peer review. The 2010 Universities Act transformed academic staff from state civil servants to university employees, raising the threshold for appealing recruitment decisions. In practice, the ombudsman institutions serve as primary oversight mechanisms, issuing non-binding statements that influence practices through publicity rather than legal enforcement. Analysis of seven Equality Ombudsman statements reveals that disputes center on merit comparison flaws: application of different criteria for candidates, omission of qualifications, problems with objective assessment, deviation between published and applied criteria, and non-transparency of processes. Since the Equality Ombudsman is authorised to make judgements about the employers' compliance with the Equality Act in Finland, claimants are required to frame their appeals as gender-based discrimination. The Ombudsman emphasizes systematic assessment using common matrices and interview structures, equal treatment throughout processes, and objective comparison of merits related to position requirements. However, the Ombudsman does not evaluate the scientific merits of candidates or reverse recruitment decisions. Its influence comes from publicly identifying procedural problems, not from legal authority to impose sanctions. Parliamentary Ombudsman statements (2014-2017) address good

governance principles, emphasizing that despite university autonomy, recruitment must comply with Constitutional requirements for equal treatment and Administrative Procedure Act standards. Key interventions concern binding nature of formal eligibility requirements, prohibition against changing assessment criteria mid-process, and ensuring qualification periods are not granted selectively to favor specific candidates.

The comparative analysis reveals different approaches to institutionalizing responsibility in research assessment. Italy externalizes responsibility through formalization and judicialization: procedural compliance becomes a proxy for fairness, with litigation functioning as a corrective mechanism when internal governance fails. Courts serve as active guarantors ensuring committees exercise discretion within transparent, principled boundaries. In Finland, responsibility lies with universities through institutional practices, managerial discretion, and professional norms, with limited external enforcement. This approach relies on universities embedding responsible assessment principles in organizational practices rather than assuming systematic ex post adjudication will ensure compliance.

Neither model fully resolves tensions between academic autonomy and accountability. Excessive Italian proceduralization may marginalize peer judgment and redirect conflicts toward courts, generating adversarial dynamics that consume institutional resources and potentially hinder reform. High thresholds for legal intervention in Finland, on the other hand, may leave contested evaluations largely insulated from challenge, raising questions about whether responsible assessment principles are effectively implemented without external pressure.

These findings have important implications for CoARA implementation. The agenda's emphasis on institutional autonomy, transparency, and qualitative judgment resonates differently across legal systems. In highly judicialized environments, expanding evaluative discretion without rethinking procedural safeguards may intensify litigation rather than reduce it. In systems with broad institutional autonomy and few legal remedies, the main challenge is to ensure that principles of responsible assessment are truly put into practice, rather than being assumed to operate automatically. The comparison demonstrates that responsible research assessment cannot be implemented through uniform models, but requires alignment with existing legal and institutional architectures, acknowledging that the balance between autonomy, discretion, and accountability is structurally conditioned by national legal contexts.

The study reveals courts and ombudsman institutions as significant, yet overlooked actors in shaping how research assessment principles translate into practice. Their interventions, whether through annulling procedures, demanding transparent reasoning, or publicizing procedural deficiencies, actively construct practical meanings of merit, fairness, and transparency in academic careers. Understanding these legal dimensions is essential for advancing responsible research assessment across diverse European contexts, ensuring that reform efforts account for how different legal frameworks enable or constrain the realization of shared principles.

References

Cavallo Perin, R. (2023). 'L'Università come istituzione di libera scienza'. *Diritto amministrativo*, 3, pp. 549–590.

D'Angelosante, M. (2024). 'Valutazione della ricerca scientifica in sede concorsuale o abilitativa e potere decisionale'. *Giornale di diritto amministrativo*, 2, pp. 264–271.

Estermann, T. and Nokkala, T. (2009). *University autonomy in Europe I: Exploratory study*. European University Association. Available at: https://www.eua.eu/downloads/publications/university%20autonomy%20in%20europe%201%20-%20exploratory%20study%20.pdf

European University Association (EUA) (2025). *How universities can protect and promote academic freedom: EUA principles and guidelines*. Available at: https://www.eua.eu/images/publications/Publication_PDFs/Academic_freedom_principles.pdf

Gmurzyńska, E. (2021). 'Analysis of the causes of conflicts at universities and alternative methods of resolving them. Part I: Mediation in academic disputes'. *Studia Iuridica Lublinensia*, 30(1), pp. 55–101. Available at: https://doi.org/10.17951/sil.2021.30.1.55-101

Kallio, T.J., Lehtivuori, A., Kallio, K.M. et al. (2025). 'Academic systems, career models, and experienced performance pressure: A comparative study of Sweden and Finland'. *Higher Education*. Available at: https://doi.org/10.1007/s10734-025-01528-7

Kivistö, J., Pekkola, E. and Pausits, A. (2019). 'Academic careers and promotions in Finland and Austria: System and institutional perspectives', in Mahat, M. and Tatebe, J. (eds.), *Achieving academic promotion*. Emerald Group Publishing, pp. 105–125. Available at: https://doi.org/10.1108/978-1-78756-899-020191009

Mäntysaari, P. (2022). 'The recruitment of full professors according to pre-determined criteria in four Nordic countries'. *Nordic Journal on Law and Society*, 5(1). Available at: https://doi.org/10.36368/njolas.v5i01.242

Teigen, M. (1999). 'Documenting discrimination: A study of recruitment cases brought to the Norwegian Gender Equality Ombud'. *Gender, Work & Organization*, 6, pp. 91–105. Available at: https://doi.org/10.1111/1468-0432.00073

UNDER EVALUATION PRESSURE: HOW RESEARCH EVALUATION SHAPES ACADEMICS' COMMUNITY ENGAGEMENT PRACTICES

VALENTINA CARAZZOLO, EMANUELA REALE, ANDREA ORAZIO SPINELLO
National Research Council - Research Institute on Sustainable Economic Growth (CNR- IRCrES), Italy

Keywords: Community engagement, Third mission, Academic autonomy

Background

Over the past two decades, universities have increasingly been positioned as key actors in addressing societal challenges and generating public value (Mazzucato, 2016). Alongside their traditional missions of teaching and research, higher education institutions are now expected to contribute more visibly to social, economic, and environmental objectives. This growing emphasis on societal relevance has reshaped the discourse on the role of universities, linking institutional legitimacy to their capacity to demonstrate forms of impact that extend beyond academic outputs and disciplinary boundaries (Goddard et al., 2016).

Within this context, many higher education systems have promoted the expansion of universities' activities under the umbrella of the so-called Third Mission, encompassing a broad range of initiatives aimed at strengthening relationships with external actors and producing social value (Pinheiro et al., 2015). Public engagement and, more specifically, community engagement, have become central components of this broad concept, increasingly framed as processes of knowledge co-creation, collaboration, and mutual learning between academia and society (Benneworth, 2012). Through these practices, universities seek not only to disseminate research results, but also to contribute directly to the addressing of concrete social problems and to position themselves as socially embedded institutions (Sánchez-Barrioluengo & Benneworth, 2019). At the same time, however, engagement remains characterised by substantial heterogeneity in terms of actors involved, modes of interaction, and expected outcomes, raising persistent questions about how such activities can be recognised, valued, and governed within academic systems (Grand et al., 2015). Parallel to this expansion of universities' societal roles, research and higher education have become objects of systematic and continuous evaluation. National and transnational assessment frameworks now constitute a stable feature of contemporary higher education governance, shaping funding allocation, institutional strategies, and academic careers (Whitley, 2007). A substantial body of literature has shown that evaluation does not merely measure performance but actively structures academic work by defining what counts as legitimate output (Atkinson, 2001, Frey, 2003, Frey et al, 2009), establishing incentives (Butler, 2003, Pinar, Horne, 2021), and orienting professional priorities. These dynamics have been associated with a range of effects on researchers' practices, including changes in publication strategies, time allocation, collaboration patterns, and the types of activities perceived as career-relevant (Reale, 2025).

This paper lies at the intersection of these two issues and examines how research evaluation affects academics' autonomy to initiate and sustain community engagement within their institutions. Community engagement constitutes a particularly revealing domain in this respect: although increasingly promoted at the institutional level, it remains difficult to accommodate within standardised, indicator-driven assessment frameworks and is often still evaluated in fragmented, non-standardised ways. The analysis draws on the theoretical framework developed within the PRIN2022 project PLACES, which conceptualises community engagement as grounded in reciprocity, an emancipative orientation, and a critical epistemology, and emphasises forms of participation based on involvement and co-production with external actors (Vargiu, 2014). On this basis, the paper investigates whether, and through which institutional mechanisms, research evaluation generates pressures on academics and reshapes their community engagement practices across different university contexts.

Research questions

In current national research evaluation exercises, community engagement is increasingly evaluated as part of universities' third mission, primarily to verify the extent to which these activities generate societal impact. Although engagement has thus gained formal recognition within evaluation frameworks, the indicators used to assess it remain highly heterogeneous, reflecting the diversity and complexity of engagement practices.

Against this background, this paper focuses on the effects that evaluation may produce on academic work. Among the various consequences of evaluation processes, we concentrate on a specific dimension: their impact on academics' autonomy to initiate community engagement activities. The central research question is therefore: How the evaluation of engagement, embedded within broader systems of institutional assessment, affects researchers' autonomy to define priorities, allocate time, and shape the forms and objectives of their engagement?

The paper situates these dynamics within the theoretical framework of institutional pressures (Oliver, 1991) generated by evaluation (Dahler-Larsen, 2011), to control whether evaluation reshapes engagement through organizational expectations, incentives and accountability mechanisms, thus circumscribing the autonomous initiatives of the academics.

Method

The analysis adopts a comparative multiple case study approach based on six universities in three national contexts: France, Italy, and the United Kingdom. The selected institutions are medium-to-large universities characterised by established community engagement practices, which display different degrees of institutionalisation across the three countries and reflect distinct evaluation traditions and approaches. Empirical evidence combines documentary analysis, secondary data, and qualitative interviews. Specifically, the study draws on international datasets and research infrastructures (e.g. EUROSTAT, OECD, EUA, ETER, Eurobarometer, the Academic Freedom Index), as well as on policy documents and reports produced by ministries, evaluation agencies, funding bodies, and universities, in order to contextualise national evaluation frameworks and institutional arrangements. This material is complemented by in-depth interviews with key actors at multiple levels (including national policy and evaluation bodies, intermediary organisations, and university leadership), as well as with academics involved in community engagement initiatives.

Building on a multi-level comparative case study approach (Spinello et al., 2025), the analysis examines how engagement practices are shaped by interactions between national policy environments (macro level) and institutional governance structures (meso level). The triangulation of these sources allows us to investigate how evaluation frameworks are translated into organisational expectations and how they affect academics' community engagement practices across different university contexts.

Results

The findings highlight clear cross-national differences in the degree of autonomy academics retain in initiating and conducting community engagement activities. While engagement has gained visibility within evaluation frameworks in all three contexts, the ways in which evaluation is translated into institutional expectations and incentives vary substantially, producing differentiated configurations of pressure and professional discretion. These differences shape not only whether academics engage with external communities, but also how such activities are framed, prioritised, and sustained within academic work. In the UK cases, results point to a relatively high level of soft pressure, with universities actively promoting engagement through incentives, recognition mechanisms, and strategic alignment, while formally preserving individual choice. In the French cases, academics report a high degree of autonomy: although national funding schemes have supported engagement initiatives, these activities are not subject to systematic national evaluation, and are largely undertaken on a voluntary basis, driven by personal motivations rather than assessment requirements. The Italian cases display a more ambivalent pattern. On the one hand, academics perceive increasing pressure stemming from national evaluation exercises to account for engagement activities; on the other hand, they emphasise the limited availability of institutional incentives, resources, and organisational support. Overall, the results suggest that the growing evaluative attention to community engagement does not produce uniform effects but reconfigures academic autonomy in context-specific ways.

References

Atkinson, M. (2001). 'Peer review culture'. *Science and Engineering Ethics*, 7(2), pp. 193–204. Available at: https://doi.org/10.1007/s11948-001-0040-8

Benneworth, P. (2012). 'University engagement with socially excluded communities: Towards the idea of "the engaged university"', in Benneworth, P. (ed.), *University Engagement with Socially Excluded Communities*. Springer, Dordrecht, pp. 3–31.

Butler, L. (2003). 'Explaining Australia's increased share of ISI publications — the effects of a funding formula based on publication counts'. *Research Policy*, 32(1), pp. 143–155. Available at: https://doi.org/10.1016/S0048-7333(02)00007-0

Dahler-Larsen, P. (2011). *The Evaluation Society*. Stanford University Press, Stanford.

Frey, B.S. (2003). 'Publishing as prostitution? Choosing between one's own ideas and academic success'. *Public Choice*, 116(1), pp. 205–223. Available at: https://doi.org/10.1023/A:1024208701874

Frey, B.S., Eichenberger, R. and Frey, R.L. (2009). 'Editorial ruminations: Publishing Kyklos'. *Kyklos*, 62(2), pp. 151–160. Available at: https://doi.org/10.1111/j.1467-6435.2009.00428.x

Goddard, J., Hazelkorn, E. and Vallance, P. (eds.) (2016). *The Civic University: The Policy and Leadership Challenges*. Edward Elgar Publishing, Cheltenham.

Grand, A., Davies, G., Holliman, R. and Adams, A. (2015). 'Mapping public engagement with research in a UK university'. *PLOS ONE*, 10(4), e0121874. Available at: https://doi.org/10.1371/journal.pone.0121874

Mazzucato, M. (2016). 'From market fixing to market creating: A new framework for innovation policy'. *Industry and Innovation*, 23(2), pp. 140–156. Available at: https://doi.org/10.1080/13662716.2016.1146124

Oliver, C. (1991). 'Strategic responses to institutional processes'. *Academy of Management Review*, 16(1), pp. 145–179. Available at: https://doi.org/10.2307/258610

Pinar, M. and Horne, T.J. (2022). 'Assessing research excellence: Evaluating the Research Excellence Framework'. *Research Evaluation*, 31(2), pp. 173–187. Available at: https://doi.org/10.1093/reseval/rvab042

Pinheiro, R., Langa, P.V. and Pausits, A. (2015). 'One and two equals three? The third mission of higher education institutions'. *European Journal of Higher Education*, 5(3), pp. 233–249. Available at: https://doi.org/10.1080/21568235.2015.1044552

Reale, E. (ed.) (2025). *Gli effetti della valutazione sulla ricerca accademica. Produzione della conoscenza e problemi metodologici*. FrancoAngeli, Milano.

Sánchez-Barrioluengo, M. and Benneworth, P. (2019). 'Is the entrepreneurial university also regionally engaged? Analysing the influence of university's structural configuration on third mission performance'. *Technological Forecasting and Social Change*, 141, pp. 206–218. Available at: https://doi.org/10.1016/j.techfore.2018.10.017

Spinello, A.O., Carazzolo, V., Finardi, U., Frongia, S., Ghibellini, V., Reale, E. and Vargiu, A. (2025). *Designing a multiple case study approach for studying universities' community engagement*. CNR-IRCrES Working Paper 5/2025. Istituto di Ricerca sulla Crescita Economica Sostenibile, Moncalieri. Available at: https://doi.org/10.23760/2421-7158.2025.005

Vargiu, A. (2014). 'Indicators for the evaluation of public engagement of higher education institutions'. *Journal of the Knowledge Economy*, 5(3), pp. 562–584. Available at: https://doi.org/10.1007/s13132-014-0194-7

Whitley, R. (2007). 'Changing governance of the public sciences', in Whitley, R. and Gläser, J. (eds.), *The Changing Governance of the Sciences: The Advent of Research Evaluation Systems*. Springer, Dordrecht, pp. 3–27.

STABILITY AND CHANGE: DISCIPLINARY HETEROGENEITY AND REGIONAL DYNAMICS IN THE RESEARCH EVALUATION OF THE HUMANITIES AND SOCIAL SCIENCES IN JAPAN

MASARU KARUBE, YUKI MIYAZAWA
Hitotsubashi University, Japan
YASUSHI HARA
Kobe University, Japan
LEI YU
The University of Kitakyushu, Japan

Keywords: Research evaluation, Social Sciences and Humanities, Disciplinary heterogeneity, Japan

Introduction

It is neither realistic nor appropriate to view the evaluation of research in the humanities and social sciences as equivalent to that in the natural sciences or to directly apply the evaluation criteria and systems of the natural sciences to the humanities and social sciences. As Reale et al. (2018) point out, unlike STEM research, it is difficult to establish a single, unambiguous definition of the dominant paradigm underpinning evaluation in the humanities and social sciences with respect to scientific, societal, and political impact.

Additionally, societal entities and the definition of society itself are multifaceted and not always explicitly defined when considering societal impact. Regarding political impact, it is not always easy to identify how humanities and social science research influences policy deliberation content or process. Causal relationships are nonlinear, and translating research findings into policy or institutional change takes time. Additionally, incorporating outcomes into policy documents involves issues of indirectness and translatability.

This study examines the structural stability and temporal change of research practices across disciplines in the humanities and social sciences in Japan. It focuses on how disciplinary heterogeneity and regional dynamics shape research evaluation systems and their outcomes. Integrating large-scale bibliometric data with qualitative interviews allows us to investigate variations in research outputs, evaluation criteria, and societal engagement across ten subfields. This approach highlights discipline-specific and Japan-specific characteristics that challenge conventional evaluation frameworks derived from the natural sciences.

Research Gaps

A body of research has been established that discusses the differences between natural science research and humanities/social science research. However, extant studies that have empirically demonstrated structural differences between disciplines within the humanities and social sciences are still limited. Notable exceptions include Sivertsen

& Larsen (2018), which empirically showed differences in publication formats, languages, and internationality by discipline. Hammarfelt (2018) focused on differences in evaluation metrics and publication cultures across fields. Engels, Ossenblok, & Spruyt (2012) examined temporal changes in publication cultures over time. And Ochsner, Hug, & Daniel (2016), which pointed out the lack of uniformity in research evaluation criteria.

A closer examination of these extant studies reveals two key issues that remain to be fully elucidated. First, the question must be raised of whether the structural differences between fields and their temporal changes, as pointed out in the aforementioned studies, are unique to Europe or represent characteristics observable in regions and countries outside Europe. Secondly, the extent to which the structural differences observed between fields in Europe reflect regional and historically dependent characteristics of European countries must be determined.

Sample data and method

To address these challenges, it is necessary to prepare new datasets that are distinct from those of Europe. In addition, it is essential to identify patterns of interdisciplinary research outcomes in the humanities and social sciences in regions and countries outside Europe. Furthermore, an examination of the underlying differences in research cultures is required, as well as an assessment of the homogeneity and heterogeneity across disciplines in the research evaluation systems that guide research activities.

To address this issue, a quantitative approach was employed, with two studies (Study 1 and Study 2) utilizing experimental methods, and a third study (Study 3) employing qualitative methods. These studies targeted Japan's humanities and social sciences.

Study 1 utilized the Elsevier Scopus database, while Study 2 employed Researchmap, a researcher database containing the research activities and outputs of nearly all scientists affiliated with Japanese research institutions, including humanities and social sciences researchers. The target books numbered 740,000, and the papers comprised 1.3 million articles written in Japanese and other foreign languages, including English.

For Study 3, we conducted semi-structured interviews with 135 researchers in the humanities and social sciences. These researchers were randomly sampled and agreed to participate. Responses were obtained from researchers belonging to 65 out of 69 subfields within the humanities and social sciences. The total interview duration was 130 hours, 24 minutes, and 9 seconds. The text data, amounting to approximately 2.5 million characters, was coded by four evaluators. The primary coding process yielded 2,783 items, and subsequent comparison of these codes through the code evaluator revealed the presence of approximately 200 common characteristics.

Findings

The following four points emerged from three quantitative and qualitative studies: The humanities and social sciences field, which comprises 10 distinct subfields (philosophy/ethics, history, literature, linguistics, psychology, sociology, cultural anthropology/folklore studies, law/political science, economics/business administration, education), contains 69 detailed classifications and is not monolithic.

First, the "primary research outcomes" that have been established as valid within each field vary considerably. Second, discrepancies in characteristics between fields are influenced by research objectives, methodology, and the selection of research subjects. Third, discrepancies between fields are indicative not only of variations in field-

specific characteristics but also of regional characteristics that are unique to Japan. Fourth, as structural factors explaining interdisciplinary differences, diverse influences were identified: the internationality of each field (languages used and languages of publication), path-dependent institutionalization, the growth potential of research communities, the potential for generational succession in research, and perspectives on the relationship between research and society.

Notably, this study offers unique findings by highlighting not only publications in international peer-reviewed journals and research books in native languages, but also foundational data-building activities supporting regional studies and research communities, as well as contributions to museums and performing arts.

Conclusion

To establish and operate research evaluation criteria, evaluation systems, and the supporting governance framework appropriately, it is necessary to shed further light not only on diversity across disciplines, linguistic and regional diversity, and methodological diversity, but also on the multifaceted relationship between research activities and society.

References

Engels, T.C.E., Ossenblok, T.L.B. and Spruyt, E. (2012). 'Changing publication patterns in the social sciences and humanities, 2000–2009'. *Scientometrics*, 93(2), pp. 373–390. Available at: https://doi.org/10.1007/s11192-012-0680-2

Hammarfelt, B. (2018). 'What is a discipline? The conceptualization of research areas and their bibliometric representation'. *Scientometrics*, 116(1), pp. 1–20. Available at: https://doi.org/10.1007/s11192-018-2765-z

Ochsner, M., Hug, S.E. and Daniel, H.-D. (2016). *Research Assessment in the Humanities: Towards Criteria and Procedures*. Springer, Cham. Available at: https://doi.org/10.1007/978-3-319-29016-4

Reale, E., Avramov, D., Canhial, K., Donovan, C., Flecha, R., Holm, P., Larkin, C., Lepori, B., Mosoni-Fried, J., Oliver, E., Primeri, E., Puigvert, L., Scharnhorst, A., Schubert, A., Soler, M., Soós, S., Sorde, T., Travis, C. and Van Horik, R. (2018). 'A review of literature on evaluating the scientific, social and political impact of social sciences and humanities research'. *Research Evaluation*, 27(4), pp. 298–308.

Sivertsen, G. and Larsen, B. (2018). 'Publication patterns in the social sciences and humanities: Evidence from eight European countries'. *Scientometrics*, 116(1), pp. 251–276. Available at: https://doi.org/10.1007/s11192-018-2711-0

2.4B GOVERNANCE AND POLITICS OF RESEARCH SYSTEMS

UNIVERSITY JOURNALS IN A STRATIFIED EVALUATION REGIME: AUTONOMY, EVALUATION, AND STATE CONTROL OF SCHOLARLY PUBLISHING IN UKRAINE (AS A CASE)

MARYNA NAZAROVETS
Open Science Lab, TIB Leibniz Information Centre for Science and Technology, Hannover, Germany

Keywords: University journals, Research evaluation, Stratified evaluation regime, Research governance, Ukraine

The governance of scholarly publishing

Research evaluation has become a central site of governance in contemporary science, increasingly functioning not only as a mechanism of quality assurance, but as a politically charged instrument through which states shape research behaviour, institutional priorities, and publishing infrastructures (e.g. Korytkowski and Kulczycki, 2019; Oancea, 2019). While the effects of evaluation regimes on researchers and research organisations have been widely discussed, the subject of scholarly publishing itself, and the journals as objects of regulation, has received less attention.

This presentation uses Ukraine as a focused analytical case to examine state control over scholarly publishing, as exercised through formal journal requirements. We analyse how these requirements, which are mandated by the state and directed at journals, reshape publishing practices and undermine scholarly communication, particularly within the context of institutionally embedded university publishing.

The Ukrainian journal landscape

University journals occupy a central position within many national publication systems, particularly in the social sciences and humanities. Typically embedded within academic and research departments, they form part of universities' research communication infrastructure. Ukraine's journal landscape is predominantly institution-based and non-commercial. Estimates of the number of scholarly journals vary across database sources from approximately 1,500 to 3,000. Ulrichsweb-based mapping identifies about 1,500 active titles, of which roughly 950 are university journals (Nazarovets, 2024; Nazarovets et al., 2026). Unlike journals published by large commercial publishers, university journals often have limited resources and rely on voluntary editorial labour. Their sustainability therefore depends heavily on institutional conditions and regulatory frameworks.

Stratification and delegated sovereignty

Across countries, a common technique is the curation and ranking of publication channels, which then feeds into evaluation, funding allocation, or credentialing. For instance, Norway operates a Norwegian Register for Scientific Journals, Series and Publishers (2026) and distinguishes between levels in ways that directly link publishing venues to evaluative weight. Finland's Publication Forum (2026) similarly

classifies journals and publishers into levels to support assessment and funding-related metrics. ENRESSH (2024) has documented and synthesised these "publication channel list" approaches as instruments of national research governance, explicitly recognising them as policy tools with consequences for scholarly communication. Outside Europe, China provides another well-studied example of governance through journal lists (Teixeira da Silva et al., 2024).

Ukraine provides a particularly illustrative case of such regulation. The Procedure for Forming the List of Professional Publications[1] establishes criteria that national journals must meet in order to be recognised for use in academic procedures, including doctoral certification, academic promotion, and institutional reporting (MESU, 2018). Within this framework, national journals indexed in the Web of Science Core Collection and/or Scopus[2] receive a *category A*[3], while journals that meet formal requirements are recognised as *category B*[4] (Nazarovets, 2024). Both categories are legitimate scholarly journals but carry unequal evaluative weight in academic procedures. In practice, this arrangement partially delegates journal quality recognition to commercial indexing databases whose inclusion criteria become embedded in national evaluation procedures. This effectively deprives universities of the right to independently determine the value of scientific work, creating a rigid hierarchy where non-indexed journals are automatically labelled as 'second-rate' regardless of their actual scholarly contribution.

Structural effects and institutional traps

Several structural effects follow from this regulatory design.

First, because academic certification and promotion procedures require publications in recognised journals, universities have strong incentives to establish new journals in order to accommodate publication demand. This contributes to the rapid growth of formally recognised titles. At the same time, journals outside the recognised list lose institutional relevance and gradually cease publication. Centrally organised evaluation reduces diverse scholarly practices to a limited set of measurable indicators, which can render previously legitimate forms of publishing institutionally irrelevant (Münch, 2008). The result is a restructuring of the journal landscape in which the expansion of recognised titles occurs alongside the disappearance of other scholarly venues.

Second, the system appears to generate structural fragility within university-based publishing. While empirical data on journal financing in Ukraine remain limited, no stable, centralised funding system has yet been established. Many journals operate with heterogeneous resource bases, combining partial institutional support with author fees. Meanwhile, editorial work often relies on voluntary contributions and goodwill (Zhenchenko & Dunaievska, 2025). The war intensifies these constraints. Recent policy changes have introduced a competitive funding scheme for category A open access journals. The first funding cycle supported only 19 titles (including nine university journals), providing grants of up to approximately 7,000 euros per year for two years (MESU, 2026). Given the scale of the national journal landscape, such

[1] See https://nfv.ukrintei.ua/

[2] Only about 200 Ukrainian journals are indexed in Scopus or the Web of Science Core Collection.

[3] 190 journals as of 23 December 2025, including 93 published/co-published by universities (higher education institutions).

[4] 1,473 journals as of 23 December 2025, including 1,168 published/co-published by universities (higher education institutions).

support represents a highly selective intervention rather than a systemic funding solution.

Thirdly, journals must comply with an increasingly formalised set of editorial, technical and ethical requirements introduced through the evaluation framework, which applies largely uniform standards across disciplines and institutional capacities. Such dynamics contribute to unstable publication trajectories, including irregular publication schedules and journal closures, sometimes followed by the loss of access to previously published content (Laakso, Matthias, & Jahn, 2021; Naxera, 2024).

Erosion of autonomy and the credentialing function

Finally, the MESU requirements change the function of journals from facilitating scholarly communication to producing credentials. This reorients authors' behaviour towards 'countable' outputs and encourages universities to treat journals as internal instruments embedded in reporting and incentive structures, rather than as platforms for scholarly communication (e.g. Adiprasetio, 2025; Nazarovets, 2023; Tomaselli, 2018).

A significant conflict of interest arises: as the publisher, the university strives for high status for the journal, but as an educational institution, it needs a 'convenient' channel to protect the needs of its staff, which often leads to a reduction in real requirements in order to maintain formal appearances. As editorial priorities shift from gatekeeping and community building to formal compliance, the system becomes increasingly susceptible to compliance-seeking practices, such as strategic low-substance publishing, questionable behaviour and paper mill supply, because the primary reward is procedural acceptance rather than communicative value. The result is an erosion of university autonomy in governing publishing infrastructure, as well as a weakening of the role of journals as venues for meaningful scholarly exchange.

The Ukrainian case illustrates how journal regulation embedded in academic certification can transform publishing infrastructures into administrative instruments. Reducing this effect would require limiting the role of the List of Professional Publications in promotion procedures while shifting attention from journal status toward the sustainability of publishing infrastructures. In such a configuration, universities would maintain journals primarily where they serve disciplinary communication, rather than certification demand. At the same time, maintaining transparent editorial governance and external participation would remain essential in order to mitigate risks of local capture.

References

Adiprasetio, J. (2025). 'Knowledge, numbers and scholarly integrity: The shallow research culture in Indonesian communication education'. *Journal of Organizational Ethnography*, ahead-of-print. Available at: https://doi.org/10.1108/JOE-07-2025-0097

ENRESSH (2024). *Report on Research Assessment*. Available at: https://enressh.eu/wp-content/uploads/2020/04/Recommendations-for-publication-channel-lists.pdf

Korytkowski, P. and Kulczycki, E. (2019). 'Examining how country-level science policy shapes publication patterns: The case of Poland'. *Scientometrics*, 119, pp. 1519–1543. Available at: https://doi.org/10.1007/s11192-019-03092-1

Laakso, M., Matthias, L. and Jahn, N. (2021). 'Open is not forever: A study of vanished open access journals'. *Journal of the Association for Information Science*

and Technology, 72(9), pp. 1099–1112. Available at: https://doi.org/10.1002/asi.24460

Ministry of Education and Science of Ukraine (MESU) (2018). *On the Approval of the Procedure for the Formation of the List of Professional Publications of Ukraine*. Order from 15.01.2018 № 32. Legislation of Ukraine. Available at: https://zakon.rada.gov.ua/laws/show/en/z0148-18

Ministry of Education and Science of Ukraine (MESU) (2026). *MES completed competitive selection of Ukrainian scientific journals for financial support*. 3 February. Available at: https://mon.gov.ua/news/mon-zavershylo-konkursnyi-dobir-ukrainskykh-naukovykh-vydan-dlia-nadannia-finansovoi-pidtrymky

Münch, R. (2008). 'Stratifikation durch Evaluation: Mechanismen der Konstruktion von Statushierarchien in der Forschung / Stratification by evaluation: Mechanisms of constructing status hierarchies in research'. *Zeitschrift für Soziologie*, 37(1), pp. 60–80. Available at: https://doi.org/10.1515/zfsoz-2008-0104

Naxera, V. (2024). '"Death of the journal": An editor's view'. *Learned Publishing*, 37, e1603. Available at: https://doi.org/10.1002/leap.1603

Nazarovets, M. (2024). 'Unlocking the hidden realms: Analysing the Ukrainian journal landscape with Ulrichsweb'. *Learned Publishing*, 37(3), e1605. Available at: https://doi.org/10.1002/leap.1605

Nazarovets, M., Laakso, M. and Taşkın, Z. (2026). 'University journals in the global academic publishing landscape: Mapping over 19,000 diverse titles based on Ulrichsweb data'. *Scientometrics*, 131, pp. 703–731. Available at: https://doi.org/10.1007/s11192-025-05535-4

Nazarovets, S. (2023). 'The growth of Ukrainian conference papers in computer science and engineering: Leadership or cheating?', in *27th International Conference on Science, Technology and Innovation Indicators (STI 2023)*. Available at: https://dapp.orvium.io/deposits/6437ef266705231188156d7d/view

Norwegian Register for Scientific Journals, Series and Publishers (2026). Available at: https://kanalregister.hkdir.no/en

Oancea, A. (2019). 'Research governance and the future(s) of research assessment'. *Palgrave Communications*, 5, 27. Available at: https://doi.org/10.1057/s41599-018-0213-6

Publication Forum (2026). Available at: https://julkaisufoorumi.fi/en/publication-forum

Teixeira da Silva, J.A., Nazarovets, S., Daly, T. and Kendall, G. (2024). 'The Chinese Early Warning Journal List: Strengths, weaknesses and solutions in the light of China's global scientific rise'. *The Journal of Academic Librarianship*, 50(4), 102898. Available at: https://doi.org/10.1016/j.acalib.2024.102898

Tomaselli, K.G. (2018). 'Perverse incentives and the political economy of South African academic journal publishing'. *South African Journal of Science*, 114(11/12). Available at: https://doi.org/10.17159/sajs.2018/4341

Zhenchenko, M. and Dunaievska, O. (2025). 'Exploring state and institutional support for sustainable scholarly journal publishing'. *Journal of International Studies*, 18(3), pp. 144–162. Available at: https://doi.org/10.14254/2071-8330.2025/18-3/8

STRATEGIC FUNDING PRIORITIES, SCIENTIFIC SOVEREIGNTY, AND THE POLITICAL ECONOMY OF SCIENCE: THE BRAZILIAN EXPERIENCE IN SCIENCE, TECHNOLOGY, AND INNOVATION POLICY

ALDA MARIA NAPOLITANO SANCHEZ
Federal University of ABC, Brazil; Lund University, Sweden

Keywords: Science, technology and innovation policy, Funding priorities and evaluation criteria, Mode 1/Mode 2 knowledge production, Research evaluation, Brazil, Social Sciences and Humanities

Over the past two decades, Science, Technology, and Innovation (STI) systems have increasingly been shaped by strategic funding priorities, redefining what counts as legitimate and valuable research. In this context, funding allocation mechanisms, legal frameworks, and evaluation systems have become central arenas in which scientific sovereignty, academic freedom, and development strategies are negotiated.

Brazil offers a particularly illustrative case due to the scale of its public STI system, its historical reliance on state-led scientific development, and the significant policy reorientation introduced since the early 2000s. (Ferrari 2002) Building on the consolidation of its national Science and Technology (S&T) system, initiated in the late 1940s with the creation of institutions such as CNPq, CAPES, FINEP, and FNDCT, Brazil progressively incorporated innovation as a central organizing principle of STI policy.

From this perspective, the incorporation of innovation-oriented policies into national Science, Technology, and Innovation (STI) systems, via legal frameworks, funding instruments, and evaluation regimes, is not a neutral administrative process. Instead, it functions as a strategic governance mechanism through which states actively redefine research priorities and reshape the boundaries between scientific autonomy, societal demand, and economic competitiveness.

This study is theoretically grounded in the distinction between Mode 1 and Mode 2 knowledge production, originally formulated by Gibbons et al. (1994), and later expanded in debates on the governance, evaluation, and political economy of science. Mode 1 refers to disciplinary, academically driven research, organized within institutional settings such as universities and public research institutes, and primarily evaluated through peer review and scientific merit. Mode 2, by contrast, is characterized by transdisciplinarity, contextualization, problem-orientation, and closer interaction with social and economic actors, particularly firms and government agencies. Knowledge production in Mode 2 is often evaluated according to criteria of relevance, applicability, and impact, rather than exclusively by academic standards.

In the Brazilian case, the expansion of innovation-centered legal frameworks since the early 2000s, most notably the Innovation Law (Law nº 10,973/2004), the “Lei do Bem” (Law nº 11,196/2005), and the consolidation of Sectoral Funds, can be interpreted as an institutional attempt to promote Mode 2 dynamics. These reforms

sought to redirect public investment toward applied research, firm-level innovation and projects connected to productive structures.

Although innovation-oriented policies are formally neutral across disciplines, their evaluation criteria and funding logics disproportionately affect the Humanities and Social Sciences. As funding instruments increasingly prioritize technological risk, market potential, and firm participation, research in fields such as political science, whose societal contributions are primarily analytical, critical, and policy-oriented, faces structural disadvantages in access to resources and recognition. In this sense, funding allocation operates as an implicit evaluation system, reshaping what counts as valuable knowledge and reinforcing tensions between academic autonomy and strategic state priorities.

This paper addresses the following research question: How have Brazil's innovation-oriented legal reforms, funding instruments, and evaluation frameworks shaped the allocation and valuation of research across academic (Mode 1), applied, and firm-oriented (Mode 2) activities, particularly regarding the position of the Social Sciences and Humanities? More specifically, it examines whether innovation-driven STI policies, largely oriented toward applied research and technological development, have contributed to relegating disciplinary research in fields such as Political Science to projects perceived as less directly aligned with national science and technology priorities.

The analysis further explores how these dynamics affect research evaluation criteria, funding stability, and academic autonomy, as well as their implications for international collaboration and long-term knowledge production. By doing so, the paper situates the Brazilian case within broader global debates on scientific sovereignty, the governance and evaluation of research, and the political economy of STI systems.

Preliminary evidence reveals a highly unstable funding trajectory, with sustained growth between 2005 and 2013, when GERD approached 1.3% of GDP, followed by sharp contraction after 2015, coinciding with fiscal austerity and political instability. (Agion & Howitt 1996) This downturn disproportionately affected public research institutions and funding agencies central to academic science, such as CAPES, CNPq, and FINEP, leading to declines in scholarship availability, research continuity, and international collaboration. Although funding levels began to recover after 2019 and more markedly after 2022, the recovery has been uneven across instruments and regions. (Piketty 2014; Arbix & Miranda 2017)

The analysis further shows that innovation-oriented legal reforms did not produce a linear or consistent shift toward private-sector-led R&D. While business-oriented instruments expanded in certain periods, non-reimbursable funding to public Scientific, Technological and Innovation Institutions (STI) remained central to Brazil's STI system, particularly in socially oriented and infrastructure-intensive projects. (Figueiredo 2023) This persistence reflects both the structural characteristics of Brazilian capitalism and the enduring role of the state in sustaining scientific capacity in a context marked by high income inequality and limited private R&D investment.

However, this study adopts a more nuanced perspective that emphasizes coexistence, hybridity, and tension between modes of knowledge production. In economies such as Brazil, where private R&D capacity remains structurally limited, public research institutions play a central role in sustaining scientific activity. As a result, public research organizations and universities continue to absorb the majority of STI funding and to anchor national scientific capabilities.

As strategic priorities gain prominence, funding agencies such as FINEP become key intermediaries in negotiating what types of knowledge are deemed legitimate, fundable, and valuable. Funding and implementation systems thus operate as arenas where Mode 1 and Mode 2 logics intersect and compete, balancing criteria such as scientific excellence, technological risk, market opportunity, and social relevance.

By analyzing longitudinal funding patterns alongside legal and regulatory changes, this paper examines how innovation-oriented policies reshape the internal composition of Brazil's STI system and influence evaluation practices. Using a qualitative and descriptive document analysis of innovation-oriented legal frameworks and public funding instruments (FINEP funding lines) the study compares the prioritization of company-led, problem-solving, and applied research projects with funding opportunities for critical, theory-driven, and knowledge-oriented research, more typical of the Social Sciences and Humanities. Rather than assuming a simple replacement of basic science by applied research, the paper shows how shifts in funding priorities reflect deeply political choices that affect academic autonomy, societal relevance, and the relative valuation of different forms of knowledge within national development strategies. (Pinho & Gomes 2017)

Ultimately, the Brazilian experience illustrates how the governance of science within national development aspirations produces a layered STI system, in which Mode 1 and Mode 2 coexist under unequal conditions. These asymmetries are not only expressed in funding volumes but also embedded in evaluation criteria that tend to privilege applied, problem-solving, and market-oriented research over critical, theory-driven knowledge production, particularly in the Social Sciences and Humanities. Understanding these dynamics is essential for assessing how strategic funding priorities reshape scientific sovereignty, academic freedom, and the long-term capacity of STI systems to support inclusive and reflexive socioeconomic development.

References

Aghion, P. and Howitt, P. (1996). 'Research and development in the growth process'. *Journal of Economic Growth*, 1, pp. 49–73. Available at: https://doi.org/10.1007/BF00163342

Araújo, A.A. de (2012). 'Desafios da Finep e o fomento à inovação'. *Revista USP*, 93, pp. 113–126. Available at: https://doi.org/10.11606/issn.2316-9036.v0i93p113-126

Arbix, G. and Miranda, Z. (2017). 'Políticas de inovação em nova chave'. *Estudos Avançados*, 31(90), pp. 49–73. Available at: https://doi.org/10.1590/s0103-40142017.3190004

Ferrari, A.F. (2002). 'O Fundo Nacional de Desenvolvimento Científico e Tecnológico — FNDCT e a Financiadora de Estudos e Projetos — FINEP'. *Revista Brasileira de Inovação*, 1(1), p. 151. Available at: https://doi.org/10.20396/rbi.v1i1.8648855

Figueiredo, P.N. (2023). 'Estratégia nacional de inovação: Uma breve contribuição para sua efetividade sob a perspectiva de acumulação de capacidade tecnológica'. *Revista de Administração Pública*, 57(5), e2022-0418. Available at: https://doi.org/10.1590/0034-761220220418

Gibbons, M. et al. (1994). *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*. Sage, London.

Grabowski, R. (2017). 'Premature deindustrialization and inequality'. *International Journal of Social Economics*, 44(12), pp. 1, 638–1,654. Available at: https://doi.org/10.1108/IJSE-07-2015-0197

Piketty, T. (2014). *Capital in the Twenty-First Century*. Harvard University Press, Cambridge.

Pinho, G.A. de and Gomes, R. (2017). 'Análise da evolução dos editais Finep de subvenção econômica à inovação'. *Blucher Engineering Proceedings*, setembro, pp. 527–544. Available at: https://doi.org/10.5151/enei2017-30

Rodrik, D. (2016). 'Premature deindustrialization'. *Journal of Economic Growth*, 21(1), pp. 1–33. Available at: https://doi.org/10.1007/s10887-015-9122-3

FROM SDGS TO ISO STANDARDS: AN OPPORTUNITY MORE THAN A CONSTRAINT

EMMANUELLE PENSEC
ARIANE, France
GEOFFREY WILLIAMS
Université Grenoble-Alpes, France

Keywords: ISO standards, Societal impact, Quality management, Responsible research and innovation

This paper poses a very simple question, how can humanities researchers be helped to obtain better funding at both national and European levels? The solution proposed is through dedicated training, and through the application of ISO standards, so as to create a framework to analyse proposals prior to submission. In other words, we apply the same techniques as used in companies, which cannot afford to lose markets. It is highly pragmatic approach that has been tested over a number of years and has proved successful in both gaining awareness from participants and from their success in project proposals. This is not research per se but shows how an evaluation procedure can solve a genuine problem with a pragmatic solution. Initial actions used Sustainable Development Goals (SDGs) (United Nations, 2017) as a means of highlighting societal impact. What we do here is to go to beyond that in order to create a framework using two standards, ISO 9001 (International Standards Organisation, 2015), which is concerned with quality, and ISO 26000 (International Standards Organisation, 2010) which is concerned with social aspects of management. The latter standard is linked to the application of SDGs in companies, and institutions.

In practical terms, a response to the above question is to develop a paradigm for responsible research based on the two ISO standards mentioned. We shall illustrate this through research and training carried out with the ARIANE digital humanities consortium in France (Pensec & Williams, 2024a, 2024d), and the CEHUM research centre for the humanities at the university of Minho, Portugal (Pensec & Williams, 2024b, 2024c, 2025).

In analysing the practice in both centres, it is clear how various factors raise preoccupations concerning a recognition of societal issues in research policy. It is equally as clear that current responses from governance tend to be top down. However, if we look at the principles of responsible research, such as ethics, societal impact, transparency and stakeholders' expectations, it is clear that it is necessary to elaborate means so that researchers and their research units can be accompanied in tackling these questions themselves. This means widening the research parameters so as to taking into account how and for whom we carry out research, with what impact and with what governance. All of these aspects are clearly implied in both cited ISO standards.

Now SDGs and ISO Standards are obviously two very different things. However, they can be linked thereby improving the valorisation of SSH research, without falling into the trap of the SSH being reduced to simply an add-on to STEM. The two research

paradigms studied are obviously very different, but the solutions are similar. Both require training through collective intelligence, and as we had found both require the use of an analytical framework that can be applied prior to submitting project applications.

In building an analytical framework, the challenge is both to develop and apply a tool that is acceptable to researchers who are reticent about any evaluative activity and to encourage the appropriation of this new tool by SSH researchers themselves. To do this, we must first, build a tool that is usable by consultants and then, adapt it to the specific needs of research communities. It is important to not even mention ISO standards at the stage, because too many humanities researchers will consider ISO standards are far to linked to business environments.

The issues raised by ISO 9001 concern research quality, which is again something which is controversial. However, it can be used to demonstrate all the aspects of scientific rigour that are inherent in FAIR. It also means highlighting management skills concerning organisational reliability, such as team building and management, as well as risk management. Social impact is clearly found in ISO 26000 and its use enables researchers to adapt their research proposal to the key questions that financers are going to ask. These include social relevance, ethical acceptability and stakeholder dialogue, all three of which must be clearly illustrated if they are taken on board. This first step has been done within the ResQ framework, which we have adapted as ResQ tools in order to refine it in collaboration with researchers.

Research evaluation is a highly controversial subject with many SSH researchers, and this is very apparent in feedback we have received during training sessions. It is therefore very important to demonstrate that our aim is not to adapt a project to societal issues, but to demonstrate how their research paradigms operate within an evaluation environment and how they can be used to valorise SSH research.

References

International Standards Organisation (2010). *ISO 26000: Guidance on Social Responsibility*. ISO.

International Standards Organisation (2015). *ISO 9001: Quality management systems — Requirements*. ISO.

Pensec, E. and Williams, G. (2024a). *Analyser pour gagner: Comment argumenter votre dépôt de projet ANR pour augmenter vos chances d'obtenir un financement*. HAL-SHS.

Pensec, E. and Williams, G. (2024b). *Beyond citizen science: SDGs as a communication tool*.

Pensec, E. and Williams, G. (2024c). *Getting funded: Strategies to build relevant and attractive projects*.

Pensec, E. and Williams, G. (2024d). *Piloter efficacement un projet*. HAL-SHS. Available at: https://doi.org/10.13140/RG.2.2.27422.75845

Pensec, E. and Williams, G. (2025). *Getting Funded*.

United Nations (2017). *Sustainable Development Goals*. Available at: https://unstats.un.org/sdgs

LONELINESS AS AN EVALUATIVE CHALLENGE FOR PUBLIC INSTITUTIONS: TOWARDS A FRAMEWORK FOR RECOGNISING CO-CREATED RESEARCH CONTRIBUTIONS

SARA MARTÍNEZ-CARDAMA

Department of Library and Information Science/Agustín Millares Institute, Universidad Carlos III de Madrid (UC3M), Spain

NÚRIA BAUTISTA-PUIG

Institute of Philosophy, Spanish National Research Council (CSIC), Spain

LISANDRA OTERO BORGES

Department of Library and Information Science/Agustín Millares Institute, Universidad Carlos III de Madrid (UC3M), Spain

Keywords: Unwanted loneliness, Social inclusion, Public libraries, Community spaces, Research evaluation

Introduction

Unwanted loneliness has emerged as a major social concern across Europe, particularly among young people, with significant consequences for mental health, social participation, and community cohesion. It is commonly defined as the distressing experience that occurs when an individual's social relationships are perceived as insufficient in quantity or quality (Perlman & Peplau, 1981). While early research primarily framed loneliness as an individual and psychological condition, recent studies have demonstrated that unwanted loneliness is deeply embedded in broader structural dynamics, including socioeconomic inequalities, digitally mediated social environments, and changing forms of social interaction (Casal-Rodríguez et al., 2023; Ruiz-Villafranca et al., 2024). Building on this shift, contemporary research has further conceptualized loneliness as a socioecological phenomenon shaped by neighborhood environments, access to social infrastructure, and unequal opportunities for social connection, thereby linking individual experiences of loneliness to the spatial, institutional, and relational conditions in which they are produced (Litt et al., 2024). From this perspective, interventions addressing loneliness increasingly emphasize community-level and infrastructural approaches, recognizing that sustainable impacts depend on relational processes such as participation, trust-building, and the availability of inclusive shared spaces (Sones et al., 2022; Department for Digital, Culture, Media & Sport, UK Government, 2023).

In Spain, the social and economic costs associated with unwanted loneliness are increasingly visible, prompting public administrations to develop targeted policies and intervention programmes. At the same time, cultural and community institutions (particularly public libraries) have gained prominence as spaces that foster social connection, participation, and inclusion. Libraries are increasingly conceptualized as social infrastructures: places where everyday interactions contribute to social resilience, trust, and collective well-being (Klinenberg, 2018).

Research on community building in libraries and cultural institutions shows that participatory practices play a key role in strengthening social relationships, increasing visibility, and fostering a shared sense of ownership over these spaces (Hernández-Pérez et al., 2020). Through these everyday interactions and connections, libraries become part of the social and infrastructural environments in which loneliness can both emerge and be addressed, aligning with social connectedness perspectives (Sones et al., 2024)

Co-creation is increasingly understood as a structured and iterative approach in which researchers, practitioners, and community members collaboratively define problems, design interventions, and reflect on outcomes. Recent co-created projects explicitly addressing loneliness demonstrate how such approaches generate locally grounded, prevention-oriented responses while strengthening community capacity and policy relevance (Corsini et al., 2024). Evidence from evaluations of loneliness interventions consistently shows that relational and process-based outcomes (such as empowerment, changes in participation, and local networks) are central to addressing loneliness, even though these effects often develop over time and are difficult to capture through conventional evaluation frameworks (DCMS, 2023). While such approaches align closely with recent calls for socially relevant, inclusive, and engaged research, they raise important questions for research evaluation. This contribution examines unwanted loneliness as a case study to reflect on the limits of conventional research evaluation and the need for more inclusive and context-sensitive assessment approaches. Drawing on empirical material from the SOLIBCM project, the paper explores how co-created research practices generate forms of value that are largely invisible within prevailing evaluation systems and discusses how these contributions could be more adequately recognized within responsible assessment frameworks.

Conceptual background

Unwanted loneliness has been linked to reduced civic engagement, weakened social ties, and diminished trust in institutions, particularly among younger populations navigating hybrid online vs offline social environments (Cuccu & Stepanova, 2021). Recent research highlights the role of digital practices, fear of missing out (FOMO), and algorithmically mediated sociality in reshaping experiences of belonging and exclusion (Ibáñez-Marco & Martínez-Cardama, 2024; Santaolalla-Rueda, 2025).

Research on social connectedness and healthy cities has emphasized the role of neighbourhood environments and shared community spaces as key mediators of social relationships (Sones et al., 2022). Public libraries, alongside other forms of social infrastructure, have been reconceptualized as active social agents rather than neutral service providers. From a Habermasian perspective, they function as spaces that support dialogue, visibility of social issues, and collective sense-making (Larsen, 2020). Empirical studies increasingly document their role in promoting social inclusion, mental well-being, and public health (Lenstra & Peritore, 2024; Dolley, 2025). These developments challenge dominant understandings of research quality and impact. Research embedded in social infrastructures often prioritizes relational work, trust-building, and reflexivity over scalable outputs. As Martínez-Cardama (2025) argues, projects addressing unwanted loneliness through libraries tend to produce “soft” outcomes (such as empowerment, visibility of lived experiences, and strengthened community ties) that resist quantification but are central to their societal value.

Data and methods

This study draws on qualitative and documentary analysis of the SOLIBCM project, a research initiative focused on unwanted loneliness among young people and the role of public libraries as spaces for social connection and co-creation. The empirical material includes project documentation, methodological reports, participatory methodologies, and reflective outputs produced during the research process (see more details in Martínez-Cardama, 2025). The project adopted a co-creative research design in which young participants, librarians, educators, and researchers collaboratively defined the research focus, identified relevant dimensions of loneliness, and co-designed activities aimed at fostering social interaction. Rather than treating participants as data sources, the project emphasized mutual learning, reflexivity, and shared ownership of both the process and the outcomes.

Preliminary findings

Preliminary analysis from the ongoing analysis reveals a clear asymmetry between the richness of the research process and its representation through conventional evaluative indicators. While the project has so far generated a limited number of formal academic outputs, it is producing a wide range of non-traditional results, including co-created narratives, participatory tools, local networks of collaboration, and enhanced visibility of young people's experiences of loneliness.

Moreover, the project highlights how impact in co-created research unfolds over time and across settings. Changes in participants' sense of belonging, confidence, and engagement often emerged after the formal end of project activities, making them difficult to capture through short-term impact indicators. Such temporal and relational dimensions of impact remain largely invisible within prevailing assessment frameworks. Building on these observations, the study develops an analytical framework that makes visible these otherwise overlooked contributions, by systematically capturing relational, processual, and community-level forms of value generated through co-created research.

Implications for research evaluation

The case of unwanted loneliness research through libraries illustrates several limitations of dominant research evaluation models. First, it demonstrates that societally embedded research cannot be adequately assessed through output-centric metrics alone.

Second, the findings underscore the need for evaluation approaches that recognize participation and co-creation as constitutive elements of research, rather than isolated activities. In projects such as SOLIBCM, engagement is not an add-on but a core epistemic condition for producing relevant knowledge about lived social experiences.

Finally, this case points to the importance of narrative and qualitative assessment tools that allow researchers to articulate contributions that exceed conventional indicators. Such approaches align with broader calls to reform research assessment in ways that promote inclusivity, diversity of contributions, and societal relevance.

Conclusions

By examining unwanted loneliness as a case of co-created, societally embedded research, this study contributes to ongoing debates on the reform of research evaluation. It shows that addressing complex social issues through participatory approaches generates forms of value that are poorly captured by dominant assessment systems, yet are central to public trust, social cohesion, and research legitimacy.

The analysis supports the argument that responsible research assessment must move beyond narrow output-based indicators and develop frameworks capable of recognizing community-level contributions.

References

Casal-Rodríguez, B., Rivera-Castiñeira, B. and Rodríguez-Míguez, E. (2023) *El coste de la soledad no deseada en España*. Observatorio Estatal de la Soledad No Deseada (SoledadES), Fundación ONCE and Nextdoor. Available at: https://www.soledades.es/estudios/el-coste-de-la-soledad-no-deseada-en-espana (Accessed: 16 April 2026).

Corsini, N. *et al.* (2024) 'A project co-created with the community to mitigate loneliness in midlife women', *Frontiers in Public Health*, 12, 1425641.

Cuccu, L. and Stepanova, E. (2021) *Loneliness & social and civic behaviours* (EUR 30929 EN). Luxembourg: Publications Office of the European Union. Available at: https://doi.org/10.2760/802268

Dolley, J. (2025) 'Libraries: activated public spaces reducing social isolation', in Dolley, J., Hardy, K. and Matthews, T. (eds.) *Public space and the sustainable development goals: inclusion, safety, culture and nature*. Cheltenham: Edward Elgar Publishing, pp. 8–18. Available at: https://doi.org/10.4337/9781035322411

Hernández-Pérez, O., Vilariño, F. and Domènech, M. (2022) 'Public libraries engaging communities through technology and innovation: insights from the library living lab', *Public Library Quarterly*, 41(1), pp. 17–42.

Ibáñez-Marco, M. and Martínez-Cardama, S. (2024) 'Fear of missing out (FOMO) en estudiantes universitarios: desafíos informativos e implicaciones para la alfabetización digital', *Revista General de Información y Documentación*, 34(1), pp. 85–97. Available at: https://doi.org/10.5209/rgid.93742

Klinenberg, E. (2018) *Palaces for the people: how social infrastructure can help fight inequality, polarization, and the decline of civic life*. New York: Crown.

Larsen, H. (2020) 'The public sphere and Habermas: reflections on the current state of theory in public library research', *Journal of Documentation*, 77(1), pp. 251–258. Available at: https://doi.org/10.1108/JD-05-2020-0075

Leask, C.F. *et al.* (2019) 'Framework, principles and recommendations for utilising participatory methodologies in the co-creation and evaluation of public health interventions', *Research Involvement and Engagement*, 5(1), p. 2.

Lenstra, N. and Peritore, N. (2024) 'Public library partnerships for public health: health in all policies (HiAP) as a new conceptual framework for LIS teaching and research', *Journal of Education for Library and Information Science*, 65(2), pp. 216–225. Available at: https://doi.org/10.3138/jelis-2023-0003

Litt, J.S. *et al.* (2024) 'Nature-based social interventions for people experiencing loneliness: the rationale and overview of the RECETAS project', *Cities & Health*, 8(3), pp. 418–431.

Martínez-Cardama, S. (2025) 'Bibliotecas públicas frente a la soledad no deseada: reflexiones iniciales desde el proyecto SOLIBCM', *Anuario ThinkEPI*, 19. Available at: https://doi.org/10.3145/thinkepi.2025.e19a15

Perlman, D. and Peplau, L. (1981) 'Toward a social psychology of loneliness', in *Personal relationships*, 3, pp. 31–56.

Ruiz Villafranca, R. *et al.* (2024) *Estudio sobre juventud y soledad no deseada en España*. Observatorio Estatal de la Soledad No Deseada (SoledadES), Fundación ONCE.

Santaolalla-Rueda, P. (2025) 'Alone and unaware: an ethnography of unidentified and unacknowledged youth loneliness', *Ethnography* (online first). Available at: https://doi.org/10.1177/14661381251335848

Sones, M., Firth, C.L., Fuller, D., Holden, M., Kestens, Y. and Winters, M. (2022) 'Situating social connectedness in healthy cities: a conceptual primer for research and policy', *Cities & Health*, 6(6), pp. 1179–1192. Available at: https://doi.org/10.1080/23748834.2021.1926657

Sones, M., Holden, M., Kestens, Y., King, A.C., Rennie, M. and Winters, M. (2024) '(Dis)connected by design? Using participatory citizen science to uncover environmental determinants of social connectedness for youth in under-resourced neighbourhoods', *BMC Public Health*, 24(1), p. 3104.

UK Government, Department for Digital, Culture, Media & Sport (DCMS) (2023) *Evaluating interventions that tackle loneliness*. Available at: https://www.gov.uk/government/publications/exploring-interventions-to-tackle-loneliness/evaluation-of-interventions-to-tackle-loneliness (Accessed: 16 April 2026).

2.5 SOCIETAL IMPACT AND PUBLIC ENGAGEMENT

ADVANCING RESEARCH ON THE KNOWLEDGE EXCHANGE OF THE SSH WITH SOCIETY

GUNNAR SIVERTSEN
Nordic Institute for Studies in Innovation, Research and Education, Norway
NATAŠA JERMEN
The Miroslav Krleza Institute of Lexicography, Croatia
ELEA GIMÉNEZ TOLEDO
Centre for Human and Social Sciences (CCHS) of the Spanish National Research Council (CSIC), Spain

Keywords: Societal impact, Trusted media, Professional communication, Non-scholarly publications, Social Sciences and Humanities

Recognizing the limitations of international indexing services such as Web of Science and Scopus in representing scholarly publishing in the social sciences and humanities, the research community contributing to the RESSH conferences has investigated the more comprehensive coverage of such publications in national research information systems and repositories (Sile et al., 2018). In turn, these more comprehensive data sources could be used for cross-country comparisons of the publication patterns in the social sciences and humanities (Kulczycki et al, 2018), also with a particular focus on scholarly book publishing (Engels et al., 2018; Giménez-Toledo et al., 2019) and on multilingualism (Sivertsen, 2018; Kulczycki et al, 2020).

The attention to the specific scholarly publication patterns of the social sciences and humanities has also led to renewed interest in the "non-scholarly" publications for the knowledge exchange between academia and society (Giménez-Toledo et al., 2022). Focusing mainly on book publishing in the SSH, we found that books are also important within the wider range of written genres used by the SSH for another purpose, to be in direct interaction with professions, policymakers, or the general public. For the interaction with society, some specific genres are used that are influenced by disciplinary standards, such as articles in encyclopedias, commentaries to artistic works, historical exhibitions or classical texts, handbooks for target professions, reference books for public management and legal procedures, books contributing with evidence to topics of political or cultural concern, and textbooks for education at all levels. In addition, scholars in the SSH are highly visible in the literatures of enlightenment presented in the non-fiction departments of general bookstores and in public libraries. SSH scholars also often appear in interviews or with contributions to the media used by the general public. These contributions are not defined by disciplinary standards but by media professionals such as journalists (Sivertsen, 2022). These contributions should not be viewed as knowledge transfer only. It is about enlightening and participating in democratic discourse.

The purpose of our contribution to RESSH 2026 is to raise a discussion at the conference and in further research about the wide range of literature, as described above, that appears in external interactions with society. We suggest focusing on this

other context while reiterating two of the research questions that were originally raised by the RESSH community in studying the scholarly literatures of the SSH:

1. Where do we find documentation of the literature outside of the main international commercial databases of metadata? To what degree do the national databases we found to cover SSH scholarly publications more completely also cover non-scholarly literature systematically? Have other data sources (national and international) appeared in the meantime that may have relevance for this new focus?
2. How are the literatures of societal interaction treated in institutional or national research assessment systems and in official statistics?

We regard the two questions as having gained new relevance with the movement towards reform of research assessment, as expressed by CoARA (www.coara.org/agreement/the-agreement-full-text), with the agreement's focus on a broader range of capacities and achievements than those expressed in scholarly publications.

We also find reasons to ask two new research questions that are particularly adequate for the domain of societal interaction and for recent developments in this area:

3. What are the specific characteristics of the different genres of societal interactions, and how are they developing over time when responding to new societal needs and conditions, political changes, and technological developments?
4. What is the role of specific channels for communication with and within society as the question of trust in academic knowledge is challenged by the communication in some of these channels?

The third question is obviously related to new pressures on academic freedom, general freedom of expression, and the digitalization of social communication. The fourth question is related to the following observation.

Academia traditionally trusts mainstream edited media such as respected book publishers, professional journals, widely read newspapers, and public broadcasting organizations for knowledge exchange with society. The advent of online social media as sources of news and knowledge has gradually changed mass communication, raising new discussions about the role of responsible editing, credible journalism, and balanced factual information (Kohring, & Matthes, 2007; Thorson et al, 2010; Salaudeen & Onyechi, 2020). This situation creates new challenges to the knowledge exchange between academia and society (Anderson et al., 2021) and an increased interest within academia in identifying and using trusted media for communication.

In the response to these challenges, we suggest taking a new perspective on the data sources that also include 'non-scholarly' publications and to build indicators for overviewing and stimulating knowledge exchange through trusted media. The traditional focus on the quality of publication channels *within* scholarly publishing can be renewed to cover societal interaction as well. The volume of knowledge exchange through specific publication channels within trusted media can be measured and compared among research organizations within and across countries. Similar procedures as are now used to promote scholarly publishing in valuable channels in collaboration with the research communities (Pölönen et al., 2020) can also be used for assessing *societal* publication channels.

Our aim for the RESSH conference is to present the ideas and to stimulate responses. We also hope to enlighten research question 1 in dialogue with experts from different

countries. As a case study for our own further research, we will focus on the organizations and publication channels within the European Encyclopedia Network (encyclopedianetwork.eu/). They represent the co-ordinated efforts, so far not recognized in official statistics, within some countries to provide knowledge exchange between academia and society through well-organized online encyclopaedias where researchers are asked to write and update research-based articles in the spirit of the enlightenment. Our data from Norway shows that, unexpectedly, the *Great Norwegian Encyclopedia*, which is online with open access (snl.no), publishes more articles from academia than any of the most read national daily newspapers.

References

Anderson, J.T.L., Howell, E., Xenos, M.A., Scheufele, D.A. and Brossard, D. (2021). 'Learning without seeking? Incidental exposure to science news on social media and knowledge of gene editing'. *Journal of Science Communication*, 20(4), A01.

Engels, T.C.E., Starcic, A.I., Kulczycki, E., Pölönen, J. and Sivertsen, G. (2018). 'Are book publications disappearing from scholarly communication in the social sciences and humanities?'. *Aslib Journal of Information Management*, 70(6), pp. 592–607. Available at: https://doi.org/10.1108/AJIM-05-2018-0127

Giménez-Toledo, E., Jermen, N. and Sivertsen, G. (2022). 'Towards proper evaluation of book publishing in the social sciences', in Engels, T.C.E. and Kulczycki, E. (eds.), *Handbook on Research Assessment in the Social Sciences*. Edward Elgar Publishing, Cheltenham, pp. 295–314.

Giménez-Toledo, E., Mañana-Rodríguez, J., Engels, T.C.E., Guns, R., Kulczycki, E., Ochsner, M., Pölönen, J., Sivertsen, G. and Zuccala, A.A. (2019). 'Taking scholarly books into account, part II: A comparison of 19 European countries in evaluation and funding'. *Scientometrics*, 118(1), pp. 233–251. Available at: https://doi.org/10.1007/s11192-018-2956-7

Kohring, M. and Matthes, J. (2007). 'Trust in news media: Development and validation of a multidimensional scale'. *Communication Research*, 34(2), pp. 231–252.

Kulczycki, E., Engels, T.C.E., Pölönen, J., Bruun, K., Duskova, M., Guns, R., Nowotniak, R., Petr, M., Sivertsen, G., Starcic, A.I. and Zuccala, A. (2018). 'Publication patterns in the social sciences and humanities: Evidence from eight European countries'. *Scientometrics*, 116(1), pp. 463–486.

Kulczycki, E., Guns, R., Pölönen, J., Engels, T.C.E., Rozkosz, E.A., Zuccala, A.A., Bruun, K., Eskola, O., Starčič, A.I., Petr, M. and Sivertsen, G. (2020). 'Multilingual publishing in the social sciences and humanities: A seven-country European study'. *Journal of the Association for Information Science and Technology*, 71(11), pp. 1371–1385. Available at: https://doi.org/10.1002/asi.24336

Pölönen, J., Guns, R., Kulczycki, E., Sivertsen, G. and Engels, T.C.E. (2020). 'National lists of scholarly publication channels: An overview and recommendations for their construction and maintenance'. *Journal of Data and Information Science*.

Salaudeen, M.A. and Onyechi, N. (2020). 'Digital media vs mainstream media: Exploring the influences of media exposure and information preference as correlates of media credibility'. *Cogent Arts & Humanities*, 7, 1837461.

Sīle, L., Pölönen, J., Sivertsen, G., Guns, R., Engels, T.C.E., Arefiev, P., Dušková, M., Faurbæk, L., Holl, A., Kulczycki, E., Macan, B., Nelhans, G., Petr, M., Pisk, M., Soós, S., Stojanovska, J., Stone, A., Šušol, J. and Teitelbaum, R. (2018). 'Comprehensiveness of national bibliographic databases for social sciences and humanities: Findings from a European survey'. *Research Evaluation*, 27(4), pp. 310–322.

Sivertsen, G. (2018). 'Balanced multilingualism in science'. *BiD: Textos Universitaris de Biblioteconomia i Documentació*, 40. Available at: http://dx.doi.org/10.1344/BiD2018.40.25

Sivertsen, G. (2022). 'Publishing in the social sciences and its representation in research evaluation and funding systems', in Engels, T.C.E. and Kulczycki, E. (eds.), *Handbook on Research Assessment in the Social Sciences*. Edward Elgar Publishing, Cheltenham, pp. 238–261.

Thorson, K., Vraga, E. and Ekdale, B. (2010). 'Credibility in context: How uncivil online commentary affects news credibility'. *Mass Communication and Society*, 13(3), pp. 289–313.

THE FUTURE OF RESEARCH SOCIETAL IMPACT IN UNIVERSITIES AND PUBLIC RESEARCH ORGANIZATIONS: WHAT CHANGES WE NEED?

EMANUELA REALE
National Research Council - Research Institute on Sustainable Economic Growth (CNR- IRCrES), Italy

Keywords: Research societal impact, Research funding, Impact evaluation

The paper wants to open a broad discussion on what the future of evaluating research impact could be in the forthcoming decade, starting from the contents of the paper "Transformative Research Assessment – Integrating Societal Impacts into Evaluation Framework" produced by the CoARA Working Group on Societal impact in November 2025, whose findings are based on a large body of literature on impact evaluation.

The proposition the paper wants to discuss is that it is not possible to fully understand and elaborate on how evaluation should evolve to assess societal impact if we don't take a holistic perspective on changes that should affect three main R&D policy pillars, namely funding policies, evaluation methods and instruments, academic research culture of different epistemic communities. The main points to raise are as follows.

Change the logic: from generating impact to enabling impact

It is largely acknowledged that it's time to integrate impact into evaluation framework and to this aim we must move from detecting actual impact to settle conditions that can help/enable impact. It needs three combined actions on the structure of funding instruments, evaluation tools and methods, and cultural changes of institutions and academics.

Build new policies for research funding

One of the most important points is to be moving away from performance-based funding of the core/basic funding for universities and research that was largely ineffective both for quality and for impact. How can we deal with this important shift? Does impact under this logic produce a further compression of the autonomy of research organization and a challenge for academic freedom?

A second issue is how to re-design, in different national and institutional contexts, competitive project funding, putting a new attention on the balance between schemes of individual grants (especially for young scholars) and efforts for oriented R&D programs, and the different types of impact one might expect from the different types of funding schemes?

One further point on funding is that working for innovation in the design of funding mix – impact needs follow-up funding because its timeline goes beyond any project duration. We must supply evidence about this important element that cannot be ignored if we are concerned about the generation of impact.

Finally, it is important to recognize different possibilities and configuration of impact depending on the type of funding (embedding social impact in funding logics must be different between instruments' aim and scope).

Evaluation must strengthen theoretical foundation to assess impact

Engaging in research and its impact means increasing awareness of the inherent uncertainty in predicting the value of the results any scientific effort might achieve and the impact it might have on society. Uncertainty is a dimension that must enter in the consideration of how far we can detect a causal attribution of effects to research in different epistemic domains.

This new awareness implies first the need to design evaluation from the very beginning of research efforts to capture impact generative mechanisms, not necessarily impact achieved, but also to consider several other elements that contribute to the results especially as far as impact is concerned. In this respect the categories of 'stakeholders' and 'civil society' need some special reflections because we are speaking of *actors* involved in the process of generating impact that have very diverse needs and attitude in local contexts.

Therefore, it is compelling a new effort to find appropriate indicators, methods, to adapt to different contexts, fields, stakeholders and civil society. But also, new evaluation approaches/model with sound theoretical foundation beyond good practices.

Cultural change of all the actors in the R&D system

Cultural change is the most complex issue of the transformative process we are discussing. It implies to act at different levels: at government level, with intermediaries such as funding agencies, research institutions, and academics in different epistemic circles. Many research results from the recent literature indicate that building institutional support for impact – to individuals and for institutional innovation is extremely important. However, the cultural shift we need is bigger and it is related to the acknowledgment of the need to overcome segmentation of universities in missions to acknowledge that universities and research should be transformative agents of society.

The design of specific measures needs to be analysed with plenty of evidence: professionals' support for impact need also the involvement of researchers to be effective not only for society but also on the research side. It is urgent to pursue the academics' recognition for impact (visibility and academic career) and to find right ways to manage the inherent tensions between excellence and impact that steering and evaluation instruments can produce on the scholars.

Are scholars, research organizations and policy makers ready to change?

References

Bayley, J. and Phipps, D. (2019). 'Extending the concept of research impact literacy: Levels of literacy, institutional role and ethical considerations'. *Emerald Open Research*, pp. 1–14. Available at: https://doi.org/10.35241/emeraldopenres.13140.2

Bleischwitz, R., Sord, M.T., Ciarli, T., Ferr, M., Estradivari, Matt, M., Binot, A., Soler-Gallart, M. and Chavarro, D. (2025). *Transformative Research Assessment: Integrating Societal Impacts into Evaluation Frameworks*. Available at: https://doi.org/10.5281/zenodo.17722382

Edler, J., Matt, M., Polt, W. and Weber, M. (eds.) (2025). *Transformative Mission-Oriented Innovation Policies*. Edward Elgar Publishing. Available at: https://www.e-

elgar.com/shop/gbp/transformative-mission-oriented-innovation-policies-9781803929514.html

Feller, I. (2022). 'Assessing the societal impact of publicly funded research'. *The Journal of Technology Transfer*, 47(3), pp. 632–650. Available at: https://doi.org/10.1007/s10961-017-9602-z

Joly, P.-B. and Matt, M. (2022). 'Towards a new generation of research impact assessment approaches'. *The Journal of Technology Transfer*, 47(3), pp. 621–631. Available at: https://doi.org/10.1007/s10961-017-9601-0

Ochsner, M. and Bulaitis, Z.H. (2023). *Accountability in Academic Life: European Perspectives on Societal Impact Evaluation*. Edward Elgar Publishing. Available at: https://doi.org/10.4337/9781800885738

Reale, E. (2022). 'Factors enabling social impact: The importance of institutional entrepreneurship in social science research'. *Science and Public Policy*, 49(4), pp. 632–642. Available at: https://doi.org/10.1093/scipol/scac014

Schneider, F., Buser, T., Keller, R., Tribaldos, T. and Rist, S. (2019). 'Research funding programmes aiming for societal transformations: Ten key stages'. *Science and Public Policy*, 46(3), pp. 463–478. Available at: https://doi.org/10.1093/scipol/scy074

Soler-Gallart, M. and Flecha, R. (2022). 'Researchers' perceptions about methodological innovations in research oriented to social impact: Citizen evaluation of social impact'. *International Journal of Qualitative Methods*, 21. Available at: https://doi.org/10.1177/16094069211067654

Traag, V., Brady, E., Simón-Martínez, C., Rafols, I. and Schneider, J. (2026). *Targeted Funding and Changing Research Landscapes*. RoRI Working Paper No. 18, February 2026. Available at: https://doi.org/10.6084/m9.figshare.31282411

BEYOND OUTPUTS: MEASURING PUBLIC ENGAGEMENT IN SOCIAL SCIENCES

NÚRIA BAUTISTA-PUIG
Instituto de Filosofía, National Research Council (CSIC), Spain
ENRIQUE ORDUÑA-MALEA
The iMetrics Lab. Department of Audiovisual Communication, Documentation and History of Art, Universitat Politècnica de València, Spain
RODRIGO LOZANO
Instituto de Filosofía. National Research Council (CSIC), Spain; Chair of Business Administration, especially Environmental Management, Dresden University of Technology, Germany

Keywords: Citizen science, Sciences, Public engagement, Societal impact, Responsible research assessment

Since the late 20th century, scientific practice has shifted from a post-war model grounded in one-way communication (Mode-1) knowledge production towards more dialogic, participatory, and reflexive approaches (Mode-2) (Nowotny et al., 2003). This transformation reflects a growing emphasis on societal relevance, accountability, and responsiveness to public values and needs (Gregory and Lock, 2008; Skarlatidou and Haklay, 2021). In the early 2000s, public engagement gained prominence as a response to scientific controversies and increasing demands for more democratic forms of research governance (Bauer, 2009). These developments have contributed to a broader rethinking of how research quality and impact should be understood, documented, and assessed (Bornmann, 2013), particularly within frameworks of responsible research assessment (RRA).

Public participation in knowledge production has increasingly been framed as a key element of responsible and societally embedded research practices, promoting inclusive, transparent, and socially robust scientific processes (Stilgoe et al., 2014; Skarlatidou and Haklay, 2021). Involving non-professional actors can enhance the legitimacy and accountability of research, and foster mutual learning between science and society (Marzuki, 2015; Monzón Alvarado et al., 2020). From the perspective of RRA, such forms of participation challenge narrow output-oriented views of scientific contribution by emphasising relationships, and modes of knowledge production that extend beyond academia.

Within this context, Citizen Science (CS) has emerged as a practical approach to operationalising public participation in research. CS encompasses a wide range of initiatives in which citizens contribute to scientific activities, moving beyond research conducted on or about people towards research conducted with or by them (Cornwall and Jewkes, 1995; Serbe-Kamp et al., 2023). Through citizens' involvement, CS seeks to foster dialogue, shared understanding, and in some cases, co-production of knowledge oriented towards action (Bonney et al., 2016; Tõnisson et al., 2021). As such, CS provides a concrete method to examine how participatory values are translated into research practice.

Citizen science has expanded across many research fields, particularly, environmental and life sciences (Bautista-Puig et al., 2019; Bautista-Puig and Orduña-Malea, 2024; Kumar and Singh, 2021; Pelacho et al., 2020). In the social sciences, CS is not only understood as a consolidated set of participatory methodologies conducted outside the laboratory, but also as an approach centred on social issues and concerns raised by groups of citizens, and on the ways in which these concerns contribute to the production of new scientific knowledge (Albert et al., 2021). Citizen social science therefore offers particular potential for introducing reflexive dimensions into research practice, building on the long-standing legacy of participatory and co-productive methods in the social sciences.

Despite its rapid spread and its frequent connection to policy discourses on responsible research, the actual depth of citizen involvement within CS initiatives remains highly uneven. While some projects involve citizens as collaborators or co-creators, the dominant form of engagement continues to rely on contributory models focused on data collection, largely because such designs are easier to implement and scale (Turrini et al., 2018; Serbe-Kamp et al., 2023). This heterogeneity raises important questions for research assessment, where citizen participation is often reported as a checkbox indicator of societal impact, without distinguishing between different levels of involvement or their epistemic and societal implications.

A growing body of literature has proposed frameworks to classify public participation according to its scope and intensity, ranging from early hierarchical models of citizen control (Arnstein, 1969) to later approaches emphasizing information flows, deliberation, and decision-making authority (Rowe and Frewer, 2005; Cornwall, 2008). Within CS, engagement is commonly conceptualized along a range from contributory models to full co-creation (Haklay, 2013), complemented by task-, goal-, and role-based typologies (Shirk et al., 2012; Wiggins and Crowston, 2011, 2012; Finke and Laszlo, 2014).

Despite this diversity of conceptual approaches, previous studies report a gap between participatory ideals and their implementation in practice. Project-level analyses show that most CS initiatives remain concentrated at lower levels of engagement, even in domains where higher involvement might be expected (Schleicher and Schmidt, 2020; Vasiliades et al., 2021). This gap is partly explained by methodological limitations: many studies rely on small or discipline-specific samples, and few large-scale data sources provide sufficiently detailed descriptions of participant tasks. As a result, the depth of citizen involvement remains difficult to assess and to document in ways that are meaningful for research evaluation.

Methods

We analysed 1,746 CS projects collected from the SciStarter platform, a global CS hub that provides project-level descriptions of participant tasks and modes of involvement. For each project, we extracted the project title, task description, goal, general description, and 'field' classification. These project-level descriptions were treated as proxies of how citizen participation is formally embedded within research design. While such descriptions do not capture participants' experiences, they reflect how public involvement is conceptualized by project coordinators, an aspect that is particularly relevant for research assessment. The projects were coded using Haklay's (2013) four-level typology to analyse engagement depth.

Coding was based on the task field, supported by goals and general descriptions where necessary. When multiple tasks were present, the project was assigned to the highest engagement level. Beyond categorical coding, we developed a composite Engagement

Indicator (EI) to compare engagement depth across fields while controlling for field size. Where available, we also recorded project-reported outputs by linking SciStarter projects to funding acknowledgement data indexed in Web of Science.

Preliminary results

Across the full dataset, CS projects concentrate strongly in low-engagement models: Levels 1–2 account for more than 90% of projects. This confirms that, at scale, CS most often relies on contributory designs characterized by data provision and limited interpretation. At the same time, large environmental categories dominate in volume but cluster at lower average engagement levels, while several smaller fields show higher relative shares of Levels 3–4, indicating deeper forms of participation.

Focusing on the Social Science category, a different profile emerges. Although the absolute number of projects is comparatively small (111 projects), social-science labelled initiatives show a higher relative presence of deeper engagement compared to other fields (81% projects in Level 1 and 2). This suggests that when CS is implemented in social science contexts, it more frequently involves participatory designs that extend beyond data contribution towards shared interpretation. The Engagement Indicator value (2.04) supports this. This pattern may reflect the epistemological features of social sciences, where research centres on social actors and collective practices, fostering deeper involvement even when it does not result in co-authored outputs. At the same time, preliminary results show that these projects translate into relatively few publications, and that recognition remain limited, with citizen contributors rarely credited through authorship or explicit acknowledgements.

These findings have two key implications for RRA. First, they show that public involvement is heterogeneous and cannot be captured by a single indicator of engagement or societal impact, particularly in the social sciences, where reflexive and participatory traditions are integral to research practice. Second, the results point to concrete ways of recognising diverse contributions beyond journal-based indicators, using engagement depth as contextual evidence in narrative CVs to document how engagement is organised, the roles citizens assume, and how knowledge circulates. This is particularly relevant for ex post research evaluations, where traditional indicators often fail to capture the forms of value generated through participatory processes. Recent work suggests that engagement depth may also signal the credibility of societal impact in research proposals through structured interaction with non-academic actors (Benneworth and Olmos-Peñuela, 2024).

References

Albert, A., Balázs, B., Butkevičienė, E., Mayer, K. and Perelló, J. (2021). 'Citizen social science: New and established approaches to participation in social research', in Vohland, K. et al. (eds.), *The Science of Citizen Science*. Springer, pp. 119–138. Available at: https://doi.org/10.1007/978-3-030-58278-4_7

Arnstein, S.R. (1969). 'A ladder of citizen participation'. *Journal of the American Institute of Planners*, 35(4), pp. 216–224. Available at: https://doi.org/10.1080/01944366908977225

Bauer, M.W. (2009). 'The evolution of public understanding of science: Discourse and comparative evidence'. *Science, Technology & Society*, 14(2), pp. 221–240. Available at: https://doi.org/10.1177/097172180901400202

Bautista-Puig, N., De Filippo, D., Mauleón, E. and Sanz-Casado, E. (2019). 'Scientific landscape of citizen science publications: Dynamics, content and presence in social media'. *Publications*, 7(1), 1. Available at: https://doi.org/10.3390/publications7010001

Bautista-Puig, N. and Orduña-Malea, E. (2024). 'How do citizens engage with science: Challenges and recommendations'. *Profesional de la Información*, 33(6). Available at: https://doi.org/10.3145/epi.2024.0601

Benneworth, P. and Olmos-Peñuela, J. (2024). 'An openness framework for ex ante evaluation of societal impact of research'. *Research Evaluation*, 33, rvac023.

Bonney, R., Phillips, T.B., Ballard, H.L. and Enck, J.W. (2016). 'Can citizen science enhance public understanding of science?'. *Public Understanding of Science*, 25(1), pp. 2–16. Available at: https://doi.org/10.1177/0963662515607406

Bornmann, L. (2013). 'What is societal impact of research and how can it be assessed? A literature survey'. *Journal of the American Society for Information Science and Technology*, 64(2), pp. 217–233.

Cornwall, A. (2008). 'Unpacking "participation": Models, meanings and practices'. *Community Development Journal*, 43(3), pp. 269–283. Available at: https://doi.org/10.1093/cdj/bsn010

Cornwall, A. and Jewkes, R. (1995). 'What is participatory research?'. *Social Science & Medicine*, 41(12), pp. 1667–1676.

Finke, P. and Laszlo, K. (2014). *Citizen Science: Das unterschätzte Wissen der Laien*. Oekom Verlag.

Gregory, J. and Lock, S.J. (2008). 'The evolution of "public understanding of science": Public engagement as a tool of science policy in the UK'. *Sociology Compass*, 2(4), pp. 1252–1265. Available at: https://doi.org/10.1111/j.1751-9020.2008.00137.x

Haklay, M. (2013). 'Citizen science and volunteered geographic information: Overview and typology of participation', in Sui, D.Z., Elwood, S. and Goodchild, M. (eds.), *Crowdsourcing Geographic Knowledge*. Springer, pp. 105–122. Available at: https://doi.org/10.1007/978-94-007-4587-2_7

Kumar, A. and Singh, A. (2021). 'Analysis of citizen science scientific publications: A scientometric study'. *Library Philosophy and Practice*, pp. 1–24.

Marzuki, A. (2015). 'Challenges in the public participation and the decision-making process'. *Sociologija i Prostor*, 53(1), pp. 21–39.

Monzón Alvarado, C.M., Zamora Rendón, A. and Vázquez Pérez, A. del S. (2020). 'Integrating public participation in knowledge generation processes: Evidence from citizen science initiatives in Mexico'. *Environmental Science & Policy*, 114, pp. 230–241. Available at: https://doi.org/10.1016/j.envsci.2020.08.007

Nowotny, H., Scott, P. and Gibbons, M. (2003). 'Introduction: "Mode 2" revisited: The new production of knowledge'. *Minerva*, 41(3), pp. 179–194.

Pelacho, M., Ruiz, G., Sanz, F. et al. (2020). 'Analysis of the evolution and collaboration networks of citizen science scientific publications'. *Scientometrics*, 124, pp. 1–29. Available at: https://doi.org/10.1007/s11192-020-03724-x

Rowe, G. and Frewer, L.J. (2005). 'A typology of public engagement mechanisms'. *Science, Technology, & Human Values*, 30(2), pp. 251–290. Available at: https://doi.org/10.1177/0162243904271724

Schleicher, K. and Schmidt, C. (2020). 'Citizen science in Germany as research and sustainability education: Analysis of the main forms and foci and its relation to the Sustainable Development Goals'. *Sustainability*, 12(15), 6044. Available at: https://doi.org/10.3390/su12156044

Schrögel, P. and Kolleck, A. (2018). 'The many faces of participation in science'. *Science & Technology Studies*, 31(2), pp. 77–99.

Serbe-Kamp, É., Bemme, J., Pollak, D. et al. (2023). 'Open citizen science: Fostering open knowledge with participation'. *Research Ideas and Outcomes*, 9, e96476. Available at: https://doi.org/10.3897/rio.9.e96476

Shirk, J.L., Ballard, H.L., Wilderman, C.C. et al. (2012). 'Public participation in scientific research: A framework for intentional design'. *Ecology and Society*, 17(2), 29. Available at: https://doi.org/10.5751/ES-04705-170229

Skarlatidou, A. and Haklay, M. (2021). 'Citizen science impact pathways for a positive contribution to public participation in science'. *Journal of Science Communication*, 20(6), A02. Available at: https://doi.org/10.22323/2.20060202

Stilgoe, J., Lock, S.J. and Wilsdon, J. (2014). 'Why should we promote public engagement with science?'. *Public Understanding of Science*, 23(1), pp. 4–15. Available at: https://doi.org/10.1177/0963662513518154

Tõnisson, L., Voigtländer, J., Weger, M. et al. (2021). 'Knowledge transfer with citizen science: Luft-Leipzig case study'. *Sustainability*, 13(14), 7855. Available at: https://doi.org/10.3390/su13147855

Turrini, T., Dörler, D., Richter, A. et al. (2018). 'The threefold potential of environmental citizen science: Generating knowledge, creating learning opportunities and enabling civic participation'. *Biological Conservation*, 225, pp. 176–186. Available at: https://doi.org/10.1016/j.biocon.2018.03.024

Vasiliades, M.A., Hadjichambis, A.Ch., Paraskeva-Hadjichambi, D. et al. (2021). 'A systematic literature review on the participation aspects of environmental and nature-based citizen science initiatives'. *Sustainability*, 13(13), 7457. Available at: https://doi.org/10.3390/su13137457

Wiggins, A. and Crowston, K. (2011). 'From conservation to crowdsourcing: A typology of citizen science', in *Proceedings of the 44th Hawaii International Conference on System Sciences*. IEEE. Available at: https://doi.org/10.1109/HICSS.2011.207

Wiggins, A. and Crowston, K. (2012). 'Goals and tasks: Two typologies of citizen science projects', in *Proceedings of the 45th Hawaii International Conference on System Sciences*. IEEE. Available at: https://doi.org/10.1109/HICSS.2012.295

INTEGRATING SOCIETAL IMPACTS INTO EVALUATION FRAMEWORKS: INTRODUCTION TO AND SSH-RELATED COMMENTS ON THE REPORT FROM THE COARA WG TRANSFORMATIVE RESEARCH ASSESSMENT

ESTRADIVARI
Leibniz Centre for Tropical Marine Research (ZMT), Germany
MICHAEL OCHSNER
ENRESSH; FORS, Switzerland

In collaboration with:
RAIMUND BLEISCHWITZ
Leibniz Centre for Tropical Marine Research (ZMT), University of Bremen, Germany
TERESA SORDÉ MARTÍ
Autonomous University of Barcelona, Spain
TOMMASO CIARLI
The United Nations University – Maastricht Economic and Social Research Institute on Innovation and Technology (UNU-MERIT), Netherlands
MARIE FERRÉ
French Agricultural Research Centre for International Development (CIRAD), France
MIREILLE MATT
French National Research Institute for Agriculture, Food and Environment (INRAE), France
AURÉLIE BINOT
French Agricultural Research Centre for International Development (CIRAD), France
MARTA SOLER-GALLART
University of Barcelona, Spain
DIEGO CHAVARRO
Research Policy Solutions, Colombia

Keywords: Societal impact, Interdisciplinarity, Transdisciplinarity, Principles

Background

The Coalition for Advancing Research Assessment (CoARA) is a response to the fact that many scholars have become critical regarding dominant research assessment practices. The Coalition, founded in 2022, brings together scholars, university administrators, funders and other stakeholders of many disciplines to reform research assessment. Since 2023, different Work Groups have explored specific topics to improve research assessment. This presentation reports on the main output of the subgroup "Societal Impact" that is part of the Work Group "Towards Transformations" (Bleischwitz et al., 2025). More than twenty members across a diverse range of

disciplines and regions contributed to the report. It thus puts forward a pluralistic understanding of both scientific and societal impacts, emphasising that research generates value in diverse, non-linear pathways (see, e.g., Muhonen et al., 2020), and often through co-created processes and productive interactions (e.g., Spaapen & van Drooge, 2011).

The goal of the report is to strengthen the role of societal impacts in research assessment by offering a shared perspective on what impacts entail and encouraging more inclusive, transparent, and context-sensitive conventions for valuing research.

Being an interdisciplinary project, it is important to define the boundaries of the terms applied. The report understands *assessment* as system-level judgments tied to decisions (e.g. hiring, promotion, funding) and *evaluation* as the systematic gathering and analysis of evidence to support both learning and accountability. *Impact* is understood as evidenced contributions to societal improvements, acknowledging shared causality and time lags.

The report offers an overview of different approaches and tools to evaluate or assess the societal impact of research currently applied in Europe. It then develops principles for the integration of societal impact into research assessment frameworks and presents steps for different stakeholders to bring those principles into action. The presentation will focus on the principles and actions and will discuss these with regard to the assessment and evaluation of SSH research and their relation to principles and procedures developed within the SSH evaluation community.

While so far societal impact was discussed on the EU-level mainly in the SSH (EU-Presidency Conferences on Societal Impact in the SSH in Vienna 2018 (https://www.ssh-impact.eu) and on Research Evaluation in Paris 2022, HCÉRES, 2023), it has reached higher attention through CoARA culminating in the EU-High-Level Conference on Reforming Research Assessment in Copenhagen in 2025 (https://www.cerra.aau.dk) covering all disciplines. It is therefore important to examine similarities and differences in recommendations of the CoARA Work Group and the SSH-related recommendations.

Principles

Six principles are proposed to guide the integration and assessment of societal impacts across research systems in ways that are context-sensitive, inclusive and aligned with public value. The first principle poses to (1) *prioritise societal relevance* without compromising scientific autonomy. This can be achieved by embedding relevance to institutional missions. The second principle is to (2) *embrace pluralism* as societal impacts come in many shapes and forms, through different pathways, both direct and indirect. Pluralism should be treated as a strength for capturing complexity rather than a weakness. The third principle asks for (3) *planning for impacts from the start*. The research process should include from proposal stage to post-project reflection the identification and collaboration with stakeholders. The fourth principle asks to (4) *co-create impact*. Impactful research is rarely produced in isolation. Instead, collaboration among researchers, civil society, policymakers, businesses and other actors is essential. Co-creation should be meaningful, implemented in all stages and aware of power dynamics. Funders and institutions must act as partners, not only as initiators. The fifth principle reminds to (5) *assess both benefits and risks*. Impact is not always positive (Derrick et al., 2018), unintended or negative consequences must be accounted for, and potential environmental, ethical or social trade-offs must be considered. It is thus important to embed ethical reflection and anticipatory governance into evaluation processes. The sixth principle asks for (6) *balancing*

accountability with learning. It is advised to combine summative and formative evaluation procedures as the first provide transparency while the second enable reflection, adaptation and responsiveness.

Actions

These principles are transformed into coordinated action points for different stakeholders in order to bring change. *Funders* should create calls that combine scientific excellence with societal relevance and support co-creation. Furthermore, they should invest in evaluator capability. *Research organisations* must integrate aspects of societal impact in hiring and promotion procedures, establish broker roles and provide protected time for stakeholder engagement. *Researchers* should engage with stakeholders throughout the research process. Specifically, early career researchers should receive impact mentorship, training and realistic expectations, while senior researchers should champion institutional change in evaluation and act as mentors. *Businesses* can drive societal impact by partnering and co-funding research as well as by upholding transparency and independence. *The civil society and policy makers*, finally, should co-define research agendas, validate findings and ensure that research addresses real societal needs through consultation and partnership.

Conclusion

The report proposes six principles forming a framework that translates to coordinated action points that can be implemented by several stakeholders. The insights and recommendations of this report will be compared with similar works from within the SSH disciplines, such as Bulaitis (2020) regarding the value of humanities research, Sigurðarson (2022) on capabilities approach, Ochsner & Bulaitis (2023) comparing different national practices of impact evaluation and Benneworth & Olmos-Peñuela (2018) on tensions of utilization, amongst others.

References

Benneworth, P. and Olmos-Peñuela, J. (2018). 'Reflecting on the tensions of research utilization: Understanding the coupling of academic and user knowledge'. *Science and Public Policy*, 45(6), pp. 764–774. Available at: https://doi.org/10.1093/scipol/scy021

Bleischwitz, R., Sordé Martí, T., Ciarli, T., Ferré, M., Estradivari, Matt, M., Binot, A., Soler-Gallart, M. and Chavarro, D. (2025). *Transformative Research Assessment: Integrating Societal Impacts into Evaluation Frameworks*. Zenodo. Available at: https://doi.org/10.5281/zenodo.17722382

Bulaitis, Z. (2020). *Value and the Humanities: The Neoliberal University and Our Victorian Inheritance*. Palgrave Macmillan, Cham. Available at: https://doi.org/10.1007/978-3-030-37892-9

Derrick, G.E., Faria, R., Benneworth, P., Pedersen, D.B. and Sivertsen, G. (2018). 'Towards characterizing negative impact: Introducing Grimpact', in Costas, R., Franssen, T.P. and Yegros-Yegros, A. (eds.), *STI Conference Proceedings*. Leiden University, Leiden, pp. 1199–1213. Available at: https://research.utwente.nl/en/publications/towards-characterizing-negative-impact-introducing-grimpact

HCÉRES (2023). 'Evaluation of social sciences and humanities in Europe. Proceedings of the Conference organized by HCERES at Paris IAS, 16–17 May 2022'.

PIAS Proceedings of the Paris Institute for Advanced Study, 17. Available at: https://paris.pias.science/

Muhonen, R., Benneworth, P. and Olmos-Peñuela, J. (2020). 'From productive interactions to impact pathways: Understanding the key dimensions in developing SSH research societal impact'. *Research Evaluation*, 29(1), pp. 34–47. Available at: https://doi.org/10.1093/reseval/rvz003

Sigurðarson, E.S. (2020). 'Capacities, capabilities, and the societal impact of the humanities'. *Research Evaluation*, 29(1), pp. 71–76. Available at: https://doi.org/10.1093/reseval/rvz031

Spaapen, J. and van Drooge, L. (2011). 'Introducing "productive interactions" in social impact assessment'. *Research Evaluation*, 20(3), pp. 211–218. Available at: https://doi.org/10.3152/095820211x12941371876742

3.1 OPEN RESEARCH AND INFRASTRUCTURE

INFRASTRUCTURAL CONSTRAINTS ON THE MOVE TOWARDS OPEN RESEARCH INFORMATION SYSTEMS: ANALYSING THE INVISIBLE WORK OF METADATA CURATORS

AUBE RICHEBOURG, DIDIER TORNY
Mines Paris, PSL University, Centre for the sociology of innovation (CSI), i3 UMR9217 CNRS, France

Keywords: Barcelona Declaration, Metadata curation, Open science, Research assessment reform, Research information platforms

Since the 2010's, the transnational research assessment reform movement has been rising awareness around "responsible metrics" in research evaluation. As goals of open science were incorporated in the "responsible research assessment movement", open metadata appeared progressively in the framing texts of the movement as one of their main features (Rushforth & Hammarfelt, 2023, CoARA, 2022, Barcelona Declaration, 2024).

Following the movement in the national context, French universities and research organisations took significant actions to comply with the first National Plan for Open Science (2018) regarding research assessment, namely "to bring publicly funded research out of the confines of closed databases" (National Plan for Open Science, 2018). Alongside signing the San Francisco Declaration on Research Assessment (DORA), the Paris Call on Research Assessment and joining the Coalition for Advancing Research Assessment (CoARA), typical actions include cancelling subscriptions to research information (RI) commercial services. Moving towards self-administered information systems and using open metadata sources are commonly part of the "overarching goal to maximize the quality and impact of research" (Sorbonne Université, 2022). In this respect, CNRS, Sorbonne Université, Université de Lorraine, and others "broke free" (CNRS, 2025) from commercial research information platforms such as Web of Science or Scopus to favour non-commercial alternatives like OpenAlex. Their aim is to produce better knowledge of scientific output and fairer research evaluation metrics. However, translating these commitments into technological implementation raises one specific problem commonly overlooked by decision-makers: the curation of scientific products metadata (author names, affiliations, funding, etc.). How does the implementation of open data policies transform metadata curation into a central issue for research evaluation?

This presentation will provide empirical insights into the transformation of metadata curators' work in the context of an institutional move towards open science. As a part of the comparative research project "Between Economy and Democracy: Reorganizing Research Evaluation through Metadata in the Digital Era" funded by four national research agencies, this presentation is based on early-stage postdoctoral research. We will present empirical data either public (material press releases and online publications, official reports, communication materials) or collected through

interviews with intermediaries in research evaluation processes, such as librarians and bibliometricians. As data curators and managers, they play a key albeit mostly invisible part and undervalued (Moore, 2025) in cleaning the metadata to achieve maximal coverage of academic activities, while querying the databases to produce research reports for evaluation. They sometimes also play a role in the implementation of open science requirements making the shift to open databases technically possible.

Our presentation will thus address the following questions: How does open data transform their tasks? How can they curate across multiple platforms and software? How do open science or research evaluation policies frame these tasks? As it received limited scholarly attention compared with other actors of research evaluation (Dagiené et al., 2025), analysing the work of these intermediaries reveals the infrastructural constraints inherent in implementing open science policies.

Described as "an evolving sociotechnical practice" (Tammaro et al., 2019), data curation in research has received significant attention, especially in library and information science (Chao et al., 2015; Lee & Stvilia, 2014; Minamiyama et al., 2024). Less documented is how the same objective—to classify, exemplify, and rationalize data curation activities—applies among professionals dealing with research information metadata. In recent years, many data specialists such as librarians and data analysts have published practical guides about collecting, curating, and using such metadata for bibliometric purposes (Bach et al., 2025; Bordignon, 2025; de Jonge & Kramer, 2025). These self-help practices question, on one hand, the role of higher-level supervision of such crucial activities, which focus more on aggregated tools and indicators rather than metadata. On the other hand, they demonstrate professionals' capacity to work around or manage infrastructural constraints at a horizontal level, highlighting their crucial role as intermediaries in implementing open science policies.

However, "metadata curation", as a set of activities, remains largely black-boxed and out of view. The presentation will focus on one example of technological implementation of research assessment reform in the French context: the shift to OpenAlex in two universities. Following the signature of the *Declaration of Barcelona*, both universities decided to dedicate human and time resources to investigate OpenAlex as a legitimate alternative to commercial databases. But the "data quality management" ("mise en qualité des données"), as it is called, entails *frictions*. Drawing on the ethnomethodological approach of technical practices studies (Lynch, 1985), interviews focusing on practices and practical reasoning, will be conducted with five librarians and bibliometricians in each institution selected as they participated in and documented this endeavour will aim at exploring the process of metadata curation tasks and the role of intermediary platforms ("works magnet" and "LODEX" both developed by public actors). Tasks such as ROR cleaning ("nettoyage du ROR"), spotting ("repérage"), affiliation disambiguation ("désambiguïsation des affiliations"), affiliation enrichment ("enrichissement d'affiliation") or alignment ("alignement") reveal the underlying infrastructure of research assessment while becoming *problematic* in the context of other political entanglements.

In the first part, we demonstrate that metadata curation is an activity entangled with three frameworks: institutional mergers, open science public policies, and the national infrastructuring of research information, notably the French national archive HAL and *ad hoc* software tools. In the second part, we show that librarians and bibliometricians, as producers, managers and users of metadata, play a significant role in the shift toward open platforms. In pursuing the goal of improving their institutional representation on databases used in research evaluation, they manage the transmission of information between platforms. In this respect, analysing metadata

curators and their activities reveals both the infrastructural constraints of changing research information platforms, according to research assessment reform principles, and potential pathways to address them.

References

Bach, C., Bourguignon, L., Guélé, C., Houdry, P. and Kremer, A. (2025). *Bibliométrie prête à l'emploi avec OpenAlex: Retour d'expérience*. Available at: https://cnrs.hal.science/hal-05003502

Barcelona Declaration on Open Research Information (2024). Kramer, B., Neylon, C. and Waltman, L. *Barcelona Declaration on Open Research Information*. Available at: https://doi.org/10.5281/zenodo.10958522

Bordignon, F. (2025). *OpenAlex au service de l'évaluation de la recherche: L'idéal d'ouverture à l'épreuve de la qualité des données*. Evaluation de la recherche et outils alternatifs. Available at: https://enpc.hal.science/hal-05363827

Chao, T.C., Cragin, M.H. and Palmer, C.L. (2015). 'Data Practices and Curation Vocabulary (DPCVocab): An empirically derived framework of scientific data practices and curatorial processes'. *Journal of the Association for Information Science and Technology*, 66(3), pp. 616–633. Available at: https://doi.org/10.1002/asi.23184

CoARA (2022). *Agreement on Reforming Research Assessment*. Available at: https://doi.org/10.5281/zenodo.13480728

CNRS (2025). *Le CNRS s'émancipe du Web of Science*. Available at: https://www.cnrs.fr/fr/actualite/le-cnrs-semancipe-du-web-science

Dagienė, E., Waltman, L. and Dix, G. (2025). 'Multi-actor policy dynamics in research evaluation: Experts, databases, and academics'. *Higher Education Policy*. Available at: https://doi.org/10.1057/s41307-025-00397-0

de Jonge, H. and Kramer, B. (2025). *Manuscript submission systems and metadata completeness in Crossref: Patterns and associations*. OSF. Available at: https://doi.org/10.31222/osf.io/ndx3f_v1

Lee, D.J. and Stvilia, B. (2014). 'Data curation practices in institutional repositories: An exploratory study'. *Proceedings of the American Society for Information Science and Technology*, 51(1), 14. Available at: https://doi.org/10.1002/meet.2014.14505101085

Lynch, M. (1985). *Art and Artifact in Laboratory Science: A Study of Shop Work and Shop Talk in a Research Laboratory*. Routledge Kegan & Paul, London.

Minamiyama, Y., Takeda, H., Hayashi, M., Asaoka, M. and Yamaji, K. (2024). 'A study on formalizing the knowledge of data curation activities across different fields'. *PLOS ONE*, 19(4), e0301772. Available at: https://doi.org/10.1371/journal.pone.0301772

Ministère de l'Enseignement supérieur et de la Recherche (2018). *Plan national pour la Science Ouverte*. Available at: https://www.ouvrirlascience.fr/plan-national-pour-la-science-ouverte

Moore, S. (2025). *Publishing Beyond the Market: Open Access, Care, and the Commons*. University of Michigan Press. Available at: https://doi.org/10.3998/mpub.11781635

Rushforth, A. and Hammarfelt, B. (2023). 'The rise of responsible metrics as a professional reform movement: A collective action frames account'. *Quantitative Science Studies*, 4(4), pp. 879–897. Available at: https://doi.org/10.1162/qss_a_00280

Sorbonne Université (2022). *Press release: Sorbonne University signs the CoARA Agreement*. Available at: https://www.sorbonne-universite.fr/en/presse/reforming-research-assessment

Tammaro, A.M., Matusiak, K.K., Sposito, F.A. and Casarosa, V. (2019). 'Data curator's roles and responsibilities: An international perspective'. *Libri*, 69(2), pp. 89–104. Available at: https://doi.org/10.1515/libri-2018-0090

CURRENT PRACTICES AND FUTURE CHALLENGES IN OPEN PEER REVIEW AMONG CANADIAN SCHOLARLY JOURNALS

LUCÍA CÉSPEDES, VINCENT LARIVIÈRE
Érudit; University of Montréal, Canada
ALEXANDRA MICHAUD, ARILYS JIA, ANNE-MARIE FORTIER, RENÉ AUDET, VANESSA ALLNUTT
Laval University, Canada
ADRIEN SAVARD-ARSENEAULT
University of Montréal, Canada

Keywords: Open peer review, Scholarly publishing, Canadian journals, Literature review, Survey

In this presentation, we share the results of a comprehensive report commissioned by the *Réseau québécois de recherche et de mutualisation pour les revues scientifiques* (Réseau Circé), which aimed to assess the current state of peer review among scholarly journals, focusing, in particular, on open peer review practices. Our objective was to explore open peer review as a lever for transforming scholarly communication and research assessment.

The research process leading to the report was threefold: first, we conducted a multilingual narrative literature review in French (Cairn, HAL and Érudit), English (Google Scholar), Spanish and Portuguese (Scielo and Redalyc). This part of the study allowed us, on the one hand, to document the historical evolution of peer review as a cornerstone of scientific legitimacy, as the effectiveness of traditional double-blind models is increasingly being questioned. On the other hand, this literature review revealed the most common arguments for and against different aspects of openness, as well as a plurality of definitions, conceptions, and practices across the publishing cultures of disciplinary and national fields. This highlights the need for a common understanding among editors, authors, reviewers and publishers.

In this sense, as a second step, we drew a portrait of the most prevalent and increasingly adopted open peer review practices, namely, open identities, open reviews, open participation and open interaction (Ross-Hellauer, 2017). These were thoroughly depicted in terms of the scientific and ideological issues involved, their impact on editorial workflows, and the tools and resources available or required to implement each modality. We acknowledge that no single list of practices can cover all the concerns and potential needs of diverse scholarly communities. Therefore, this part of our study aims to inform scholarly journals in Québec and Canada seeking to reconceptualize their peer review models, and to support them in determining the best ways to implement different peer review practices in a context-sensitive fashion.

Finally, we distributed a survey among editors of Canadian scholarly journals. Previous similar studies have included surveys among editors, reviewers and authors of international journals (Ware, 2016; Ross-Hellauer et al., 2017), contributors and editors of Spanish journals (Segado-Boj et al., 2018; Abadal & Melero, 2023),

stakeholders involved in a Brazilian medical journal (Fontenelle & Sarti, 2021), and editors of journals indexed in the DOAJ (Maia & Farias, 2025). To the best of our knowledge, no such survey had been conducted in Canada. The questionnaire was created using Google Forms, with versions both in English and French, and sent to 768 Canadian scholarly journals identified from van Bellen's list (2025). The survey was open from April 23rd to May 23rd, 2025. A total of 141 responses was initially received (15 % response rate), 84 in French and 58 in English. After deduplication of repeated entries, we arrived at a total of 134 valid responses, representing journals published across nine Canadian provinces.

Time emerged as the scarcest and most difficult to manage resource for editors. Different problematic issues reinforce each other: the growing difficulty to find expert reviewers translates into an overburdening of those who usually accept to perform this task. In turn, this affects the quality and punctuality of reviews written under pressure. The "thematic dossier" or "special issue" stands as a space for innovation and careful introduction of new practices, which may later be integrated into the journals' regular editorial policies.

Nevertheless, open peer review remains a marginal practice in the Canadian context. Our results show a minimal adoption of open peer review practices and an adherence to traditional closed and anonymous evaluation models regardless of discipline, language of publication, access model or age of the journal. The vast majority do not disclose information about reviewers to authors, do not publicly reveal reviewers' identities, do not publish article reviews, nor do they enable any interaction between reviewers and authors or among reviewers. This indicates a certain conservatism (stemming from disciplinary or editorial traditions), a reticence to abandon anonymous review models, or a lack of resources to enact "experimental" review formats. However, at the same time, almost one third of the respondents stated that their journals are reassessing their peer review practices and even considering implementing changes.

The surveyed journal editors do not tend to consider changes towards open identities or open reports, but rather towards improving support, mentorship and training opportunities offered to authors and reviewers, even if these activities would be a priori more time-consuming. At least ten journals (from the social sciences, humanities, arts, literature, economy, and professional fields) mentioned their intentions in this direction. Editors manifested their intentions to foster participation of early career researchers or those whose first languages are not French nor English. These proposed forms of mentorship also seek to expand the geographical scope of journals in terms of a greater inclusion of researchers from the Global South, either as authors or reviewers. Given the reported difficulties to find relevant reviewers, the underrepresentation of researchers outside North America and Europe in the pool of potential reviewers available for Canadian journals and given that this pedagogical and formative aspect is probably one of the less evident and neglected forms of open research, this seems like a promising path to explore.

Thus, this report sheds light on Canadian journals' perspectives on peer review and constitutes a timely contribution for editorial strategic planning, as demands for openness and transparency increase within broader conversations regarding the reform of certain scientific practices. While still marginal and not without challenges, open peer review stands as a promising avenue to re-engage with the dialogic, communal, and formative aspects of knowledge construction, assessment and dissemination, in line with the core tenets and values of open science.

References

Abadal, E. and Melero, R. (2023). 'Open peer review: The point of view of scientific journal editors'. *JLIS.It*, 14(1), pp. 60–70. Available at: https://doi.org/10.36253/jlis.it-507

Fontenelle, L.F. and Sarti, T.D. (2021). 'Attitudes toward open peer review among stakeholders of a scholar-led journal in Brazil'. *Transinformação*, 33, e200072. Available at: https://doi.org/10.1590/2318-0889202133e200072

Maia, F.C. de A. and Farias, M.G.G. (2025). 'Percepções de editores quanto à implantação da revisão aberta em periódicos científicos'. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, 30, pp. 1–31. Available at: https://doi.org/10.5007/1518-2924.2025.e99922

Ross-Hellauer, T. (2017). 'What is open peer review? A systematic review'. *F1000Research*, 6(588), pp. 1–40. Available at: https://doi.org/10.12688/f1000research.11369.2

Ross-Hellauer, T., Deppe, A. and Schmidt, B. (2017). 'Survey on open peer review: Attitudes and experience amongst editors, authors and reviewers'. *PLOS ONE*, 12(12), e0189311. Available at: https://doi.org/10.1371/journal.pone.0189311

Segado-Boj, F., Martín-Quevedo, J. and Prieto-Gutiérrez, J.J. (2018). 'Attitudes toward open access, open peer review, and altmetrics among contributors to Spanish scholarly journals'. *Journal of Scholarly Publishing*, 50(1), pp. 48–70. Available at: https://doi.org/10.3138/jsp.50.1.08

van Bellen, S. (2025). 'Revues savantes canadiennes/Canadian scholarly journals (version 5.0)'. Borealis. Available at: https://doi.org/10.5683/SP3/9ONCEU

Ware, M. (2008). *Peer Review: Benefits, Perceptions and Alternatives*. Publishing Research Consortium, London.

OPEN SCIENCE AND ITS ROLE IN RESEARCH EVALUATION: CULTURAL CHANGE, OPEN INFRASTRUCTURES AND SCIENTIFIC AUTONOMY AT THE UNIVERSITY OF MILAN

PAOLA GALIMBERTI, CAROLINA MANFREDINI, MARGHERITA CRIVELLER
University of Milan, Italy

Keywords: Open science, Scientific autonomy, Research evaluation, Open research infrastructures, Cultural change in assessment

In contemporary research systems, scientific autonomy is increasingly negotiated within evaluation regimes shaped by global power asymmetries, commercial information infrastructures, and emerging algorithmic tools. From this perspective, Open Science should not be perceived as a technical reform or a set of best practices, but as a governance space in tension, where questions of control, legitimacy, and responsibility in knowledge production and assessment are continuously contested. In contexts marked by weak public coordination, Open Science exposes unresolved tensions concerning the role of the state, the autonomy of universities, and the criteria through which research quality is defined and assessed. This proposed paper presents an institutional case study focused on the University of Milan, analysing how Open Science has been governed and sustained over nearly two decades within the Italian research system (Bolelli Gallevi et al., 2025a).

Italy exemplifies a prolonged institutional transition: although Open Science principles are formally endorsed in national policy documents, this alignment has not been matched by the development of shared public infrastructures, stable funding mechanisms, or coordinated monitoring systems. As a result, Open Science remains largely driven by individual institutions, is organisationally fragmented, and interacts with a research evaluation system that continues to rely on proprietary data sources and performance-based metrics (Galimberti 2026). This situation places scientific autonomy under pressure, especially in social sciences and humanities, and limits how institutions control research information. Thus, the analysis here proposed adopts a process-oriented institutional perspective, based on administrative experience, to trace how Open Science policies, infrastructures, and services were progressively shaped through synergistic interaction between the internal Open Science Division and the university's Open Science Commission. The Commission, composed of representatives from each department and supported by internal policy units, functions as a plural consultive forum that brings together diverse disciplinary perspectives on Open Science. It has played an active role in formulating policies, monitoring practices, and recommending actions to governing bodies, while also contributing to designing and endorsing training programmes tailored to departmental needs. This internal governance interplay highlights how institutional actors collectively negotiate the transformation of evaluation cultures, rather than enact top-down mandates (Bolelli Gallevi et al., 2025b).

The case is not shown as a best practice, but as a lens to study how universities (or dedicated bodies within them) increasingly compensate for structural gaps in public coordination. Rather than following a predefined strategy, this trajectory emerged through the cumulative layering of decisions shaped by external constraints and limited margins of autonomy. Indeed, a central focus of the paper is research evaluation as an arena in which power relations, quality standards, and academic freedom are negotiated. In Italy, evaluation has become a dominant instrument of system governance, often delegating quality judgments to indicators produced by commercial platforms and opaque algorithmic tools. These dynamics are now amplified by the adoption of AI-driven assessment procedures, which increase efficiency at the cost of transparency, interpretability, and epistemic diversity. In response, the integrated body composed by the OS Division and the OS Commission of the University of Milan pursued a strategy of reclaiming informational sovereignty through the systematic use of open infrastructures, open research information, and open data (Galimberti, 2025).

A pivotal decision was to embed Open Science governance within the institutional unit responsible for research evaluation and quality assurance, rather than treating it as a technical or library-based function. This positioning enabled open infrastructures to directly inform evaluation practices. The adoption of Dataverse for research data and AIR as the institutional repository for publications provided the basis for transparent, reproducible, and institutionally controlled representations of research outputs. In parallel, the development of an internal research dashboard based on OpenAlex open data demonstrated how open bibliographic sources can support strategic analysis of collaborations, outputs, and research dynamics without reliance on proprietary databases. These infrastructural choices were underpinned by an explicit commitment to open information principles, reinforced by the University's engagement with the *Barcelona Declaration on Open Research Information* and the adoption of OpenAIRE MONITOR for reporting, strategic decision making, and monitoring. Together, these initiatives signal a shift from compliance- oriented openness to governance-driven openness, where infrastructures and metadata stewardship become critical instruments for rethinking evaluation itself.

Crucially, this infrastructural transformation occurred within the context of a broader cultural shift. Through the Commission and OS Division's deliberations and the co-design of training activities – ranging from FAIR data management to evaluation literacy for SSH and STEM – openness has been reinterpreted as an ethical and epistemological dimension of research practice rather than a mere administrative requirement. This cultural change enabled the challenge of the naturalisation of metric-based evaluation and the promotion of more plural, contextual and reflexive assessment practices. Open infrastructures supported a gradual rebalancing of quantitative indicators and qualitative judgement, creating space to recognise interdisciplinary research, collaborative practices and socially engaged scholarship, which are poorly captured by standard metrics. Overall, the case illustrates how strategic use of open infrastructures and research information, combined with deliberative institutional collaboration and cultural investment, can foster alternative visions of research evaluation and help safeguard scientific autonomy within fragmented and politicised research systems. Open Science thus emerges not as a technical solution, but as a cultural and governance project capable of reshaping how quality, responsibility, and trust are understood in research evaluation.

References

Bolelli Gallevi, S., Basset, D. and Galimberti, P. (2025a). 'Dati della ricerca aperti all'Università degli studi di Milano: Partecipazione attiva al contesto internazionale'. GenOA Week 2025. Zenodo. Available at: https://doi.org/10.5281/zenodo.17711229

Bolelli Gallevi, S., Filadoro, R. and Galimberti, P. (2025b). 'La Commissione Open Science dell'Università degli studi di Milano: Attività e ruolo'. GenOA Week 2025. Zenodo. Available at: https://doi.org/10.5281/zenodo.17711466

Galimberti, P. (2025). 'Fonti aperte per la valutazione della ricerca: Un'esperienza di transizione. Risorse aperte per il monitoraggio e la valutazione della ricerca'. COARA Italia training seminar. Zenodo. Available at: https://doi.org/10.5281/zenodo.15848477

Galimberti, P. (2026). 'Open science in assenza di un coordinamento pubblico. Appunti da un'esperienza istituzionale in un contesto di lunga transizione'. *Bollettino Telematico di Filosofia Politica*. Zenodo. Available at: https://zenodo.org/records/18184982

OPEN SCIENCE AND NEW RESEARCH ASSESSMENT CRITERIA: A TWO-WAYS PATH

ELENA GIGLIA
University of Turin, Italy

Keywords: Open science, Open research assessment, CoARA, Research assessment reform

The Open Science-research assessment binomial is bi-directional and can be read in both ways: what can a new research assessment framework do for Open Science and what Open Science can do for a new research assessment framework.

Aim of this talk is to explore how Open Science can at the same time enable the change in the current research assessment framework and be fostered by it.

Research evaluation criteria deeply inform the choices made by researchers and can drive their attention and time to privilege some type of outputs or practices (Coalition S, 2018). This is a common reaction every Open science advocate or trainer gets after a session presenting different Open science practices: "Fine, but we are still evaluated on the Impact factor of the journal". Quantitative, journal-based metrics prevent researchers from being actually free in their research practices, as they have to focus on mainstream topics to get published in the so called "prestigious journals", and they have to restrain themselves to journal articles as they are evaluated only upon the number of publications (Ederer, 2024). The pressure to publish reduces also the time dedicated to describe a sound methodology or to find solutions for appropriate data management, creating a reproducibility crisis (Ross-Hellauer 2023).

These criteria can also trigger so called "adaptive behaviours" in researchers (Baccini et al. 2019), and lead to questionable practices like dividing a single article in several ones, just to get more lines on their curricula, or even shift entire communities habits, like it happened to many researchers in the humanities who abandoned books – which have been for centuries the humanists benchmark, expression of their long reasoning and argumentation -, for articles, shorter and quicker to produce, again as they need more lines on curricula. The same logic triggered the phenomenon of predatory publishing (Neylon, 2015) with a huge economic damage, too, as this hypertrophic production of papers is draining billions of public money into the commercial publishing system, which is no longer functional to science (Beigel et al., 2025).

This hypercompetitive research assessment framework based on quantitative indicators can lead, in the end, also to serious unintended consequences, which have been under scrutiny for years now (Pizzolato, 2024; Guardian, 2024; Rahman, 2024): retractions (Fang & Casadevall, 2011; Freijedo et al., 2024; Oranski 2025), frauds (Biagioli &Lippman, 2020), citation cartels (Catanzaro, 2024), paper mills (Bishop, 2023; Conroy, 2024), evidence of scientific misconduct are in the news almost on a daily basis (O'Grady, 2025). Artificial Intelligence has put this system "on steroids", as Ivan Oransky said in an interview (Oranski 2025), with AI generated papers (Cabanac et al. 2025; Kabel, 2025), bibliographies, even reviews, which beware, can often be only realistic, not real (Sakai, 2026).

The European Council Conclusions formally recognised the issue: "(the Council) RECALLS that the current research assessment systems are nowadays to a great extent too focused on the use of some quantitative journal- and publication-based indicators and the evaluation of a narrow range of research outputs; CONSIDERS that such an approach may lead to negative biases in terms of research quality, reproducibility and integrity" (Council of the European Union, 2022). These Conclusions are considered as a legal basis for the COARA initiative on reforming research assessment, which is a milestone for the purpose of this talk.

COARA's Agreement signatories commit to recognise the diversity of contributions, and to abandon the inappropriate use of journal and citation based metrics like Impact Factor and h-index. In principle, the new research assessment should "reward early sharing and open collaboration" (COARA, 2022).

In this presentation we shall start with an overview of the worst examples and the major concerns about research integrity linked to the current research assessment criteria, to then explore the promising bi-directional relationship with Open Science and the imperative to "slow down", as Elizabeth Bik puts it (Olaizola, 2024).

On one hand, the shift – pledged by COARA signatories - towards a system where the entire research process has its value, where every step of the research workflow can – and should – be recognized as an "output", where collaboration and interdisciplinarity are rewarded is crucial to foster and spread the adoption of Open Science practices: here is the role for evaluation in making Open Science not only possible but "the new normal", as we read in the EOSC Association SRIA - Strategic Research and Innovation Agenda (EOSC Association, 2023). Changing the research evaluation criteria is the only way we have to truly guarantee academic freedom, and let researcher be curious, never sacrificing again what's right for their research for what's right for their career to abide to the publish or perish culture and the Impact factor dictatorship.

On the other end, Open Science practices like preregistration, open peer review, open lab notebooks, can play a role in ensuring a better quality and integrity of the research process through transparency, and improving research quality is ultimately the aim of any exercise of research assessment. FAIR data management, thanks to the support of the new high profile professional figure of data stewards, can solve the security concern – so widespread nowadays - at the source, as good data stewardship first addresses security questions and find suitable solutions according to the principle "as open as possible, as closed as necessary" (Slifka, 2024). Actually, unmanaged data are a liability, instead of an asset. All the more so, quality FAIR data can avoid to incur in the "rubbish in, rubbish out" situation in training AI models (Moseley, 2024). Access to methods, protocols, code, can also enhance reproducibility (Hosseini et al, 2025). And for research evaluation purposes, they can enlarge the concept of "research output", getting rid of the tyranny of publications, to better align research assessment to the new digital world, data driven, AI pervaded, less linear, more collaborative and multi- and cross-disciplinary to tackle the global changes we are facing.

These principles have been translated into reality in different European countries: from the Norwegian NOR-CAM framework including all the research aspects into evaluation (NOR-CAM, 2021), to the Room for everyone's talent Dutch initiative (Room for everyone's talent, 2022), to the Finnish FAIRer research approach (Pölönen & Mustajoki, 2022), we shall present how Open Science and a new research evaluation can benefit from each other.

References

Baccini, A., De Nicolao, G. and Petrovic, E. (2019). 'Citation gaming induced by bibliometric evaluation: A country-level comparative analysis'. *PLOS ONE*, 14(9), e0221212. Available at: https://doi.org/10.1371/journal.pone.0221212

Beigel, F., Brockington, D., Crosetto, P. et al. (2025). *The drain of scientific publishing*. ArXiv preprint. Available at: https://doi.org/10.48550/arXiv.2511.04820

Biagioli, M. and Lippmann, A. (eds.) (2020). *Gaming the Metrics: Misconduct and Manipulation in Academic Research*. MIT Press, Boston. Available at: https://doi.org/10.7551/mitpress/11087.001.0001

Bishop, D. (2023). 'Defence against the dark arts: A proposal for a new MSc course'. *Bishopblog*, November. Available at: https://deevybee.blogspot.com/2023/11/defence-against-dark-arts-proposal-for.html

Cabanac, G., Labbé, C. and Joelving, F. (2025). 'Problematic Paper Screener: Trawling for fraud in the scientific literature'. *The Conversation*, 29 January. Available at: https://theconversation.com/problematic-paper-screener-trawling-for-fraud-in-the-scientific-literature-246317

Catanzaro, M. (2024). 'Citation cartels help some mathematicians — and their universities — climb the rankings'. *Science*, 383(6682). Available at: https://www.science.org/content/article/citation-cartels-help-some-mathematicians-and-their-universities-climb-rankings

Coalition S (2018). *Science Without Publication Paywalls: A Preamble to cOAlition S for the Realisation of Full and Immediate Open Access*. Position paper. Available at: https://www.coalition-s.org/wp-content/uploads/cOAlitionS_Preamble.pdf

CoARA (2022). *Agreement on Reforming Research Assessment — ARRA*. Available at: https://www.coara.org/agreement/the-agreement-full-text/

Conroy, G. (2024). 'Surge in number of "extremely productive" authors concerns scientists'. *Nature*, 625, pp. 14–15. Available at: https://doi.org/10.1038/d41586-023-03865-y

Council of the European Union (2022). *Research Assessment and the Implementation of Open Science*. Council Conclusions June 2022, 10126/22. Available at: https://www.consilium.europa.eu/media/56958/st10126-en22.pdf

Ederer, P. (2024). 'Perspective on scientific truth versus scientific evidence: Maintaining integrity in global food systems'. *Animal Production Science*, 64, AN23331. Available at: https://doi.org/10.1071/AN23331

EOSC Association (2023). *Strategic Research and Innovation Agenda 1.2*. Available at: https://eosc.eu/wp-content/uploads/2023/12/20231114_SRIA_1.2_final2.pdf

ERAC Standing Working Group on Open Science and Innovation (2021). *Guideline Report on Research Integrity and Open Science*. European Union, Brussels. Available at: https://data.consilium.europa.eu/doc/document/ST-1207-2021-INIT/en/pdf

European Commission, Directorate-General for Research and Innovation (2017). *Evaluation of Research Careers Fully Acknowledging Open Science Practices: Rewards, Incentives and/or Recognition for Researchers Practicing Open Science*.

EU Publications Office, Luxembourg. Available at: https://data.europa.eu/doi/10.2777/75255

Fang, F.C. and Casadevall, A. (2011). 'Retracted science and the retraction index'. *Infection and Immunity*, 79. Available at: https://doi.org/10.1128/iai.05661-11

Freijedo-Farinas, F., Ruano-Ravina, A., Pérez-Ríos, M., Ross, J. and Candal-Pedreira, C. (2024). 'Biomedical retractions due to misconduct in Europe: Characterization and trends in the last 20 years'. *Scientometrics*, 129, pp. 2867–2882. Available at: https://doi.org/10.1007/s11192-024-04992-7

The Guardian (2024). '"The situation has become appalling": Fake scientific papers push research credibility to crisis point'. *The Guardian*, 3 February. Available at: https://www.theguardian.com/science/2024/feb/03/the-situation-has-become-appalling-fake-scientific-papers-push-research-credibility-to-crisis-point

Hosseini, M., Horbach, S.P.J.M., Holmes, K. and Ross-Hellauer, T. (2025). 'Open science at the generative AI turn: An exploratory analysis of challenges and opportunities'. *Quantitative Science Studies*, 6, pp. 22–45. Available at: https://doi.org/10.1162/qss_a_00337

Kabel, S. (2025). *A bibliography of genAI-fueled research fraud from 2025*. Available at: https://sharonkabel.com/genai-fraud/

Moseley, H. (2024). 'In the AI science boom, beware: Your results are only as good as your data'. *Nature*, 1 February. Available at: https://www.nature.com/articles/d41586-024-00306-2

Neylon, C. (2015). 'Researchers as victim. Researchers as predator'. *Science in the Open* [blog post], 7 September. Available at: http://cameronneylon.net/blog/researcher-as-victim-researcher-as-predator/

NOR-CAM (2021). *A Toolbox for Recognition and Rewards in Academic Careers*. Report. Available at: https://www.uhr.no/en/_f/p3/i86e9ec84-3b3d-48ce-8167-bbae0f507ce8/nor-cam-a-tool-box-for-assessment-and-rewards.pdf

Nosek, B., Ebersole, C., DeHaven, A.C. and Mellor, D.T. (2018). 'The preregistration revolution'. *PNAS*, 115(11), pp. 2600–2606. Available at: https://doi.org/10.1073/pnas.1708274114

O'Grady, C. (2025). 'Scientific fraud has become an "industry," alarming analysis finds'. *Science*, 4 August. Available at: https://www.science.org/content/article/scientific-fraud-has-become-industry-alarming-analysis-finds

Olaizola, B. (2024). 'Elisabeth Bik, expert in scientific integrity: "We need to slow down scientific publishing"'. *El País*, 26 April. Available at: https://english.elpais.com/science-tech/2024-04-26/elisabeth-bik-expert-in-scientific-integrity-we-need-to-slow-down-scientific-publishing.html

Oransky, I. (2025). 'Retractions increase 10-fold in 20 years — and now AI is involved'. *AAPS Newsmagazine*, February. Available at: https://www.aapsnewsmagazine.org/articles/new-page3/feb25/meetings-feb25

Pizzolato, D. (2024). 'Bad apples or systematic problem? Is Italy struggling with maintaining high level of research integrity?'. *Accountability in Research*, pp. 1–6. Available at: https://doi.org/10.1080/08989621.2024.231823

Pölönen, J. and Mustajoki, H. (2022). 'European recommendations on responsible research assessment'. EASST Conference, Madrid, 6–9 July 2022. Available at: https://doi.org/10.5281/zenodo.6850824

Rahman, T. (2024). 'Our "publish-or-perish" culture is breaking the academy'. *The Brown Daily Herald*, 12 April. Available at: https://www.browndailyherald.com/article/2024/04/rahman-26-our-publish-or-perish-culture-is-breaking-the-academy

Room for Everyone's Talent Roadmap (2022). Available at: https://recognitionrewards.nl/wp-content/uploads/2023/04/Room-for-everyones-talent-in-practice-Road-map-Recognition-Rewards.pdf

Ross-Hellauer, T. (2023). 'Strategic priorities for reproducibility reform'. *PLOS Biology*, 21(1), e3001943. Available at: https://doi.org/10.1371/journal.pbio.3001943

Sakai, Y., Kamigaito, H. and Watanabe, T. (2026). *HalluCitation Matters: Revealing the Impact of Hallucinated References with 300 Hallucinated Papers in ACL Conferences*. ArXiv preprint. Available at: https://doi.org/10.48550/arXiv.2601.18724

Slifka, J. (2024). 'Who is a data steward'. *FAIR Wizard* [blog post]. Available at: https://fair-wizard.com/blog/who-is-a-data-steward

Thorpe, H. (2024). 'Universities need to stop hiding from research integrity problems'. *Science Forever* [blog post], March. Available at: https://holdenthorp.substack.com/p/universities-need-to-stop-hiding

EU AND NATIONAL POLICIES ON RESEARCH SECURITY: IMPLICATIONS FOR OPEN SCIENCE AND RESEARCHERS' FREEDOM

PAOLO VALENTE
National Institute for Nuclear Physics (INFN), Italy
ROBERTA VIGNI
Italian Institute for Environmental Protection and Research (ISPRA), Italy

Keywords: Research security, Open science, Researchers' freedom, Science and law, Open data

Considering the evolving geopolitical landscape and the importance of research and innovation in this context, it is increasingly necessary to implement effective measures to shield researchers and European research from security risks, particularly those linked to hybrid threats. Is it possible to safeguard the integrity and security of research while preserving, at the same time, the fundamental values and rights of the European Union, policies in support of open science, and academic freedom? This challenge is at the core of the Council Recommendation of 23 May 2024 (C/2024/3510) addressed to the Commission and the Member States on strengthening research security (hereinafter "the Recommendation") and the related measures.

European Union – Regulatory and strategic framework

Since 2021, European institutions have launched several reflections on the need to protect research conducted within the EU from foreign interference (summarized in points 8, 9 and 10 of the preamble to the Recommendation). The Recommendation is the starting point for the concrete planning of operational measures, and firstly it clarifies that "research security" refers to the anticipation and management of risks relating to:

a) undesirable transfer of critical knowledge and technologies that may compromise the security of the Union and its Member States, for instance if diverted toward military or intelligence uses in third countries;

b) malicious interference in research, which may result in its instrumentalization by third countries with the aim, among others, of generating disinformation or encouraging self-censorship among students and researchers, thereby violating academic freedom and research integrity within the Union;

c) breaches of ethics or integrity, in which knowledge and technologies are used to suppress, violate, or undermine the Union's fundamental values and rights as enshrined in the Treaties.

The importance of international cooperation in research and innovation remains central, in compliance with the principle "as open as possible, as closed as necessary," ensuring that research results are findable, accessible, interoperable, and reusable (FAIR). The Recommendation stresses that responsibility for international cooperation in research and innovation lies with the organizations conducting

research activities and that academic freedom cannot be separated from the assumption of academic responsibility.

These principles form a framework aimed at ensuring a comprehensive approach to research security, promoting coherence in governmental actions and communication towards the research and innovation sector, fostering self-governance through training measures, and ensuring coordinated governmental support, including with intelligence agencies.

National initiatives (Italy)

Italy, as an EU Member State, has followed up on the Recommendation by establishing, under the coordination of the Ministry of Universities and Research and with the contribution of the National Cybersecurity Agency, an intergovernmental working group, which also included representatives from the Conference of Rectors of Italian Universities and the Council of Presidents of Public Research Bodies. The outcome of this work led, in August 2025, to the publication of the National Model (featuring both a national and a decentralized level) and the Guidelines for Research Integrity and Security, which aim to support research institutions in operational implementation.

Across these documents, it is evident that the entire framework centers on research project leaders, who, following a guided process and before launching their activities, are required to carry out a self-assessment on three potential areas of concern relating to:

a) the scientific or technological field of the activity;

b) the collaborating individuals and institutions;

c) the entities funding the activities.

Whenever the self-assessment reveals potential risks, consultation with higher levels is recommended.

Research data

Article 10 of Directive (EU) No. 1024/2019 (Open Data) defines research data as follows:

Member States shall support the availability of research data by adopting national policies and relevant actions aiming at making publicly funded research data openly available ('open access policies'), following the principle of 'open by default' and compatible with the FAIR principles. (...)

This provision has strengthened a European and national regulatory framework aimed at ensuring openness of data by default, with a limited number of well-defined exceptions.

For research data, this process has been supported by specific obligations included in programmes such as Horizon 2020, Horizon Europe, and the PNRR (FAIR data), and by dedicated infrastructures and tools (EOSC, ORE).

Conclusions

The Recommendation, and consequently the Italian national model, assigns research project leaders and scientific and academic institutions the responsibility of assessing research security risks, based on the principle that *freedom of research entails the assumption of academic responsibility*. However, the areas in which this responsibility is exercised extend beyond the academic domain, producing effects in the ethical, legal, social, economic and other spheres.

Furthermore, the rules and conditions for accessing public research resources, currently under discussion within European institutions (the new Horizon Europe programme, the ERA Act), inevitably influence both security and freedom of research.

Regarding FAIR data, the guiding principle of data openness, *"as open as possible, as closed as necessary"*, allows flexibility and adaptation to changing needs, including security needs. However, if the push for data openness, which is primarily driven in the EU regulation by the goal of generating economic development through the reuse of open data, is constrained by overriding security and confidentiality requirements, it may lose momentum and effectiveness.

In international law, as well as Italian constitutional case law, a principle has emerged, especially concerning the protection of the environment and public health, summarized by the expression *"scientific reserve"*. This principle holds that the absence of a solid scientific basis for a regulatory decision renders it unconstitutional, and that such a scientific basis must be ensured by official and reliable "reference data." This requires the scientific community to be capable of asserting its authority based on its integrity.

Those working in Open Science are well aware that the ethical conduct of researchers cannot be taken for granted. In light of systemic biases in publication, evaluation, and funding—exacerbated by the integration of AI—research integrity must be proactively defended.

We are caught in precarious balancing act between the demand for policy legitimacy and the preservation of foundational democratic values and academic freedom.

References

Council of the European Union (2024). *Council Recommendation of 23 May 2024 on Enhancing Research Security*(C/2024/3510). Available at: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32024H03510

Corte costituzionale (IT), sentenze: 26 giugno 2002, n. 28; 15 maggio 2015, n. 83; 15 novembre 2017, n. 240; 18 gennaio 2018, n. 5.

Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on Open Data and the Re-use of Public Sector Information. Available at: https://eur-lex.europa.eu/eli/dir/2019/1024/oj

ECtHR, 9 April 2024, No. 53600/20, *Verein KlimaSeniorinnen Schweiz and Others v Switzerland*.

Lobascio, C. (2024). 'Tabella riassuntiva dell'evoluzione del diritto europeo dei dati e delle piattaforme'. *Rivista Italiana di Informatica e Diritto*, 5(2), pp. 35–40. Available at: https://doi.org/10.32091/RIID0132

Ministero dell'Università e della Ricerca (MUR) (n.d.). *Modello Nazionale e Linee Guida per l'Integrità e la Sicurezza della Ricerca*. Available at: https://www.sicurezzaricerca.mur.gov.it/

Tallacchini, M. (2023). 'Scienza e potere', in Cartabia, M. and Ruotolo, M. (eds.), *Enciclopedia del Diritto. Volume: Potere e Costituzione*. Giuffrè Editore, Milano, pp. 1059–1095. Available at: https://hdl.handle.net/10807/259786

United Nations (1992). *United Nations Framework Convention on Climate Change (UNFCCC)*. United Nations, New York.

United Nations (1994). *United Nations Convention on the Law of the Sea (UNCLOS)*. United Nations, New York.

Wewerinke-Singh, M. (2025). 'Climate protection obligations under the European Convention on Human Rights: The KlimaSeniorinnen judgment: ECtHR 9 April 2024, No. 53600/20, Verein KlimaSeniorinnen Schweiz and Others v Switzerland'. *European Constitutional Law Review*, 21(2), pp. 356–374. Available at: https://doi.org/10.1017/S1574019625100576

3.2 BIBLIOMETRICS, LLMs AND RESEARCH INTEGRITY

TOPIC CONCENTRATION IN SSH FUNDING PORTFOLIOS: A LARGE LANGUAGE MODEL ASSISTED QUANTITATIVE AND QUALITATIVE ANALYSIS

ALESIA A. ZUCCALA
Department of Communication, University of Copenhagen

Keywords: Research funding portfolios, Topic concentration, Social Sciences and Humanities, Large Language Models

Introduction

Concerns about epistemic diversity, innovation, and path dependency have long shaped debates in science policy and science and technology studies (STS). Although research funding systems are normatively expected to foster pluralism, empirical research shows that funding is highly concentrated, with a small elite of researchers and institutions capturing a disproportionate share of resources across countries and disciplines (Aagaard et al., 2020; Larivière et al., 2010; Madsen & Aagaard, 2020). This concentration is reinforced by cumulative advantage dynamics, whereby early funding success increases future funding prospects largely independent of subsequent research performance (Bol et al., 2018; Liao, 2021). At the same time, the link between funding concentration and scientific output appears weak: highly funded researchers do not consistently outperform their peers, and additional funding often yields diminishing marginal returns (Larivière et al., 2010; Mongeon et al., 2016). Beyond individuals, funding systems have been shown to follow established and highly visible areas of research. Klavans and Boyack (2017) demonstrated that topic "prominence"; measured as visibility and attention, strongly predicts future funding allocations, thereby reinforcing thematic lock-in and potentially constraining epistemic diversity (Yin et al., 2017; Madsen & Aagaard, 2020).

This paper presents an exploratory analysis of topic concentration in a large corpus of funded SSH research projects (n = 1,911) using a Large Language Model, ChatGPT 5.2. Treating the model as an assistive instrument, the paper makes three contributions: it maps funded SSH fields in Denmark from 2016 to 2024, provides a thematic and qualitative inductive analysis of the contemporary funding landscape, and reflects on the affordances and limits of LLMs for research policy analysis.

Method

Data were extracted from the "Grants and Funders" database of Research Portal Denmark (2026) and focused exclusively on project descriptions within the Social Sciences and Humanities. Research Portal Denmark currently contains data from five private funders and one Danish public funder as well as metadata on grants awarded since 2016 to both Danish and international recipients. At the Grants and Funders database, the following query was used:

- Grant Year = (2024 OR 2023 OR 2022 OR 2021 OR 2020 OR 2019 OR 2018 OR 2017 OR 2016)
- AND Danish Funders= (Independent Research Fund Denmark OR Novo Nordisk Foundation OR Carlsberg Foundation OR Lundbeck Foundation OR Villum Foundation OR VELUX FOUNDATION)
- AND Danish Organisations= (All Universities)
- AND OECD= (All Social Sciences OR All Humanities and the Arts)
- AND Statistics Denmark= (All Social Sciences OR All Humanities and the Arts)
- AND Open Competition Grant=(Yes)

From the initial exported data, two types of working datasets were prepared for import to ChatGPT. One was an Excel worksheet, which included: a) Grant IDs; b) Level 2 OECD field classifications, c) the number of records categorized according to OECD Level 2; d) the normalized weight of each OECD field, e) the title of each grant proposal, and f) the grant proposal abstracts[1]. A second working dataset, compiled as a Word file, contained an unformatted version of all funding portfolio titles and abstracts. Using an iterative prompt design, we asked ChatGPT to assist us in preparing the following:

1. a field map based on the OECD Level 2 classifications, titles, and abstracts
2. a thematic analysis of the combined textual content of the funded project titles and abstracts.
3. a qualitative, inductive analysis based on the unformatted titles and abstracts

Results

1. Field Map. The first prompt design resulted in Figure 1. Words from both the granted project titles and abstracts were processed using TF–IDF weighting, and all OECD level 2 field classifications were represented by centroid vectors capturing their dominant research language. Cognitive distances between the fields were calculated using cosine distance and projected into two dimensions using multidimensional scaling. In parallel, field-level diversity was assessed using the Shannon diversity component ($-p_i \log p_i$).

Each bubble in Figure 1 represents a research subfield (OECD Level 2), with bubble size indicating how frequently it has been funded, while the distance between bubbles reflects subfield similarities or differences in research language, methods, concepts, and problem framings. The OECD subfields positioned closer together draw on similar research approaches, whereas those farther apart represent substantively different ways of doing research. The colour of each bubble indicates the subfield's contribution to overall portfolio diversity, with warmer colours (yellow/orange) signalling a stronger contribution to balance and cooler colours (green) signalling a weaker contribution.

[1] 223 Abstracts were missing or labelled as N/A in the database.

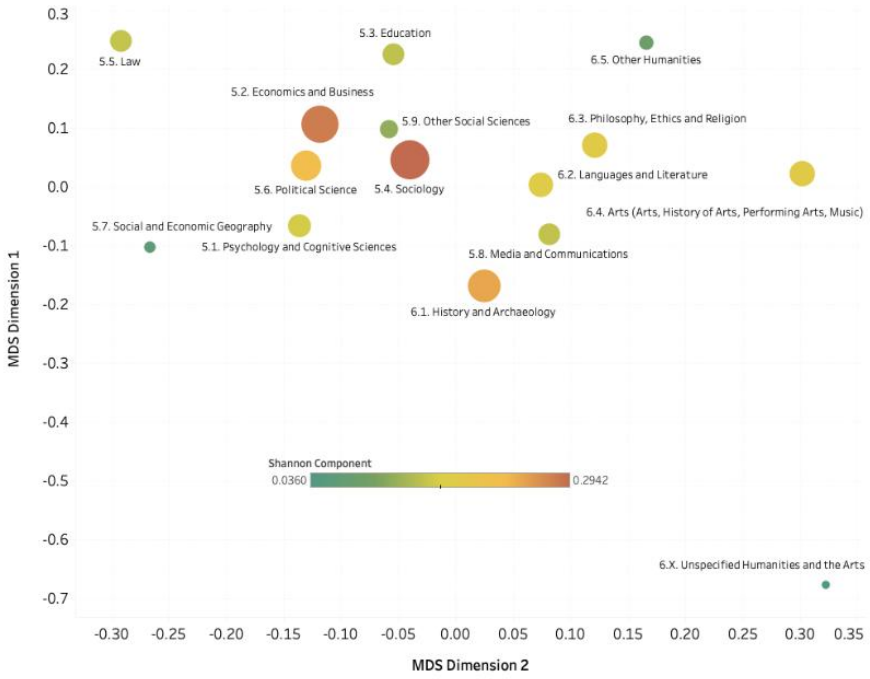

Figure 1. Map of Funded Research Fields in the Social Sciences and Humanities
(Source: ChatGPT 5.2 and Tableau desktop 2024.2)

2. Thematic Analysis. The second prompt design resulted in Tables 1 and 2, below. Table 1 shows that the corpus is highly concentrated in that nearly half of all projects situate themselves in digitalised social contexts, especially work and welfare. The heatmap, in Table 2 shows that funded societal challenge areas are predominantly addressed through qualitative, practice-based, and ethically inflected approaches, while explicit policy implementation methods remain secondary across the portfolio.

Table 1. Most Recurring Topic Areas (theory-driven coding)

Topic Area	Projects	% of Corpus	Interpretation
Digitalisation & AI-in-Society	877	45.9%	Digitalisation is the dominant *context*, not a technical focus
Work, Labour & Organisation	777	40.7%	Strong concentration on work, balance, and organisational life
Welfare, Health & Care	303	15.9%	Core concern with welfare institutions and citizen services
Ethics, Meaning & Experience	272	14.2%	Consistent humanistic and phenomenological orientation
Education & Learning	249	13.0%	Education as an applied societal domain

Table 2. Topic Concentration and Dominant Methodological Orientations (implicit)

Colour legend

- ▨ Strong association
- ▨ Moderate association
- ▨ Weak association
- ☐ Minimal / rare association

Topic Area (ordered by concentration)	Approx. % of Portfolio	Qualitative / Interpretive	Practice- & Case-oriented	Ethics / Normative	Policy / Governance / Implementation
Social inequality, vulnerability & inclusion	≈45–50%	▨	▨	▨	▨
Health, mental health & wellbeing	≈35–40%	▨	▨	▨	▨
Digitalisation, AI & data in society	≈45–46%	▨	▨	▨	▨
Climate change & sustainability	≈25–30%	▨	▨	▨	▨
Democracy, politics & governance	≈20–25%	▨	▨	▨	▨
Institutions in transformation	≈20–25%	▨	▨	▨	▨
Culture, identity & memory	≈18–22%	▨	▨	▨	☐
Education & socialisation	≈13–15%	▨	▨	▨	▨
Work, labour & economy	≈35–40%	▨	▨	▨	▨
Migration, mobility & borders	≈15–20%	▨	▨	▨	▨

Topic Area (ordered by concentration)	Approx. % of Portfolio	Qualitative / Interpretive	Practice- & Case-oriented	Ethics / Normative	Policy / Governance / Implementation
Narratives, imaginaries & futures	≈10–12%	▩	▨	▨	☐
Human–technology relations (beyond AI)	≈8–10%	▩	▨	▩	☐
Colonialism & global justice	≈8–10%	▩	▨	▩	☐

3. Qualitative Inductive Analysis. Table 3 presents the result of a third iterative prompt where ChatGPT was asked to "assume the role of a qualitative researcher" and "identify or classify as many topics to determine if there is concentration or repetition of topics funded in the Humanities and Social Sciences"

Table 3. Comparative Thematic Table Ordered by Area of Concentration (Highest → Lowest)

Rank	Thematic Area	Humanities Emphasis	Social Sciences Emphases	Area of Concentration
1	Social inequality, vulnerability & inclusion	Ethics; lived experience; stigma; marginalisation narratives; historical roots of inequality	Structural inequality; welfare systems; policy effects; institutional outcomes	Very high (central hub)
2	Health, mental health & wellbeing	Meaning of illness; recovery narratives; phenomenology; ethics of care	Health systems; service delivery; inequalities; interventions	Very high
3	Digitalisation, AI & data in society	AI as cultural interface; ethics; responsibility; creativity; meaning	AI governance; regulation; labour markets; education; welfare technologies	High and rapidly growing
4	Climate change, environment & sustainability	Environmental humanities; narratives; ethics; colonialism; human–nature relations	Energy transition; political economy; policy design; distributional effects	High
5	Democracy, politics & governance	Normative theory; discourse; legitimacy; political imaginaries	Institutions; participation; policy processes; media effects	High
6	Institutions in transformation(welfare, law, culture)	Legitimacy; ethics; symbolic change	Institutional design; reforms; unintended consequences	High (cross-cutting)
7	Culture, identity & memory	History; language; literature; art; postcolonial critique	Identity as contextual variable shaping outcomes	Moderate–high (Humanities-heavy)
8	Education & socialisation	Meaning of learning; identity; linguistic and cultural practices	Educational systems; inequality of opportunity; policy reforms	Moderate

Rank	Thematic Area	Humanities Emphasis	Social Sciences Emphases	Area of Concentration
9	Work, labour & economy	Meaning of work; ethics; critique of capitalism	Labour markets; gig economy; organisational change	Moderate
10	Migration, mobility & borders	Narratives of belonging; hospitality; identity; memory	Integration outcomes; border regimes; migration policy	Moderate
11	Narratives, imaginaries & futures	Anticipation; storytelling; imagination; uncertainty	Expectations; behaviour; policy uptake	Moderate–low but growing
12	Human–technology relations (beyond AI)	Phenomenology; embodiment; aesthetics; moral responsibility	Technology adoption; impacts on work and welfare	Moderate–low
13	Colonialism, postcolonialism & global justice	Historical depth; cultural critique; memory	Contextual factor shaping inequality and governance	Lower but consistent

Discussion

The map in Figure 1 shows that while while some OECD fields have received funding more frequently than others, they are not clustered around a single cognitive core. Instead, they occupy distinct regions of the research space, reflecting substantive differences in research language, methods, and problem framings rather than mere disciplinary labels. This interpretation is supported by the relatively high Shannon diversity score (Shannon entropy = 2.49), indicating that repeated funding of certain topics coexists with a broadly plural and balanced portfolio in which medium-sized research areas contribute significantly to overall diversity.

At the same time, both the thematic and qualitative inductive analyses (Tables 1, 2 and 3) reveal clear patterns of concentration. Across funders, the portfolio consistently centres on a small set of interrelated societal challenges (i.e., social inequality, health and wellbeing, and digitally mediated work and welfare) addressed primarily through qualitative, practice-oriented, and ethically inflected research. The methodological profile (Table 2), in particular, shows a comparatively weak orientation toward governance and implementation, suggesting that funded social science and humanities projects in Denmark articulate societal relevance less through direct policy instruments than through interpretive, critical, and agenda-setting forms of inquiry. For policy makers, this highlights the importance of aligning expectations of research impact with the epistemic strengths of SSH research, and of recognising reflexive and agenda-setting contributions alongside more instrumental forms of relevance.

Emerging themes such as artificial intelligence, climate change, and democratic governance build outward from the core topics, extending established SSH agendas rather than displacing them. Overall, the findings point to a funding landscape characterised not by topical dispersion, but by structured concentration around socially consequential problems, combined with epistemic diversity, interdisciplinarity, and a continued emphasis on reflexive rather than instrumental modes of engagement.

Conclusions

Overall, the findings indicate that responsiveness to major societal challenges coexists with structured thematic concentration, even within ostensibly bottom-up funding systems. While such concentration is often assumed to undermine epistemic diversity,

the analysis demonstrates that diversity can be maintained at the level of research language, methods, and problem framings. This suggests that portfolio balance cannot be assessed through field/subfield counts or disciplinary categories alone, but requires attention to cognitive and methodological diversity across funded research.

To address anticipated skepticism toward the use of ChatGPT, we emphasize that the model is employed to augment rather than replace expert judgement in large-scale qualitative analysis. The paper thus contributes to ongoing STS and RESSH debates on large language models as tools for critical inquiry into science systems, shifting attention from whether LLMs "get it right" to what they render visible, for whom, and under what conditions. Looking ahead, extending this approach to comparisons between funded research and the broader publication landscape would support more transparent funding strategies, while helping to identify both thematic lock-in and underexplored areas within SSH portfolios.

References

Aagaard, K., Kladakis, A. and Nielsen, M.W. (2020). 'Concentration or dispersal of research funding?'. *Quantitative Science Studies*, 1(1), pp. 117–149. Available at: https://doi.org/10.1162/qss_a_00002

Bol, T., De Vaan, M. and Van De Rijt, A. (2018). 'The Matthew effect in science funding'. *Proceedings of the National Academy of Sciences*, 115(19), pp. 4887–4890. Available at: https://doi.org/10.1073/pnas.1719557115

Klavans, R. and Boyack, K.W. (2017). 'Research portfolio analysis and topic prominence'. *Journal of Informetrics*, 11(4), pp. 1158–1174. Available at: https://doi.org/10.1016/j.joi.2017.10.002

Larivière, V., Macaluso, B., Archambault, É. and Gingras, Y. (2010). 'Which scientific elites? On the concentration of research funds, publications and citations'. *Research Evaluation*, 19(1), pp. 45–53. Available at: https://doi.org/10.3152/095820210X492495

Liao, C.H. (2021). 'The Matthew effect and the halo effect in research funding'. *Journal of Informetrics*, 15(1), 101108. Available at: https://doi.org/10.1016/j.joi.2020.101108

Madsen, E.B. and Aagaard, K. (2020). 'Concentration of Danish research funding on individual researchers and research topics: Patterns and potential drivers'. *Quantitative Science Studies*, 1(3), pp. 1159–1181. Available at: https://doi.org/10.1162/qss_a_00077

Mongeon, P., Brodeur, C., Beaudry, C. and Larivière, V. (2016). 'Concentration of research funding leads to decreasing marginal returns'. *Research Evaluation*, 25(4), pp. 396–404. Available at: https://doi.org/10.1093/reseval/rvw007

Research Portal Denmark (2026). *Grant search results*. Available at: https://grants.forskningsportal.dk/search/

Yin, Z., Liang, Z. and Zhi, Q. (2018). 'Does the concentration of scientific research funding in institutions promote knowledge output?'. *Journal of Informetrics*, 12(4), pp. 1146–1159. Available at: https://doi.org/10.1016/j.joi.2018.09.003

CAN LARGE LANGUAGE MODELS VALIDLY IDENTIFY COMPARABLE PEERS FOR RESEARCHER BENCHMARKING?

SHIR AVIV-REUVEN
Bar-Ilan University, Israel
LUTZ BORNMANN
Administrative Headquarters of the Max Planck Society, Germany
ARIEL ROSENFELD
Bar-Ilan University, Israel

Keywords: Large Language Models, Peer identification, Scholarly benchmarking

Background and motivation

In evaluative bibliometrics and research management, valid assessment depends on comparing "like with like." Benchmarking researchers against meaningful peers is central to processes such as promotion, grant allocation, and career evaluation. However, identifying truly comparable peers remains difficult. Manual peer selection is time-consuming and susceptible to selection bias, while automated approaches typically require extensive bibliometric metadata and specialized analytical expertise (Andersen et al., 2017; Cormode et al., 2014).

Recent work suggests that large language models (LLMs) may support benchmarking tasks by inferring scholarly similarity from minimal textual input (Bornmann & Lepori, 2024; Lepori et al., 2025), but evidence at the level of individual researchers remains limited.

This study examines whether LLMs can provide a valid and scalable alternative to metadata-intensive peer identification by testing whether their suggested peers align with established bibliometric indicators of comparability.

Data

The analysis draws on a large dataset of computer science researchers compiled from the Digital Bibliography & Library Project (DBLP) and enriched with metadata from OpenAlex. DBLP provides extensive coverage of computer science publications (Rosenfeld, 2023), while OpenAlex offers metadata on authors, citations, institutions, and research topics (Priem et al., 2022).

After filtering for scholars with at least five publications, the dataset comprised about 308,000 researchers from 112 countries. Extracted attributes included research topics, publication counts, mean citation counts, institutional affiliation, country, and academic age inferred from career start and end years.

Sample construction

Running the benchmarking pipeline across the full dataset is computationally expensive, particularly for the LLM-based approach. Hence, we drew a representative sample of 2,000 scholars. Representativeness was verified using Bonferroni-corrected

chi-square tests comparing the sample with the population on research topics, country, gender, and binned publication and citation counts.

Methods

We evaluated three approaches for identifying peer researchers.

LLM-based approach. Anthropic's Claude model was prompted with minimal structured information about each focal researcher: name, institutional affiliation, academic age, and a subset of research topics. To mitigate variability, the prompt was executed five times per scholar, and peers appearing most frequently across runs were retained.

Metadata-based. A similarity model was constructed. Each scholar was represented by a feature vector including academic age, country, gender similarity, and topic similarity measured via a Jaccard index. Cosine similarity between scholars was then calculated, and the ten most similar researchers were selected as benchmark peers.

Random baseline. A simple random sample of ten peers from the dataset served as a baseline.

Evaluation

Following prior work on benchmarking units in research evaluation (Bornmann & Marx, 2014; Wang & Jeppsson, 2022), we assessed the comparability of suggested peers across five dimensions:

- Collaboration proximity (co-authorship distance)
- Academic age
- Research topic similarity
- Publication output
- Citation impact

These metrics capture both structural proximity within research communities and performance-based indicators of scholarly activity.

Results

Across most dimensions, the LLM-based approach produced peers comparable to those identified by the structured Metadata-based model.

Both the LLM-based and Metadata-based approaches produced peers that were relatively close in the collaboration network. As shown in Table 1, the average collaboration distance is approximately 3.5 for the LLM approach and 3.3 for the Metadata-based approach, indicating that suggested peers are typically only a few collaboration steps away from the focal scholar. In contrast, the random baseline yields more distant peers, with an average distance of about 4.4. Similarly, about 74% of scholars have LLM-identified peers within a collaboration distance of four or less, compared with roughly 80% for the metadata-based model and only about 24% for the random baseline. Notably, since collaboration data was not included in the LLM prompt, this suggests the LLM may implicitly capture structural relationships within research communities. Top of FormBottom of Form

Table 1: Descriptive statistics for selected collaboration distance metrics

Approach	Statistic	Minimum	Maximum	Average
LLM	Mean	2.6	4.5	3.5
	Standard deviation	1.2	1.1	1.0
	Median	3.0	5.0	4.1
Metadata-based Model	Mean	2.1	4.5	3.3
	Standard deviation	1.1	1.1	0.8
	Median	2.0	4.0	3.3
Random	Mean	3.5	5.3	4.4
	Standard deviation	0.6	0.8	0.4
	Median	4.0	5.0	4.3

The relationship between the academic age of focal scholars and their benchmark peers is shown in Figure 1. For both the LLM-based and metadata-based approaches, differences in academic age are centered near zero, indicating that peers are typically at a similar career stage as the focal scholars. In contrast, the random baseline often selects more senior peers. This suggests that the LLM effectively incorporates career-stage information from the prompt.

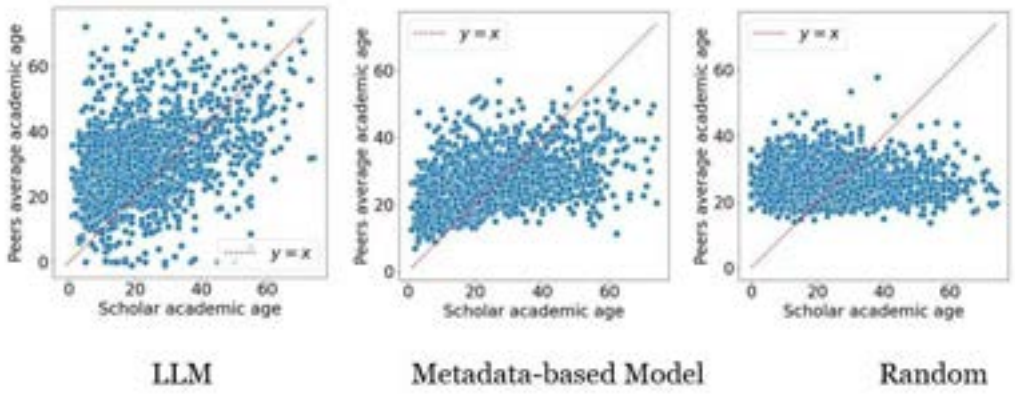

Figure 1. Scatter plot of scholar's academic age versus mean publication counts of benchmark peers

Table 2 presents the topic similarity between focal scholars and their benchmark peers. Topic alignment is high for all methods, indicating that most suggested peers work in closely related research areas. The metadata-based model shows the highest average overlap, while the LLM-based approach yields slightly lower but still strong similarity scores. The random baseline shows comparable though somewhat lower similarity, likely reflecting the topical homogeneity of the dataset.

Table 2: Descriptive statistics for Jaccard similarity measures

Approach	Statistic	Average Jaccard similarity
LLM	Mean	0.81
	Standard deviation	0.11
	Median	0.83
Metadata-based Model	Mean	0.87
	Standard deviation	0.05
	Median	0.88
Random	Mean	0.8
	Standard deviation	0.05
	Median	0.82

A notable difference emerges in publication output. As shown in Figure 2, the metadata-based model identifies peers whose publication counts closely match those of the focal scholars, with most observations near the identity line. In contrast, the LLM-based approach shows a systematic upward bias, frequently recommending peers with higher publication counts. The random baseline shows no clear alignment, with peer productivity largely reflecting the overall population distribution.

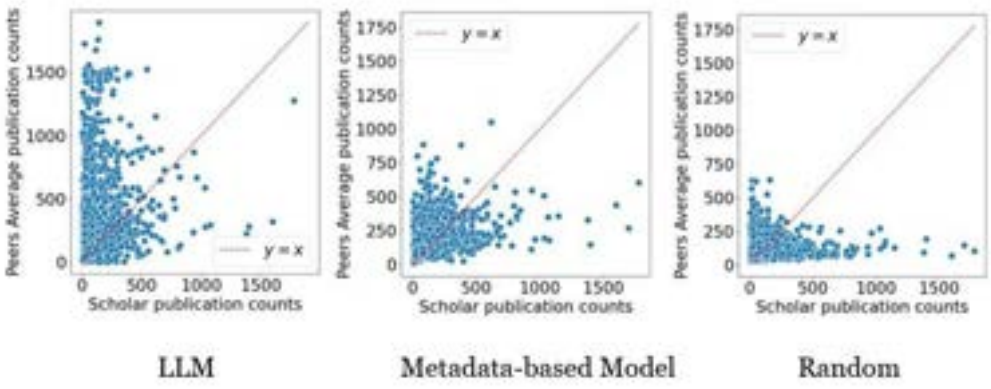

Figure 2 Scatter plot of scholar's publication counts versus mean publication counts of benchmark peers

The distribution of differences in mean citation impact between focal scholars and their benchmark peers is shown in Figure 3. Both the LLM-based and metadata-based approaches produce very similar patterns, with distributions centered near zero. This indicates that the peers identified by both methods have citation impact levels comparable to those of the focal scholars.

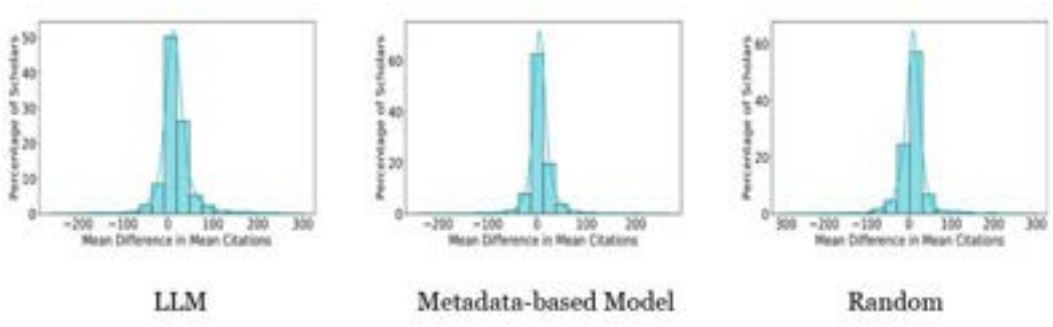

Figure 3 Distribution of mean differences in mean citations between scholars and their benchmark peers

Implications and operational protocol

The results indicate that LLMs can approximate structured bibliometric similarity models when identifying comparable scholars, Suggesting that LLM-based approaches may serve as a scalable and accessible starting point for peer identification in research evaluation contexts.

To operationalize this approach, we propose a simple protocol:

1. Define the benchmarking context.
2. Provide structured prompts.
3. Run multiple prompt variants.
4. Apply post-hoc filtering.
5. Audit candidate peers.

This workflow retains the efficiency of LLMs while adding safeguards that improve reliability.

Conclusion

This study provides empirical evidence that LLMs can identify comparable researchers for benchmarking purposes with performance close to that of structured metadata-based similarity models. Across most evaluation metrics, including topic similarity, citation impact, collaboration proximity, and academic age, LLM-generated peers were strongly aligned with traditional bibliometric benchmarks.

Rather than replacing bibliometric tools, LLMs appear best suited as a complementary, lower-cost method for generating candidate peer sets for further validation.

References

Andersen, J.P., Didegah, F. and Schneider, J.W. (2017). 'The necessity of comparing like with like in evaluative scientometrics: A first attempt to produce and test a generic approach to identifying relevant benchmark units', in *Proceedings of the 22nd International Conference on Science, Technology and Innovation Indicators (STI 2017)*. Paris, France, 6 September 2017.

Bornmann, L. and Lepori, B. (2024). 'The use of ChatGPT to find similar institutions for institutional benchmarking'. *Scientometrics*, 129(6), pp. 3593–3598. Available at: https://doi.org/10.1007/s11192-024-05039-7

Bornmann, L. and Marx, W. (2014). 'How to evaluate individual researchers working in the natural and life sciences meaningfully? A proposal of methods based on percentiles of citations'. *Scientometrics*, 98(1), pp. 487–509. Available at: https://doi.org/10.1007/s11192-013-1161-y

Cormode, G., Muthukrishnan, S. and Yan, J. (2014). 'People like us: Mining scholarly data for comparable researchers', in *Proceedings of the 23rd International Conference on World Wide Web*. pp. 1227–1232. Available at: https://doi.org/10.1145/2567948.2579038

Lepori, B., Bornmann, L. and Gay, M. (2025). 'Finding "similar" universities using ChatGPT for institutional benchmarking: A large-scale comparison of European universities'. *Journal of the Association for Information Science and Technology*, 76(9), pp. 1174–1187. Available at: https://doi.org/10.1002/asi.25010

Priem, J., Piwowar, H. and Orr, R. (2022). 'OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts'. arXiv preprint arXiv:2205.01833.

Rosenfeld, A. (2023). 'Is DBLP a good computer science journals database?'. *Computer*, 56(3), pp. 101–108. Available at: https://doi.org/10.1109/MC.2022.3181977

Wang, Q. and Jeppsson, T. (2022). 'Identifying benchmark units for research management and evaluation'. *Scientometrics*, 127(12), pp. 7557–7574. Available at: https://doi.org/10.1007/s11192-022-04413-7

EVALUATING A MACHINE-LEARNING BASED REVIEWER MATCHING TOOL ACROSS DISCIPLINES: EARLY EVIDENCE OF FIELD-SENSITIVE PERFORMANCE AND IMPLICATIONS FOR SSH

JANNE NIEMI
Research Council of Finland, Finland

Keywords: Reviewer matching, Machine learning, Disciplinary differences, Becher's typology, Research evaluation, Social Sciences and Humanities

Context and aim

Research funding organisations are under pressure to process increasing numbers of applications while maintaining fairness, consistency, and transparency. Machine-learning tools have therefore become an attractive way to support review processes. At the Research Council of Finland (RCF), this development is part of a broader effort to integrate algorithmic analysis tools into a unified platform (AIbox), supported by Sitra and aligned with ongoing reform initiatives such as the CoARA Agreement on Reforming Research Assessment. The goal is to make the most of computer-assisted text analysis to understand thematic patterns in applications, anticipate reviewer expertise needs, and streamline reviewer allocation, while respecting principles of responsible evaluation. The AIbox project has demonstrated that such tools can reduce manual workload and offer consistent support to staff, but they also raise questions of reliability, transparency and disciplinary fairness.

In this contribution, we focus on one text-based reviewer matching tool used at RCF. Our analytical data consist of the tool-generated similarity values paired with reviewer behaviour data: the rates assigned to applications, reviewers' self-assessed expertise, and the assignment decisions made by RCF staff. We analyse the RCF Winter Call 2025, a large bottom-up funding round with over 2,800 submitted research plans. Applications were submitted to 41 review panels across the natural sciences and engineering, social sciences and humanities, and biosciences, health and the environment. A total of 680 experts participated in the review, most of whom contributed both as individual reviewers and as panel members.

Our aim is to understand the tool's alignment with human judgment and to identify areas in which performance varies between disciplines. Earlier studies (such as those presented in Bouzón-Arnáiz et al. 2025) have noted that machine-learning tools often perform better in natural sciences and engineering than in humanities and social sciences. Our findings support this overall observation but also show that a more nuanced disciplinary analysis adds substantial value.

Data and method

The tool analyses application texts and compares them with reviewer expertise represented by keywords provided either by the reviewers themselves or by RCF staff. For each application–reviewer pair, it generates a similarity value based on cosine similarity, which measures how close the two text-based representations (vectors) are.

Higher similarity indicates stronger thematic alignment between the application and the reviewer's keywords.

In our data, raw similarity values vary systematically across fields. Panels in the social sciences and humanities (SSH) tend to show higher raw scores, not because of better matches but because SSH reviewers typically list more—and more heterogeneous—keywords due to less standardised disciplinary vocabulary. Longer keyword lists increase the likelihood of textual overlap and thus inflate raw cosine similarity. To address this, we normalise similarity scores within each reviewer, scaling scores by the reviewer's own average. This retains the relative ordering of applications for each reviewer while reducing inflation effects and allowing more meaningful comparison across panels.

We employ Becher's typology describing disciplinary differences in scientific cultures, presented in four axes: Hard–Soft (from tightly codified, consensus-driven, technical knowledge to interpretive, plural and context-dependent inquiry), Pure–Applied (from theory and curiosity-driven knowledge to use-oriented, problem-solving knowledge in application contexts), Convergent–Divergent (from shared, narrowing problem agendas and methods to multiple, competing schools and expanding agendas), and Urban–Rural (from dense, fast-moving, boundary-crossing "urban" research ecologies to more stable, inward-looking, slower-changing "rural" ones). This offers a more systematic and structured analysis than simple SSH vs. non-SSH categories, or arbitrary groupings of different SSH-fields.

While the typology can clearly indicate where most SSH fields tend to cluster, it also allows meaningful variation within SSH, rather than treating it as a monolithic group or relying on institutional classifications. Importantly, the same nuance applies to non-SSH fields: natural sciences and technology, for example, may both fall on the Hard side but differ in their degree of convergence, divergence, or contextual embeddedness. This makes the typology a useful tool for identifying where text-matching algorithms work well and where they face systematic challenges. In particular, degrees of codification and vocabulary standardisation vary systematically across these axes, which affects how readily textual features capture genuine expertise alignment and thus helps explain field-specific performance differences in matching.

Findings

Several preliminary findings emerge. First, similarity scores generally align with reviewer self-assessment: applications that reviewers mark as core to their expertise tend to show higher similarity than those marked as only somewhat within their expertise, and both are above cases where reviewers provide no self-assessment. This pattern appears across most typological categories and offers reassuring evidence that the tool captures meaningful aspects of technical fit.

Second, field differences are visible. Hard and Pure domains tend to show clearer separations between levels of expertise and higher average similarity values. Soft domains show lower overall similarity and narrower separations. While these results are descriptive and preliminary, they align with expectations about differences in vocabulary standardisation, conceptual heterogeneity, and the prominence of multilingual or book-based publication cultures. They suggest that SSH applications may be structurally more difficult for text-matching tools to handle, though we do not rule out other contributing factors.

Third, reviewer assignment decisions made by staff correspond closely to the similarity scores. Those ultimately assigned to review an application typically have above-average similarity, while non-assigned reviewers are typically below average.

This demonstrates how the tool already supports operational decision-making and contributes to consistency across panels.

Fourth, reviewer behaviour in rating the applications does not appear to depend strongly on similarity strength. Rates in the midrange and upper ranges do not show a consistent relationship with similarity values, suggesting that higher algorithmic similarity does not inflate reviewer rates.

In the Winter Call process, applications advance to the panel stage only if at least one of the external reviewers assigns an overall rate of 5 or 6. Rate 5 therefore functions as a threshold rating that determines whether an application proceeds to the second stage of review. Reviewers who classify an application as only somewhat within their expertise show a tendency to give the rate 5 more often than reviewers who consider it core to their expertise. This pattern appears across most disciplinary categories but is most consistent in Soft fields and in Divergent and Urban subtypes within Becher's typology. Hard fields show the tendency as well, though with more variation between subfields. This suggests that uncertainty-related escalation through the rate 5 threshold may be more common in epistemically heterogeneous areas.

Taken together, the results indicate that the tool works directionally as intended, but its performance varies across disciplinary environments. The patterns strengthen the case for developing SSH-sensitive enhancements, such as richer vocabularies, embeddings trained on SSH corpora, improved handling of methods and contextual information, and typology-specific calibration. At the same time, the results help identify where the tool already functions effectively, including in many non-SSH fields. By combining empirical observations with Becher's typology, we hope to support a more informed discussion on how algorithms can be refined to enhance fairness, transparency, and usefulness in the evaluation of diverse research fields.

References

Becher, T. and Trowler, P.R. (2001). *Academic Tribes and Territories: Intellectual Enquiry and the Cultures of Disciplines*. 2nd edn. SRHE / Open University Press, Buckingham and Philadelphia.

Bouzón-Arnáiz, I., Carbonell Cortés, C., Hagemeijer, A., Jorstad, A., Okasa, G., Thelwall, M., van den Berg, N. and Woods, H.B. (2025). *AI and Reviewer Matching in Research Funding: Three Case Studies*. Research on Research Institute. Available at: https://doi.org/10.6084/m9.figshare.30885269

PAPER MILLS: OBJECTS OF FRAUD AND MIRRORS OF MANAGERIAL RESEARCH GOVERNANCE

LAURA GIOVINAZZI
University of Milan, Italy

Keywords*:* Paper mills, Research evaluation, New public management

Modern technology allows for a series of potentially fraudulent activities which would have been unimaginable in the past. Paper mills are one of such cases: thanks to the intersection of a variety of technological objects such as the web, digital encryption and AI softwares, they profit by successfully fabricating and publishing manuscripts, selling authorships, and altering peer-reviews, undermining the basic principles of scientific research, namely its integrity and social impact. Existing research indicates that paper mill activity is widespread (Wittau et al., 2023; Candal-Pedreira et al., 2024; Abalkina et al.,2024; Abalkina et al.,2025; Sabel et al., 2026), with an increasing number of researchers using mills to publish or simply to publish more quickly. Sometimes, researchers are aware that they are engaging in fraudulent activity, but other times they are unaware victims of it (for example, when they get contacted via e-mail by fake journals). Studies carried out so far focus on finding and quantifying retractions and fake article contents, detecting and listing fake ("hijacked") journals which steal real journals' identities, and publishing all possible information on the phenomenon. Thanks to sleuths' and researchers' work, we currently have an idea of a variety of aspects on paper mills: the regional distribution, the most common practices (stealing journals' ISSN, faking reviews, etc.), the methods used for falsifying research text, data and images. At the same time, their work also urges scientific community as a whole to conduct more work on this topic. This is because it is difficult to concretely capture the true scale of the phenomenon in both quantitative and qualitative terms. Firstly, it is challenging to find a single, most representative, indicator of paper mill activity. For example, retractions count may be influence by other factors besides paper-mill activity. Secondly, paper mills can easily alter their practices or disappear from records once they are detected – especially if they use AI technology to refine practices of fabrication and/or strengthen anonymity.

In broader terms, paper mills also reflect the multifaceted political, economic and social factors which account for the rise and development of perverse incentives in research communities: they are yet another example of how there is a demand for faking science. It is not just researchers *writing* articles who are incentivised to recur to fraudulent practices, but also those researchers r*eviewing* articles. From this standpoint, paper mills are both objects of fraud, namely a fraudulent industry with a set of specific technologies and operational modalities, and a reflection of the consequences of a performance-oriented governance of research and higher education which is pushing research communities to undermine scientific integrity. This managerial character of research governance takes the form of new public-management-oriented policies, such as national-level performance-based funding (PBF) systems for higher education sectors (Jongbloed, 2001; Nobili & Turri, 2025). As resource allocation becomes performance-driven, research institutions mirror this

approach in their internal practices, creating a vicious cycle of incentives which reward publication venues at the cost of publications' research quality. Examples of this adaptation include cash-rewards for publications in indexed journals (Quan et al.,2017) and universities using the same formulae used by national evaluation agencies for distributing resources between university departments (CRUI, 2015). As this contribution discusses, evaluation systems have helped to transform bibliometric performance into an economic asset, whereby prestigious publication venues and quantities translate into higher chances of starting and continuing a research career. If we "reassemble" phenomena, following Latour (2005), we can draw connections between networks in the scientific environment and the technologies, such as platforms and indicators, which mediate research production and research behaviours. If we therefore consider both the material and social drivers of specific configurations of 'doing science', we can better understand how and why the current scientific ecosystem allows paper mills to operate and succeed. Building on the insights and key questions for addressing research on paper mills presented by researchers (Byrne et al, 2024; Wittau & Seifert, 2024), this contribution presents an overview of the challenges of studying paper mills. Secondly, it critically approaches the issue by investigating the relationship between paper-mill activity and performance incentives driving publications, considering the Chinese context as an example, where providing cash-rewards based on journal metrics and indexation is a frequent practice and where the majority of global paper-mill activity is frequently signalled (Quan et al., 2017). Thirdly, positive examples countering paper mills and quantitative evaluation are discussed. It is argued that a closer analysis of the presence and workings of performance rewarding and evaluation practices can shed more light on the perverse incentives leading to fraud in academia. In light of increasing pressures for rejecting the publish or perish dynamics, this is ever more crucial.

3.3 RETHINKING INDICATORS, METRICS AND EVALUATION TOOLS FOR SSH

COUNTING THE UNCOUNTABLE? BIBLIOMETRIC INFRASTRUCTURES, INTERDISCIPLINARITY, AND RESPONSIBLE RESEARCH EVALUATION: THE CASE OF CNR-ISPC

ANDREA DI RENZONI, COSTANZA MILIANI
National Research Council - Institute of Heritage Science (CNR-ISPC), Italy

Keywords: Responsible research assessment, Database coverage and comparability, Data disambiguation, Subject classification, Interdisciplinarity

Introduction

Bibliometric indicators have become a central component of research evaluation systems, shaping funding allocation, institutional strategies, and individual academic careers (Hicks et al. 2015). Their growing prominence rests on the promise of objectivity, comparability, and scalability. Within the Social Sciences and Humanities (SSH), however, their applicability and legitimacy remain deeply contested (Archambault & Larivière 2010; Gläser & Oltersdorf 2019): differences in publication formats, linguistic diversity, disciplinary traditions, and research goals challenge the assumption that quality and impact can be adequately captured through standardized quantitative proxies (Kulczycki et al. 2018).

These tensions lie at the core of current debates on Responsible Research Assessment, which call for evaluation practices that are transparent, context-sensitive, and aligned with the epistemic specificities of different research domains. The Coalition for Advancing Research Assessment (CoARA) – to which the National Research Council of Italy (CNR) is a signatory – explicitly promotes a more reflexive use of research indicators, particularly in SSH. Understanding how these infrastructures operate is a prerequisite for their responsible use.

This contribution examines how different bibliometric databases produce divergent representations of the same research institution – the Istituto di Scienze del Patrimonio Culturale (CNR-ISPC) – and assesses the consequences for evaluation in interdisciplinary SSH contexts. Specifically, it addresses: (1) how database coverage affects the quantification of research output; (2) how affiliation disambiguation influences attribution; (3) how subject classification shapes the visibility of interdisciplinary research; and (4) how the internal cognitive structure of the institute relates to its bibliometric visibility.

Case study

Research at ISPC extends well beyond the humanities to encompass STEM domains. Interdisciplinarity is essential to a holistic understanding of cultural heritage, it complicates bibliometric evaluation frameworks that presuppose disciplinary homogeneity and rely on standardized publication channels (Larivière & Gingras 2010; Rafols & Meyer 2010). The consequences are asymmetric. Research published in SSH-oriented journals becomes partially invisible regardless of its STEM

components, while interdisciplinary work appearing in multidisciplinary or STEM journals gains full visibility.

Problems in quantifying CNR-ISPC research output

Quantifying ISPC's scientific production highlights four interrelated categories of problems.

First, database scope and coverage vary substantially across bibliometric infrastructures. Commercial indexes such as Web of Science and Scopus adopt a journal-centred model that systematically disadvantages SSH fields, where monographs, edited volumes, and national-language publications play a central role. OpenAlex, by contrast, adopts a product-oriented approach aimed at comprehensive coverage, collecting outputs from multiple sources and web crawling. As a result, publication counts for ISPC differ markedly depending on the database consulted—raising the question of which representation should be considered authoritative for evaluation purposes.

Second, institutional affiliation disambiguation represents a critical source of uncertainty. Bibliometric databases rely on automated parsing of affiliation strings, which are often inconsistent or incomplete—a problem amplified in large organizations like CNR, where multiple hierarchical levels coexist (organization, department, institute, research unit). OpenAlex does not recognize ISPC as a standardized institutional entity: the institute can only be retrieved through raw affiliation strings, with no institutional normalization. Based on articles explicitly showing ISPC affiliation, OpenAlex retrieves 365 authors (including numerous un-disambiguated occurrences) against Scopus's 165. The median number of articles per researcher is 1 in OpenAlex and 6 in Scopus—a gap reflecting not only different coverage but divergent indexing and attribution criteria. Tellingly, the most-cited authors are those working in STEM disciplines, while many prolific SSH researchers remain nearly invisible.

Third, subject classification systems introduce further distortions. In Web of Science and Scopus, subject categories are assigned at the journal level and inherited by all articles regardless of content—particularly problematic for an interdisciplinary institute like ISPC. OpenAlex classifies individual outputs using machine-learning techniques, offering greater granularity but also opacity, with concerns about the interpretability and stability of algorithmically generated categories.

Finally, definitions of research quality differ across evaluation regimes. In Italy, national SSH assessment relies on curated lists of "scientific" and "Class A" journals produced by ANVUR, precisely because bibliometric indicators are considered inadequate for these fields. This creates a structural tension between database-driven and policy-driven notions of quality, further complicating the interpretation of bibliometric data.

Methods and analytical strategy

These problems are not randomly distributed across the institute. If bibliometric invisibility falls disproportionately on researchers whose work crosses disciplinary boundaries, then the very feature that defines ISPC's scientific identity becomes a structural disadvantage in evaluation. Assessing this requires going beyond database comparison: it demands a map of the institute's internal cognitive structure.

The study adopts a quantitative approach, complemented by manual verification of publication records against independently compiled reference lists. Bibliometric data are collected from Web of Science, Scopus, and OpenAlex using multiple query

strategies based on institutional affiliation, author identifiers, and subject filters. The analysis is anchored to a controlled sample of ISPC researchers representing different disciplinary domains, for each of whom an independently verified publication list serves as a reference, enabling the identification of missing outputs, misattributions, and classification discrepancies.

Measuring interdisciplinarity requires a dedicated analytical layer. As Marres and de Rijcke (2020) note, interdisciplinarity is not a stable attribute but a relational property emerging from the interplay between cognitive practices and evaluation infrastructures. Existing bibliometric indicators—diversity, entropy, betweenness centrality—capture different facets of interdisciplinary research (Stirling 2007; Leydesdorff & Rafols 2011), yet their consistency remains contested (Wang & Schneider 2020). We adopt a two-level strategy combining structural analysis of the institute's cognitive composition with individual bibliometric profiling.

Level 1 – Structural heterogeneity. We construct a dataset tracing each ISPC researcher's degree and PhD-granting institution, internal research group or laboratory, and self-identified ERC sector. Heterogeneity indices are computed on the distribution of disciplinary origins across the institute. A proximity network is then built on two complementary layers: links represent either shared disciplinary backgrounds or co-membership in the same research group or laboratory. This dual structure allows us to assess whether actual collaborative arrangements within the institute reflect—or cut across—researchers' disciplinary formation, and whether ISPC functions as an integrated interdisciplinary environment or as a juxtaposition of separate disciplinary clusters.

Level 2 – Individual publication profiles. For each researcher in the controlled sample, publications are assigned a disciplinary category based on journal classification (WoS/Scopus) or, where available, article-level classification (OpenAlex).

Integration. Overlaying individual profiles onto the structural network allows us to address three questions: Do researchers with heterogeneous training publish more interdisciplinarily? Does network position predict interdisciplinary output? Are there clusters where collective interdisciplinarity emerges despite individual specialisation? This approach enables a distinction between multidisciplinarity—competences juxtaposed at the institutional level—and genuine interdisciplinarity—integration at the individual or group level—a distinction often blurred in evaluation exercises.

Expected outcomes and relevance

We expect to show that database coverage systematically favours certain disciplinary clusters over others; that interdisciplinarity—often promoted as a policy goal—becomes a liability when evaluation relies on discipline-bound indicators; and that the internal network structure of the institute mediates its bibliometric visibility in ways that current evaluation frameworks do not account for. The study contributes methodological and critical tools for a more reflexive, context-sensitive approach to research assessment in the SSH—one that treats bibliometric infrastructures not as neutral mirrors of scientific activity, but as active shapers of what counts as visible, impactful, and excellent.

References

Archambault, É. and Larivière, V. (2010). 'The limits of bibliometrics for the analysis of the social sciences and humanities literature', in UNESCO (ed.), *World Social Science Report 2009/2010*. UNESCO Publishing, Paris, pp. 251–254.

Gläser, J. and Oltersdorf, J. (2019). 'Persistent problems for a bibliometrics of social sciences and humanities and how to overcome them', in Catalano, G., Daraio, C., Gregori, M., Moed, H.F. and Ruocco, G. (eds.), *Proceedings of the 17th International Conference on Scientometrics and Informetrics*. Edizioni Efesto, Rome, pp. 1056–1067.

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S. and Rafols, I. (2015). 'Bibliometrics: The Leiden Manifesto for research metrics'. *Nature*, 520(7548), pp. 429–431. Available at: https://doi.org/10.1038/520429a

Kulczycki, E., Rozkosz, E.A., Dušek, M., Sivertsen, G. and Pölönen, J. (2018). 'Publication patterns in the social sciences and humanities: Evidence from eight European countries'. *Scientometrics*, 116(1), pp. 463–486. Available at: https://doi.org/10.1007/s11192-018-2711-0

Larivière, V. and Gingras, Y. (2010). 'On the relationship between interdisciplinarity and scientific impact'. *Journal of the American Society for Information Science and Technology*, 61(1), pp. 126–131. Available at: https://doi.org/10.1002/asi.21226

Leydesdorff, L. and Rafols, I. (2011). 'Indicators of the interdisciplinarity of journals: Diversity, centrality, and citations'. *Journal of Informetrics*, 5(1), pp. 87–100. Available at: https://doi.org/10.1016/j.joi.2010.09.002

Marres, N. and de Rijcke, S. (2020). 'From indicators to indicating interdisciplinarity: A participatory mapping methodology for research communities in-the-making'. *Quantitative Science Studies*, 1(3), pp. 1041–1055. Available at: https://doi.org/10.1162/qss_a_00062

Rafols, I. and Meyer, M. (2010). 'Diversity and network coherence as indicators of interdisciplinarity: Case studies in bionanoscience'. *Scientometrics*, 82(2), pp. 263–287. Available at: https://doi.org/10.1007/s11192-009-0041-y

Stirling, A. (2007). 'A general framework for analysing diversity in science, technology and society'. *Journal of the Royal Society Interface*, 4(15), pp. 707–719. Available at: https://doi.org/10.1098/rsif.2007.0213

Wang, Q. and Schneider, J.W. (2020). 'Consistency and validity of interdisciplinarity measures'. *Quantitative Science Studies*, 1(1), pp. 239–263. Available at: https://doi.org/10.1162/qss_a_00011

BIBLIOGRAPHIES AND THE BIBLIOMETRICS OF BOOKS. NEW TOOLS FOR A FAIR RESEARCH ASSESSMENT IN HUMANITIES

ANDREA BONACCORSI
University of Pisa, Italy

Keywords: Arts and Humanities, Books and book chapters, Book bibliometrics, Bibliographies, Fair assessment of research

Introduction and research questions

The issue of a fair assessment of research in humanities is still unsolved. The nature of the problem is well known: the process of peer review cannot benefit from the background of quantitative information, adopting the more advanced methodology of "informed peer review". In turn, the lack of quantitative information is the result of several factors: the lack of a clearly defined perimeter of publications, the poor coverage of humanities in almost all bibliometric databases, the limited availability of data on books and book chapters. Under these conditions, it is impossible to compare the subjective evaluations of peer review with any background information about, say, the counting of publications or citations or any other appropriate indicators.

Faced with this puzzle, the bibliometric community has generously explored several directions: expanding the citation indexes by including multilingual journals and book, creating a book citation index, integrating national repositories and Research Information Systems at national level, using Library Catalog Analysis (catalog holdings, loans) and book reviews (Goodreads, Amazon), introducing non-bibliometrics indicators (Altmetrics). All these efforts, for one reason or another, fail to satisfy the formal requirements of a fair and transparent system, such as commprehensiveness of items (journals, books, book chapters), clear definition of the perimeter (representativeness), transparency of criteria for inclusion of publications and their regular updating and revision, multilingualism, and availability of citation data (see the Reference lista for a selection of works exploring all these options).

In a recent paper (Bonaccorsi, 2026b) I have suggested that a promising direction for the development of a fair assessment of research in humanities should explore the potential of bibliographies.

Bibliographies are systematic collections of metadata within a clearly defined field of research in humanities. Bibliographies are curated and updated by scientific communities, often with a formal editorial responsibility by academic teams or individual editors. They are published by a variety of institutions.

Table 1. A sample of biblographies in Philology, Literature, Linguistics

Name	Year of foun dation	Main characteristics	Number of entries	Editors
Année Philologique	1928	Bigliography on all aspects of Ancient Greek and Roman civilizations	>1.481.251 (*)	Pedro Pablo Fuentes Gonzalez Societé Internationale de Philologie Classique
Library of Latin Texts online	1991	Largest database on Latin texts		Centre Traditio Litterarum Occidentalium (CTLO)
Literary and critical theory	2017		50-75 new articles per year	Eugene O'Brien
Linguistics	2011		50-75 new articles per year	Mark Aronoff
Classics	2010		50-75 new articles per year	Dee L.Chapman
Gnomon Bibliographische Datenbank	1994	Classical studies, >25.000 keywords	>864.000	Jurgen Malitz, Eichstätt Univ. Gregor Weber, Augsburg Univ.
Sources Chretiennes Online (SCO)		Online indexed version of Sources Chretiennes	>650	
Kritisches Lexicon zur deutschsprachigen Gegenwartsliteratur (KLG)		Biographies and bibliographies of authors in contemporary literature in German language		
Bibliographie der deutschen Sprach und Literaturwissenschaft (BDSL)	1957	1957-1969 edited by Hans W. Eppelsheimer 1970- today	>12.000 entries per year	Universität bibliothek Frankfurt am Main
Bibliographie del französischen Literaturwissenschaft (KLAPP)	1956	1956-1990 edited by Otto Klapp 1990-today online		
Bibliografia de la literatura española (BLE)	1980	Complete bibliography on Spanihsh literature		
European Literary Bibliography		Integration of Czech, Polish, Spanish and Finnish literary bibliographies	>4.777.867	
Bibliografia generale della Lingua e della	1994	>20 volumes of bibliographic records	>500.000 (1991-2014)	Centro Pio Rajina

Name	Year of foun dation	Main characteristics	Number of entries	Editors
Letteratura Italiana (BIGLI)				
Riviste di Italianistica nel mondo	2000	Bibliographies of publications on Italian literary authors from >150 journals	>70.000	Gianfranco Crupi
International Bibliography of Books and Articles on the Modern Languages and Literatures	1920	Literature records since 1880 1957- Annual Bibliography	>3.000.000 >70.000 new records annually	

Source: Bonaccorsi (2026b)

Here is a detailed description (Isis, 2026)

> Bibliography (from Ancient Greek: βιβλίον, romanized: biblion, lit. 'book' and -γραφία, -graphía, 'writing'), as a discipline, is traditionally the academic study of books as physical, cultural objects; in this sense, it is also known as bibliology (from Ancient Greek: -λογία, romanized: -logía). English author and bibliographer John Carter describes bibliography as a word having two senses: one, a list of books for further study or of works consulted by an author (or enumerative bibliography); the other one, applicable for collectors, is "the study of books as physical objects" and "the systematic description of books as objects" (or descriptive bibliography).

What is of interest to us is the typology called “enumerative bibliography” in this definition. I started a non-systematic exploration of existing bibliographies, collecting metadata on 140 sources (Bonaccorsi, 2026b). Table 1 is an example.

In my preliminary exploration I have found a number of interesting properties.

- As a first impression, it is surprising to find one or more bibliographies in almost all fields in humanities
- In some cases the coverage of bibliographies is large, covering entire disciplines (e.g. *International Philosophical Bibliography)*
- In other cases they cover sub-fields (e.g. *Oxford Bibliography of Jewish Studies*) or individual topics and domain niches (e.g. *Birgitta Bibliografin*)
- In most cases bibliographies are curated and edited by academic editors (identified by academic affiliations) or by senior librarians and library collection editors (identified by the library affiliation)
- The size of bibliographies is highly variable, from few hundreds to several million entries.
- They include all publication items (journal articles, books, book chapters) with full bibliographical metadata
- In the case of journals, they include non-indexed journals.
- All publication items are multilingualist.

- The total amount of entries declared in the website of bibliographies is in the order of 40 million, a very large number.

Based on these preliminary evidence, it seems that a system based on the integration of metadata from bibliographies might satisfy several formal requirements, for example comprehensiveness (journals, books, book chapters) and multilingualism.

In this paper I go forward into the exploration of the potential of bibliographies addressing the following research questions:

RQ 1 Do bibliographies cover the entire field of humanities?

RQ 2 Do bibliographies include citation data?

These two research questions address the formal requirements of representativeness (i.e. we need to define a clear perimeter) and availability of citations. Please note that these are the formal requirements not addressed by the previous efforts in bibliometrics: they failed to ensure a large and validated perimeted of analysis of scientific production in humanities, and to provide citation data.

To these two questions I add another one:

RQ 3 What are the governance models underlying the production of bibliographies?

This research question is crucial to the design of a large scale bibliometric system that might complement peer review in the fair assessment of research in humanities.

Data and methodology

Bibliographies are a relatively ignored source of data in research assessment and bibliometrics. This is surprising given the role they have in scientific communities in humanities, as witnessed by a large engagement at international level and the persistence and resilience of institutional initiatives, many of which achieve secular age, or at least age of several decades.

In order to address RQ 1 (Do bibliographies cover the entire field of humanities) I adopted the following methodology.

First, we need a granular description of the fields and sub-fields in humanities. It is well known that the taxonomy of disciplines is a complex issue, for which no universal classification is available. We adopted a multi-step approach.

First, we started with the English version of the Italian taxonomy, called “Settori scientifico-disciplinari” (Scientific disciplines), as available at https://www.cun.it/documentazione/academic-fields-and-disciplines-list/. This is labelled “CUN taxonomy”, from the acronym of the National University Council (CUN), the representative body of all scientific communities that is legally responsible for the creation and revision of the taxonomy.

This is a very granular classification, organized in a tree-like hierarchical taxonomy. The great advantage of this taxonomy is not only the number of sub-categories (n= 104 at the lowest level of the hierarchy) but also the correspondence between the nomenclature of categories and the recruitment and employment of specific academic categories. Therefore, if we use, say, Anatolian Studies in our taxonomy, this does not only denote a scientific community and (plausibly) a set of journals, books and book chapters, but also a concrete employment of researchers and professors in some Italian university. We find this criterion close to our notion of bibliographies as curated and edited by “scientific communities”, which we see in concrete occupational terms.

Second, we integrate the CUN taxonomy with two largely used journal classifications, i.e. the ASJC (All Science Journal Classification) and the Web of Science journal classification. These are more coarse-grained classifications (n= 43 for WoS). We

checked the overlapping between the CUN and the journal classifications and we identified a small number of missing categories in the CUN one. These have been added to the main list. Interestingly, they most refer to the US tradition of cultural studies (Cultural Studies, Ethnic Studies, Family Studies, Folklore, Women's Studies), as well as some specifications of history of science (History of social sciences) and Literature (Poetry). Given the absence of Theology in Italian universities, as opposed to the German-speaking countries, we also add this category to our list.

Please take note that we exclude from our list the entire field of Psychology, which in the Italian classification is still included in humanities. Following the international experience (but also the Italian research assessment practice), we keep Psychology outside humanities.

With this combined classification we extended in a systematic way the search for bibliographies, carried out with structured queries on the web. We were able to identify n=202 bibliographies, extending the preliminary list (n=140) attached to Bonaccorsi (2026b).

Coverage of fields in humanities

The result of the analysis is available in Appendix 1.

We were not able to identify bibliographies in 8 sub-fields in total. The largest group of missing bibliographies is found in the fields of national language translation studies (e.g. Portuguese and Brazilian, French, English, Spainish, German). It is not clear whether they represent scientific communities separated from the respective language and literature. It is likely that they have been created in the Italian academic system to offer an academic positions to lecturers in the respective languages. To this large group we add a few cases, such as Economic and political geography, or history of philosophy, which at a better examination might reveal that bibliographies do exist but are nested into other terminologies.

Summing up, a conservative estimate is that existing bibliographies cover more than 95% of fields and sub-fields in humanities. This is a remarkable number, It tells us that if we were able to integrate metadata from these sources we were immediately ready to claim comprehensiveness and coverage for the entire field of humanities.

Therefore the answer to our first research question (RQ 1 Do bibliographies cover the entire field of humanities?) is an unconditional yes.

Bibliographies and citation data

The second research question will receive a strong negative answer. None of the existing bibliographies publishes or announces citation data. The only available data is metadata (author, title, year, keyword, publication type, with item-specific data such as journal, issue, pages for journal articles, and place, publisher for books and book chapters).

Apparently, citation data is not crucial to the current needs of scientific communities in humanities. However, I was able to collect a small number of exercises in which bibliographies or collections have been used to generate a perimeter of publications, from which to extract citations. Let me summarize the most interesting examples.

a) Using bibliographies to construct samples of publications

A bibliography may work as a population among which it is possible to extract representative samples and obtain reliable lists of publications for citation analysis. This was done by Yitzhaki (1986) using the Index Bibliographicus Biblicus to extract randomly 33 articles and monographs published in 1971.

After extracting references from these references, he examined the metadata in order to study the proportion between Ancient Near Eastern Studies internal (i.e. within the discipline) and external citations, controlling the results with control samples taken from alternative bibliographies.

b) Using bibliographies to identify long term patterns in scientific work

One of the largest bibliographies is *L'Annéè Philologique*, covering all publications on the Latin and Greek literature and published by Brepols. The importance and use of this bibliography have been extensively examined in a special issue of *Revue de philologie, de littérature et d'histoire anciennes* (Viredaz, 2021; Vandenborre, 2021; Hilbold, 2021; Fuentes González, 2021; Montanari, 2021).

On the basis of these data Scheidel (1997) examined the 1924-1992 period in search of indicators of stability or change. Based on classification of sub-fields and the prominence of authors he concluded that the long term trend is one of remarkable stability.

c) Using bibliographies to map knowledge flows from other fields

Still working with data from *L'Annéè Philologique* Dufault (2023) tested the claim that philological works have been deeply affected by the adoption of so called "theory" in classics, or more generally in humanities (after post-structuralism this term refers to approach such as postcolonial studies, geneder studies, or psychoanalysis). Bibliographical data do not support this claim, as their use is marginal and not increased over time.

d) Using bibliographies to identify cycles of attention

Guérin (2025) focused on just one Latin author (Seneca the Father, not to be confused with the more famous Lucio Anneo Seneca, the son). After identifying the bibliographical records he examined the metadata and discovered an interesting regularity in the cyclic alternation between technical and philological work, at the beginning of the discovery or rediscovery of an author, and interpretive-critical work.

e) Using bibliographies to examine research collaboration

The Modern Language Association Bibliography is one of the largest and authoritative sources in literary studies. Leane et al. (2019) made a systematic sampling of references at 5 year interval in the 1995-2015 period, extracted 51.192 references and examined the cases of co-authorship. They document a low prevalence of collaborations (approximately 4%) but an upward trend.

f) Using bibliographies to build up maps of science

Several efforts have been undertaken with the goal of building maps of science, as for example using the classical *Archäologische Bibliographie* in archaeology to idetify community structures using internal classifications (Schich and Coscia, 2011), or comparing US journals that publish studies on French history with the *Bibliographie annuelle de l'histoire de la France*, then published by the French CNRS (Ghiati et al. 2005) (Schaeper (1991). These studies are complementary to other studies that obtain ad hoc bibliographies by specialized publishers or libraries, as in the case of Kokash et al. (2024) using data from Brill on the full text of 1804 books in Classics (see also Colavizza and Romanello, 2017; 2019; Colavizza et al. 2017; Kokash et al., 2023), or in the case of Blidstein and Zhitomirsky-Geffet (2022) building an ad hoc collection of 364 books on history, religion and culture of Mediterranean countries.

Summing up, it is my contention that bibliographies constitute the next promising starting point for a fair and responsible data infrastructure for humanities. Add to this consideration the circumstance that the software for citation extraction is becoming a commodity, with declining pricing.

Bibliographies and governance

We now have to turn to the issue of governance of bibliographies, namely the institutions that are responsible fo the curation, updating and publication of the 202 bibliographies.

We adopted a classification of institutions as follows

- Universities, research institutes, research and public libraries
- Non profit foundations, research networks. Learned societies, projects
- Individual authors
- Commercial publishers

Examples of the former categories are universities and university institutes (e.g. Princeton, Saragoza, Leuven, Marburg), research institutes (e.g. Institut National pour l'Histoire de l'Art, CNRS, DLA), and libraries (e.g. Tubingen, Yale, National Library of Norway or Sweden).

Examples of the second category are philanthropic foundations (e.g. Getty Research Institute), learned societies (e.g. American Association for the History of Science, American Modern Language Association) and bottom up networks of scholars and projects (e.g. Conservation Network). Research institutions are permanent and legally established, while research associations in this second group may be transitory and do not need a legal enforcement, although in practice some of them enjoy large longevity.

The third group includes single-author reference books and bibliographies: often the work of individual erudites and senior librarians, they appear in single volumes or limited series, from a large variety of publishers. For this category the issue of legitimation from a scientfic community is more problematic and should be evaluated case by case. A peculiar difficulty with these bibliographies is that they are often printed books for which a digital edition is not available. Nevertheless in some sub-fields (particularly in niche fields) they are the most authoritative source available.

Finally, the last category includes all bibliographies that are regularly published by commercial publishers. They usually have appointed editors, who may change over time in order to ensure continuity to the series. In some cases there is evidence of previously existing bibliographies that have been acquired by commercial publishers. In include in this category some of the largest digital infrastructures, such as EBSCO and ProQuest, which include in their catalogues several bibliographies.

Table 2 Breakdown of bibliographies by type of institution

Type of institution	**Number of cases**	**%**
Universities, research institutes, research and public libraries	58	28,7
Non profit foundations, research networks. Learned societies, projects	30	14,9
Individual authors	25	12,4
Commercial publishers	89	44,1
Total	202	100,0

Table 2 shows the breakdown by nature of institution. The largest group (44,1%) is formed by commercial publishers. The second includes research institutions (universities, institutes, libraries) with 28,7%, the third includes research associations with 14,9%. Individual authors follow at 12,4%.

We might comment this distribution by placing the data in the perspective of the creation of a fair research assessment system, in which complete metadata are available for a comprehensive perimeter of publications, with associated citation data.

The question therefore is: do the institutions involved in the production of bibliographies have an incentive to produce such a system?

A first dividing line might be between public and private institutions. Re-combining the data offers an interesting perspective: if we put aside individual authors (for which the main issue is to digitalize existing volumes) the breakdown between private and public sectors is 50:50, that is 44,1% commercial publishers, and 43,6% as the sum of research institutions and associations. This means that the two sectors have similar starting conditions. Who is going to take the initiative is an open question.

Another perspective might be dependent on the size. Given that the production of data is expensive, it might be that the incentive to undertake an upfront investment is positively associated to the size of the institution. In our case, what matters is the size of commercial publishers, given that public institutions and associations are, generally speaking, producers of a single or very few bibliographies.

As a first approximation it is interesting to observe the distribution of commercial publishers by the number of separate bibliographies they publish (Table 3).

Table 3. Distribution of commercial publishers by number of bibliographies

Publisher	**Number of bibliographies published**	**%**
Brill	14	15,7
Brepols	11	12,4
Oxford University Press	11	12,4
Bloomsbury	10	11,2
De Gruyter	7	7,9
ProQuest	6	6,7
EBSCO	5	5,6
John Benjamins	5	5,6
Peter Lang	3	3,4
Routledge	2	2,2
Klostermann	2	2,2
Archeobooks	2	2,2
Casalini Libri, Manchester UP, Princeton UP, Cambridge UP, Illinois UP, Salerno Editrice, Rowman & Littlefield, Bautz Verlag, Garland, CH Bek, Biering & Brinkman	1	
Total	89	100,0

I was able to identify 23 commercial publishers.[1] From a preliminary analysis on the number of bibliographies published the landscape does not seem concentrated in a small number of publishers. An increased level of concentration has been produced by the acquisition of the Dutch company Brill by De Gruyter in 2023. The first three competitors have a cumulative market share of 40,4% (48,3% after the acquisition by De Gruyter).

This paper is a call for reflection on the future of data collection and analysis in humanities, an issue still unsolved but in need of solution. A solution that should be fair and transparent.

References

Aksnes, D.W. and Sivertsen, G. (2019) 'A criteria-based assessment of the coverage of Scopus and Web of Science', *Journal of Data and Information Science*, 4(1), pp. 1–21.

Archambault, É., Vignola-Gagnè, É., Côté, G., Larivière, V. and Gingras, Y. (2006) 'Benchmarking scientific output in the social sciences and the humanities: the limits of existing databases', *Scientometrics*, 68(3), pp. 329–342.

Ardanuy, J. (2013) 'Sixty years of citation analysis studies in the humanities (1951–2010)', *Journal of the American Society for Information Science and Technology*, 64(8), pp. 1751–1755.

Ardanuy, J., Urbano, C. and Quintana, L. (2009) 'A citation analysis of Catalan literary studies (1974–2003): towards a bibliometrics of humanities studies in minority languages', *Scientometrics*, 81(2), pp. 347–366.

Benardou, A., Constantopoulos, P., Dallas, C. and Gavrilis, D. (2010) 'Understanding the information requirements of arts and humanities scholarship', *International Journal of Digital Curation*, 5(1), pp. 18–33.

Biagetti, M.T., Iacono, A. and Trombone, A. (2018) 'Is the diffusion of books in library holdings a reliable indicator in research assessment?', in Bonaccorsi, A. (ed.) *The evaluation of research in social sciences and humanities: lessons from the Italian experience*. Cham: Springer International Publishing.

Blidstein, M. and Zitomirsky-Geffet, M. (2022) 'Towards a new generic framework for citation network generation and analysis in the humanities', *Scientometrics*, 127, pp. 4275–4297.

Bonaccorsi, A. (2018) 'Towards an epistemic approach to evaluation in SSH', in Bonaccorsi, A. (ed.) *The evaluation of research in social sciences and humanities: lessons from the Italian experience*. Cham: Springer International Publishing.

Bonaccorsi, A. (2023) 'An epistemic approach to research assessment in the social sciences', in Engels, T.C.E. and Kulczycki, E. (eds.) *Handbook on research assessment in the social sciences*. Cheltenham: Edward Elgar.

Bonaccorsi, A. (2026a) *The knowledge of humanities: a comparative epistemology of historiography, literary criticism, history of art, and history of architecture*. Turnhout: Brepols.

[1] The analysis has been carried out manually, on information available on websites (last access 20 March 2026), with diligence. We remain responsible for any error. We encourage readers to submit comments and corrections, given the lack of standards in the field.

Bonaccorsi, A. (2026b) 'Introducing bibliographies: a new source for a responsible assessment of research in humanities' (under second round review).

Butler, L. and Visser, M.S. (2006) 'Extending citation analysis to non-source items', *Scientometrics*, 66, pp. 327–343.

Chang, Y.W. and Yeh, H.T. (2025) 'Do prolific arts and humanities authors have publishing preferences?', *Scientometrics*, 130, pp. 1281–1302.

Chi, P.S. (2014) 'Which role do non-source items play in the social sciences? A case study in political science in Germany', *Scientometrics*, 101(2), pp. 1195–1213.

Chi, P.S. (2016) 'Differing disciplinary citation concentration patterns of book and journal literature', *Journal of Informetrics*, 10(3), pp. 814–829.

Chi, P.S. (2020) 'The field-specific citation and usage patterns of book literature in the Book Citation Index', *Research Evaluation*, 29(2), pp. 203–214.

Clack, D.H. (1990) *Authority control: principles, applications, and instructions*. Chicago: American Library Association.

Colavizza, G., Romanello, M. and Kaplan, F. (2017) 'The references of references: a method to enrich humanities library catalogs with citation data', *International Journal on Digital Libraries*, 19, pp. 151–171.

Colavizza, G. and Romanello, M. (2017) 'Annotated references in the historiography on Venice: 19th–21st centuries', *Journal of Open Humanities Data*, 3, p. 2. Available at: https://doi.org/10.5334/johd.9

Colavizza, G. and Romanello, M. (2019) 'Citation mining of humanities journals: the progress to date and the challenges ahead', *Journal of European Periodical Studies*, 4, pp. 36–53.

Colavizza, G., Peroni, S. and Romanello, M. (2023) 'The case for the Humanities Citation Index (HuCI): a citation index by the humanities, for the humanities', *International Journal on Digital Libraries*, 24, pp. 191–204.

Colavizza, G. *et al.* (2018) 'Linked books: towards a collaborative citation index for the arts and humanities', in *Proceedings of the Red de Humanidades Digitales*. Mexico City, pp. 178–181.

Cronin, B., Snyder, H. and Atkins, H. (1997) 'Comparative citation rankings of authors in monographic and journal literature: a study of sociology', *Journal of Documentation*, 53(3), pp. 263–273.

Dagienė, E. (2023) 'Prestige of scholarly book publishers: an investigation into criteria, process, and practices across countries', *Research Evaluation*, 32(2), pp. 356–370.

Dagienė, E. (2024a) 'Mapping scholarly books: library metadata and research assessment', *Scientometrics*, 129(9), pp. 5689–5714.

Dagienė, E. (2024b) 'The challenge of assessing academic books: the UK and Lithuanian cases through the ISBN lens', *Quantitative Science Studies*, 5(1), pp. 98–127.

Dagienė, E. (2025) 'Exploring Google Books, Open Library, and Wikipedia as sources for book metadata: the UK and Lithuanian cases', in *Proceedings of the ISSI Conference*. Erevan.

Engels, T.C.E. and Guns, R. (2018) 'The Flemish performance-based research funding system: a unique variant of the Norwegian model', *Journal of Data and Information Science*, 3(4), pp. 45–60.

Engels, T.C.E., Ossenblok, T.L.B. and Spruyt, E.H.J. (2012) 'Changing publication patterns in the social sciences and humanities, 2000–2009', *Scientometrics*, 93(2), pp. 373–390.

Federation of Finnish Learned Societies *et al.* (2019) *Helsinki initiative on multilingualism in scholarly communication*. Figshare.

Giménez-Toledo, E. (2020) 'Why books are important in the scholarly communication system in social sciences and humanities', *Scholarly Assessment Reports*, 2(1), p. 6. Available at: https://doi.org/10.29024/sar.14

Giménez-Toledo, E. and Román-Román, A. (2009) 'Assessment of humanities and social sciences monographs through their publishers: a review and a study towards a model of evaluation', *Research Evaluation*, 18(3), pp. 201–213.

Giménez-Toledo, E., Sivertsen, G. and Mañana-Rodríguez, J. (2019) 'International register of academic book publishers (IRAP): overview, current state and future challenges', in *Proceedings of the ISSI Conference*, pp. 1752–1757.

Gorraiz, J., Gumpenberger, C. and Purnell, P.J. (2014) 'The power of book reviews: a simple and transparent enhancement approach for book citation indexes', *Scientometrics*, 98(2), pp. 841–852.

Gorraiz, J., Purnell, P. and Glänzel, W. (2013) 'Opportunities and limitations of the Book Citation Index', *Journal of the American Society for Information Science and Technology*, 64(7), pp. 1388–1398.

Guérin, C. (2025) 'L'Année Philologique, una source pour l'histoire de la philologie? L'exemple de Sénèque le Père', in Bougard, F., Mantovani, D. and Sanchi, L.A. (eds.) *Voies et voix de la philologie classique*. Aubervilliers: Institut de Recherche et d'Histoire des Textes.

Hicks, D. and Wang, J. (2010) 'Coverage and overlap of the new social sciences and humanities journal lists', *Journal of the American Society for Information Science and Technology*, 62(2), pp. 284–294.

Kokash, N. *et al.* (2023) 'From books to knowledge graphs', *Journal of Data Mining & Digital Humanities*, jdmdh: 9380.

Kokash, N. *et al.* (2024) 'The Brill knowledge graph: a database of bibliographic references and index terms extracted from books in humanities and social sciences', *Research Data Journal for the Humanities and Social Sciences*, 9, pp. 1–21.

Koskas, M. (2022) 'Universal bibliographic control today: preliminary remarks', in Bergamin, G. and Guerrini, M. (eds.) *Bibliographic control in the digital ecosystem*. Florence: Firenze University Press.

Kousha, K. and Thelwall, M. (2009) 'Google book citation for assessing invisible impact?', *Journal of the American Society for Information Science and Technology*, 60, pp. 1537–1549.

Kousha, K. and Thelwall, M. (2011) 'Assessing the citation impact of books: the role of Google Books, Google Scholar, and Scopus', *Journal of the American Society for Information Science and Technology*, 62, pp. 2147–2164.

Kulczycki, E. *et al.* (2018) 'Publication patterns in the social sciences and humanities: evidence from eight European countries', *Scientometrics*, 116(1), pp. 463–486.

Kulczycki, E. *et al.* (2020) 'Multilingual publishing in the social sciences and humanities: a seven-country European study', *Journal of the Association for Information Science and Technology*, 71(11), pp. 1371–1385. Available at: https://doi.org/10.1002/asi.24336

Li, K., Dobreski, B. and Busch, M. (2025) 'How are library cataloging metadata used differently over time: a large-scale quantitative analysis of the Library of Congress catalog', *Journal of Documentation*, 81(2), pp. 385–402.

Linmans, A.J.M. (2009) 'Why with bibliometrics the humanities does not need to be the weakest link: indicators for research evaluation based on citations, library holdings, and productivity measures', *Scientometrics*, 83(2), pp. 337–354.

Nederhof, A.J., van Leeuwen, T.N. and van Raan, A.F.J. (2010) 'Highly cited non-journal publications in political science, economics and psychology: a first exploration', *Scientometrics*, 83(2), pp. 363–374.

Nederhof, A.J. (2006) 'Bibliometric monitoring of research performance in the social sciences and the humanities: a review', *Scientometrics*, 66(1), pp. 81–100.

Nederhof, A.J. *et al.* (1989) 'Assessing the usefulness of bibliometric indicators for the humanities and the social and behavioral sciences: a comparative study', *Scientometrics*, 15(5–6), pp. 423–435.

Peroni, S., Shotton, D. and Vitali, F. (2016) 'A document-inspired way for tracking changes of RDF data', in *Proceedings of the Workshop on Detection, Representation and Management of Concept Drift in Linked Open Data*. CEUR Workshop Proceedings.

Sīle, L. *et al.* (2021) 'Towards complexity-sensitive book metrics for scholarly monographs in national databases for research output', *Journal of Documentation*, 77(5), pp. 1173–1195. Available at: https://doi.org/10.1108/JD-06-2020-0107

Sīle, L. *et al.* (2018) 'Comprehensiveness of national bibliographic databases for social sciences and humanities: findings from a European survey', *Research Evaluation*, 27(4), pp. 310–322. Available at: https://doi.org/10.1093/reseval/rvy016

Sivertsen, G. (2018) 'The Norwegian model in Norway', *Journal of Data and Information Science*, 3(4), pp. 3–19. Available at: https://doi.org/10.2478/jdis-2018-0017

Torres-Salinas, D., Gorraiz, J. and Robinson-Garcia, N. (2018) 'The insoluble problem of books: what does Altmetric.com have to offer?', *Aslib Journal of Information Management*, 70(6), pp. 691–707.

Torres-Salinas, D. and Moed, H.F. (2009) 'Library catalog analysis as a tool in studies of social sciences and humanities: an exploratory study of published book titles in economics', *Journal of Informetrics*, 3, pp. 9–26.

Vandenborre, C. (2021) 'Bibliographies aujourd'hui: vestiges du passé ou instrument de recherche pour l'avenir?', *Revue de Philologie, de Littérature et d'Histoire Anciennes*, 95, pp. 191–205.

Vandewalle, E. and Arhiliuc, C. (2025) 'How well does OpenAlex cover the Flemish social sciences and humanities?', in *Proceedings of ISSI 2025*, pp. 128–141. Available at: https://doi.org/10.51408/ISSI2025_074

Zuccala, A. *et al.* (2021) 'Performance-based publisher ratings and the visibility/impact of books: small fish in a big pond, or big fish in a small pond?', *Quantitative Science Studies*, 2(2), pp. 588–615.

Zuccala, A. and Robinson-Garcia, N. (2019) 'Reviewing, indicating and counting books for modern research evaluation systems', in Glänzel, W., Moed, H. and Schmoch, U. (eds.) *Handbook of science and technology indicators*. Cham: Springer. Available at: https://doi.org/10.1007/978-3-030-02511-3_27

Zuccala, A. and White, H.D. (2015) 'Correlating libcitations and citations in the humanities with WorldCat and Scopus data', in *Proceedings of the ISSI Conference*, pp. 305–316.

APPENDIX 1 Coverage of bibliographies in sub-fields of humanities

Scientific disciplinary sector	**Number of bibliographies**
10/A1 - ARCHAEOLOGY	
PREHISTORY AND EARLY HISTORY	XXX
AEGEAN CIVILIZATIONS	XX
NUMISMATICS	X
ETRUSCOLOGY AND ITALIC ANTIQUITIES	X
CLASSICAL ARCHAEOLOGY	XX
CHRISTIAN AND MEDIEVAL ARCHAEOLOGY	XX
ANCIENT TOPOGRAPHY	X
PAPIROLOGY	X
METHODS OF ARCHAEOLOGICAL RESEARCH	X
10/B1 - ART HISTORY	
HISTORY OF MEDIEVAL ART	X
HISTORY OF MODERN ART	XX
HISTORY OF CONTEMPORARY ART	XXX
MUSEOLOGY, ART AND RESTORATION CRITICISM	X
10/C - CINEMA, MUSIC, PERFORMING ARTS, TELEVISION AND MEDIA STUDIES	
PERFORMING ARTS	X
CINEMA, PHOTOGRAPHY AND TELEVISION	XX
MUSICOLOGY AND HISTORY OF MUSIC	XXX
ETHNOMUSICOLOGY	
10/D- SCIENCES OF ANTIQUITY	
GREEK HISTORY	X
ROMAN HISTORY	XX
AEGEAN CIVILIZATIONS	X
GREEK LANGUAGE AND LITERATURE	XX
BYZANTINE CIVILIZATION	XXX

Scientific disciplinary sector	**Number of bibliographies**
NEO-HELLENIC LANGUAGE AND LITERATURE	
LATIN LANGUAGE AND LITERATURE	XX
CLASSICAL PHILOLOGY	XXX
ANCIENT CHRISTIAN LITERATURE	XX
10/E- MEDIEVAL LATIN AND ROMANCE PHILOLOGIES AND LITERATURES	
MEDIEVAL AND HUMANISTIC LATIN LITERATURE	X
ROMANCE PHILOLOGY AND LINGUISTICS	XXX
ROMANIAN LANGUAGE AND LITERATURE	X
PORTUGUESE AND BRAZILIAN LITERATURE	X
LANGUAGE AND TRANSLATION - PORTUGUESE AND BRAZILIAN	
10/F- ITALIAN STUDIES AND COMPARATIVE LITERATURE	
ITALIAN LITERATURE	XX
CONTEMPORARY ITALIAN LITERATURE	X
ITALIAN LINGUISTICS	X
PHILOLOGY OF ITALIAN LITERATURE	X
LITERARY CRITICISM AND COMPARATIVE LITERATURE	X
10/G- HISTORICAL AND GENERAL LINGUISTICS	
ITALIC, ILLIRIC AND CELTIC PHILOLOGY	XX
HISTORICAL AND GENERAL LINGUISTICS	XXX
EDUCATIONAL LINGUISTICS	XX
ALBANIAN LANGUAGE AND LITERATURE	X
FINNO-UGRIC PHILOLOGY	X
10/H- FRENCH STUDIES	
FRENCH LITERATURE	XX
LANGUAGE AND TRANSLATION – FRENCH	
10/I - HISPANIC STUDIES	
SPANISH LITERATURE	X
LATIN AMERICAN LANGUAGES AND LITERATURES	X
LANGUAGE AND TRANSLATION - SPANISH	
10/L- ENGLISH AND ANGLOAMERICAN STUDIES	
ENGLISH LITERATURE	XXX
ANGLO-AMERICAN LANGUAGES AND LITERATURES	X
LANGUAGE AND TRANSLATION - ENGLISH	
10/M- GERMANIC AND SLAVIC LANGUAGES, LITERATURES AND CULTURES	
GERMANIC PHILOLOGY	XX
GERMAN LITERATURE	XX
LANGUAGE AND TRANSLATION - GERMAN	

Scientific disciplinary sector	Number of bibliographies
NORDIC LANGUAGES AND LITERATURES	X
DUTCH LANGUAGE AND LITERATURE	X
SLAVIC STUDIES	XX
10/N- EASTERN AND AFRICAN CULTURES	
HISTORY OF ANCIENT NEAR EAST	X
EGYPTOLOGY AND COPTIC CIVILIZATION	X
ASSYRIOLOGY	XX
ANATOLIAN STUDIES	X
ANCIENT NEAR EAST ARCHAEOLOGY AND ART HISTORY	X
PHOENICIAN AND PUNIC ARCHAEOLOGY	X
SEMITIC STUDIES, ETHIOPIAN LANGUAGES AND LITERATURES	X
HEBREW	XX
AFRICAN LANGUAGES AND LITERATURES	XX
HISTORY OF ISLAMIC COUNTRIES	X
MUSLIM ARCHAEOLOGY AND ART HISTORY	X
ARABIC LANGUAGE AND LITERATURE	XX
ARMENIAN, CAUCASIAN, MONGOLIAN AND TURKISH STUDIES	XX
IRANIAN PHILOLOGY, RELIGIONS AND HISTORY	X
PERSIAN LANGUAGE AND LITERATURE	XX
INDIAN AND CENTRAL ASIAN ARCHAEOLOGY AND ART HISTORY	X
INDIAN AND CENTRAL ASIAN PHILOSOPHY, RELIGIONS AND HISTORY	X
INDOLOGY AND TIBETOLOGY	XX
MODERN LANGUAGES AND LITERATURES OF THE INDIAN SUBCONTINENT	X
EAST ASIAN ARCHAEOLOGY, ART HISTORY AND PHILOSOPHY	X
CHINESE AND SOUTH ASIAN LANGUAGES AND LITERATURES	XX
JAPANESE AND KOREAN LANGUAGES AND LITERATURES	XX
HISTORY OF EAST AND SOUTH-EAST ASIA	X
11/A – HISTORY	
MEDIEVAL HISTORY	XXX
MODERN HISTORY	XXX
CONTEMPORARY HISTORY	XXX
HISTORY OF EASTERN EUROPE	XX
SCIENCE OF BOOKS AND DOCUMENTS, HISTORY OF RELIGIONS	X
ARCHIVAL SCIENCE, BIBLIOGRAPHY AND LIBRARIANSHIP	X
PALEOGRAPHY	XXX
HISTORY OF RELIGIONS	X

Scientific disciplinary sector	Number of bibliographies
HISTORY OF CHRISTIANITY AND OF CHURCHES	X
DEMOLOGY, ETHNOLOGY AND ANTHROPOLOGY	XXX
11/B GEOGRAPHY	
GEOGRAPHY	X
ECONOMIC AND POLITICAL GEOGRAPHY	
11/C - PHILOSOPHY	
THEORETICAL PHILOSOPHY	XXX
LOGIC AND PHILOSOPHY OF SCIENCE	XX
HISTORY OF SCIENCE AND TECHNOLOGY	XX
MORAL PHILOSOPHY	XX
AESTHETICS	XX
PHILOSOPHY AND THEORY OF LANGUAGE -	X
HISTORY OF PHILOSOPHY	
HISTORY OF ANCIENT PHILOSOPHY	X
HISTORY OF MEDIEVAL PHILOSOPHY	X
HISTORY OF SOCIAL SCIENCES (*)	X
11/D - PEDAGOGY AND EDUCATIONAL THEORIES	
PEDAGOGY, THEORIES OF EDUCATION AND SOCIAL EDUCATION	X
HISTORY OF PEDAGOGY AND EDUCATION	X
METHODOLOGIES OF TEACHING AND SPECIAL EDUCATION	X
EDUCATIONAL RESEARCH	X
METHODS AND TEACHING OF MOTOR ACTIVITIES	
METHODS AND TEACHING OF SPORTS ACTIVITIES	
ADDENDUM	
CULTURAL STUDIES (*)	XXX
ETHNIC STUDIES (*)	X
FAMILY STUDIES (*)	XX
WOMEN'S STUDIES (*)	XXX
POETRY (*)	XXX
THEOLOGY (*)	XXX

Legenda

Coding (10A...11D) derived from CUN taxonomy. Area 10 ANTIQUITIES, PHILOLOGY, LITERARY STUDIES, ART HISTORY; Area 11 HISTORY, PHILOSOPHY, PEDAGOGY AND PSYCHOLOGY.

(*) Sub-fields added to the CUN taxonomy on the basis of queries in ASJC and WoS.

Number of bibliographies: X one, XX two, XXX three or more.

POINTS OF FRICTION BETWEEN COARA PRINCIPLES AND CURRENT EVALUATIVE FRAMEWORKS AT THE NATIONAL, INSTITUTIONAL AND INDIVIDUAL LEVEL: PRELIMINARY RESULTS FROM THE RIF4HSS PROJECT

IVA MELINŠČAK ZLODI
University of Zagreb Faculty of Humanities and Social Sciences, Croatia

Keywords: Social Sciences and Humanities, Evaluative frameworks, Research assessment reform, CoARA principles, Open Science

This presentation reflects on the system of rules, requirements, and incentives that currently, and in the near future, shape the evaluative framework of the Croatian research system, and examines how these frameworks relate to the principles of the Agreement on Reforming Research Assessment (CoARA). The analysis is grounded in an ongoing institutional pilot conducted within the CoARA Boost project *RIF4HSS – Responsible Indicator Framework for Humanities and Social Sciences*, implemented at the University of Zagreb Faculty of Humanities and Social Sciences (FHSS).

The project is situated in a research assessment environment that remains predominantly structured by national-level criteria with a strong emphasis on quantitative indicators, many of which are not well adapted to the epistemic practices, publication cultures, and societal roles of research in the humanities and social sciences (HSS). At the same time, the Croatian system allows for a limited degree of institutional autonomy, particularly in the context of institutional self-assessment and re-accreditation processes. RIF4HSS is situated within this institutional space and examines how principles of responsible research assessment and open science are reflected in institutional practices within a multi-layered policy environment.

A key element of the project is the development of a participatory catalogue, or matrix, of research assessment indicators. The project starts from indicators that are already in use across different evaluative regimes and levels. The first version of the matrix maps indicators applied to individuals, research groups, project teams, and institutions, drawing on four main sources: national quality standards used in the re-accreditation of higher education institutions; performance indicators applied in negotiations on multiannual institutional funding contracts with the Ministry of Science, Education and Youth; criteria used in the evaluation of project proposals by the Croatian Science Foundation; and the draft National University, Scientific and Artistic Criteria currently under adoption. Together, these frameworks shape how research quality, productivity, and impact are assessed in Croatia.

Each indicator in the catalogue is described using a structured template specifying its intended purpose, level of application, data sources, collection methods, and limitations, as well as the organisational units and individuals responsible for its use and maintenance. The catalogue also includes selected indicators from assessment frameworks developed in other European contexts where CoARA principles have been

more explicitly reflected in assessment design, such as NOR-CAM and FIN-CAM. These examples are included to support comparison and contextualisation.

The catalogue has been developed as a pilot at a single HSS-focused institution, with the aim of testing its applicability in an institutional setting. While the pilot is institution-specific, it is informed by challenges commonly observed in non-Anglophone research environments, particularly a strong reliance on publication-based metrics. It also takes into account broader issues related to the recognition and use of national research infrastructures, such as CRIS systems and open repositories. The catalogue is conceived as a revisable instrument that can be adapted to similar institutional contexts.

At the time of submitting this abstract, the project is ongoing. The first version of the indicator matrix has been completed, and one participatory workshop with internal stakeholders has been conducted. Further empirical work, including interviews and focus groups with researchers, evaluation practitioners, management, and professional support staff, is planned prior to the conference. The presentation will therefore focus on preliminary observations arising from the mapping exercise, rather than on completed analytical findings.

At this stage, the mapping points to differences between evaluative frameworks and to uneven alignment with principles of responsible research assessment and open science. Some frameworks that are more closely aligned with these principles operate primarily at the institutional or project level and have a limited direct effect on individual researchers. In contrast, frameworks that play a stronger role in career progression, employment, and funding decisions rely more heavily on formalised quantitative indicators and publication-based criteria.

The presentation will discuss how these differences appear across national, institutional, and individual levels of evaluation, and how they relate to the conditions under which research in the humanities and social sciences is assessed. While grounded in the Croatian context, the issues addressed are relevant to other HSS-oriented and non-Anglophone research environments engaged in discussions on the reform of research assessment.

References

Agencija za znanost i visoko obrazovanje (2023). *Standardi kvalitete za vrednovanje u postupku reakreditacije visokih učilišta* [Standards for quality assessment in the re-accreditation of higher education institutions]. Available at: https://www.azvo.hr/wp-azvo-files/uploads/2024/05/Standardi-kvalitete-za-vrednovanje-u-postupku-reakreditacije-visokih-ucilista.pdf

Nacionalno vijeće za visoko obrazovanje, znanost i tehnološki razvoj (2025). *Savjetovanje s javnošću o Nacionalnim sveučilišnim, znanstvenim i umjetničkim kriterijima* [Public consultation on the National University, Scientific and Artistic Criteria]. Available at: https://esavjetovanja.gov.hr/ECon/MainScreen?entityId=30553

Republika Hrvatska, Ministarstvo znanosti, obrazovanja i mladih (2023). *Katalog ciljeva i pokazatelja* [Catalogue of objectives and performance indicators]. Available at: https://mzom.gov.hr/UserDocsImages/dokumenti/Znanost/Programski%20ugovori/Katalog-ciljeva-i-pokazatelja.pdf

Söderman, M., Himanen, L., Hyrkkänen, A.-K., Kallio, M., Koivisto, E., Mustajoki, H., Niemi, L., Pietilä, M., Pölönen, J. and Sipola, T. (2026). *FIN-CAM Tool for Researcher Assessment*. Zenodo. Available at: https://doi.org/10.5281/zenodo.18345588

NEXT GENERATION SSH-EVALUATION: FROM INDICATORS TO ONTOLOGIES

JON HOLM
Research Council of Norway, Norway

Keywords: Research evaluation, Social Sciences and Humanities, Indicators, Ontologies, Artificial Intelligence

Introduction

The Research Council of Norway is mandated by the government to "cater" for the assessment of research in Norway" (Ministry of research and education 2014). To fulfil this role, the Council organizes discipline-specific evaluations approximately every ten years, involving all publicly funded research organizations across the country (RoRI 2025). These evaluations are based on informed peer review carried out by international expert panels, who assess the Norwegian research system and provide recommendations to enhance quality and efficiency.

New technologies, including AI, provide new opportunities for collecting and analysing data for research evaluation (Holm et al. 2024). This contribution discusses how evaluation data could be structured in upcoming national evaluations of Social sciences and Humanities (SSH) in Norway to best serve the formative purpose of the evaluation exercise. To operationalise this question, we will compare a traditional presentation of evaluation data as a set of indicators, with the option of establishing a formal ontology for SSH research activities that may allow more complex exploratory analysis of networks, new research trends (topic modelling) and correlations.

Policy context

The ENRESSH policy brief for Better Adapted Procedures for Research Evaluation in the SSH states that an adequate evaluation of research should consider the diversity of research practices across disciplines, type of R&D (basic, applied and experimental development) and institutions (Ochsner 2020). Recent calls for reform of research assessment voice concerns that current evaluation systems evaluate research quality in ways «that is incentivizing overly narrow outputs and goals among individuals, groups, departments, universities and other knowledge producing entities» (Rushford 2025a). In addition to concern of goal displacement affected by narrow indicators, there is also a growing realization that the widespread use of bibliometric indicators incentivise opportunistic publishing practices (Rushford 2025b). In this presentation we will show how formal ontologies can help include new data sources and allow for a broader set of analytical perspectives that considers *the diversity of research practices across disciplines, types of research and institutions*.

How to construct an ontology for SSH evaluation in Norway?

An ontology is a way to formally describe the concepts in a domain and the relationships between them, providing a shared vocabulary and structure for data (Cammarata 2024). Unlike a basic taxonomy for research disciplines or topics, an

ontology can represent richly qualified relations (beyond simple hierarchies) and attributes of entities, capturing the complexity of real-world knowledge.

The purpose of the subject-specific evaluations in Norway is not only to assess the quality of the output from SSH in terms of research results, but also to contextual factors determining research quality and relevance and how outputs are valued by non-academic audiences. In past evaluations, data on the various aspects and layers of research activities have been collected and assessed separately. Typically, access to funding, research productivity and impact have been conceptualised as separate evaluation criteria. Thus, the connections between funding, research production and societal impact remain poorly understood resulting in an understanding of research quality as an absolute feature disconnected from goals stated by funders and the needs of various beneficiaries outside of the research group.

The Norwegian Current Research Information System (formerly CRISTIN, now NVA, i.e. the Norwegian database of scientific publications) will be used as a starting point for building an ontology that links from publications to their antecedents (e.g., data, funding, and resources used in their production) and succedents (e.g., researchers, institutions, or societal actors who cite or benefit from the results). NVA also includes a national repository of full text research publications. We will further investigate the possibility of integrating open data on research from other sources, building on foundational work that has already been carried in the context of CoARA by the working group "Towards Open Infrastructures for Responsible Research Assessment (OI4RRA)". (Manola N et al. 2025)

Once the ontology is established, it can be adapted to serve our evaluative purposes. We envision two main applications:

1. Investigating specific predefined knowledge creation processes conceptualised by a knowledge graph modelling a structured network of real-world entities (institutions, researchers, funders, publications, etc.) and their semantic relationships. It could model linear relations, for example from a research grant used for collection of data that is analysed in a publication and cited by an academic peer or a policy-paper, or many-to-many relationship between specific research organisations, funding sources, research data, types of publications and publication channels and their readership.
2. The second application is to extract patterns from data (text or other sources) in an inductive way by statistical methods, machine learning or other AI techniques. Such methods may be used to identifying new classes or research networks that transcend established disciplinary or organisational boundaries. As an example, there might be cognitive overlaps between research papers traditionally classified as belonging to different disciplines based on the journals where they are published. Such overlaps may indicate new interdisciplinary fields or shared areas of application. New classes or networks may be discovered by using algorithmic methods to detect patterns in inherent properties such as vocabulary (e.g. topic modelling) or references lists (e.g. network analysis).

Strengths and weaknesses of ontologies vs indicators

Indicators provide a high level of transparency and predictability as they are defined through simple calculations, most often based on a small set of variables. This makes it relatively easy for all parties to understand why a certain evaluation unit receives a certain score on a specific indicator. Ontology-based pattern discovery — particularly

inductive applications such as topic modelling or network analysis — is harder to explain. When an algorithm identifies a network or a cluster, its calculations are distributed across model parameters that few evaluation stakeholders will be able to audit. Evaluators should thus ensure the accountability of algorithmic analysis by documenting datasets, methods and code, as well as assuring expert validation of results. For both types of analysis, the completeness of available data sources is a main concern. However, it is much harder to assess how these limitations affect the results of the analysis for ontology-based analysis because the analysis is combining different data sources with different degrees of coverage.

Evaluations often measure development over time. Such stability is readily achieved with indicators calculating with the same formula for the same variables, whereas algorithmic methods (clustering, network analysis) may produce different cluster structures depending on the state of the data at the time of analysis, making longitudinal comparison harder.

Conclusion

The planned approach of using ontologies as a basis for algorithmic network and classification analyses may provide new insights into the properties of Norwegian research in the SSH that go beyond the punctual measurements provided by traditional indicators. We see this approach as well adapted to reflect the diversity of research practices across disciplines as advocated by ENRESSH (Ochsner at al 2020).

Nonetheless, it is important to consider the limitations of ontology-based algorithmic analyses compared to traditional indicators as discussed above. We acknowledge that the two approaches are not mutually exclusive. The choice between individual indicators and ontology-based analyses should be informed by the purpose of the evaluation, the units of evaluation and its assessment criteria

References

Cammarata, E. and Maineri, A. (2024). 'Ontologies for the social sciences: What they are, and why we need them'. *ODISSEI FAIRblog*, 22 February. Available at: https://doi.org/10.5281/zenodo.10694246

Curry, S., de Rijcke, S., Hatch, A., Pillay, D., van der Weijden, I. and Wilsdon, J. (2020). *The Changing Role of Funders in Responsible Research Assessment: Progress, Obstacles and the Way Ahead*. Research on Research Institute. Available at: https://doi.org/10.6084/m9.figshare.13227914.v1

Holm, J., Newman-Griffis, D. and Petersson, G.J. (2024). 'Big data for big investments: Making responsible and effective use of data science and AI in research councils', in *Artificial Intelligence and Evaluation: Emerging Technologies and Their Implications for Evaluation*. pp. 120. Available at: https://doi.org/10.4324/9781003512493-7

Manola, N. et al. (2025). *Conceptual Architecture for the Implementation of a Responsible Research Assessment Framework Built on Open Infrastructures (v4)*. OpenAIRE, 29 April. Available at: https://doi.org/10.5281/zenodo.15487415

Ministry of Research and Education [Kunnskapsdepartementet] (2014). *Forskrift om vedtekter for Norges forskningsråd §2j*. Available at: https://lovdata.no/dokument/INS/forskrift/2016-06-10-658 (Accessed: 18 January 2026).

Ochsner, M., Kancewicz-Hoffman, N., Ma, L., Holm, J., Gedutis, A. and Šima, K. et al. (2020). *ENRESSH Policy Brief Research Evaluation*. Figshare. Available at: https://doi.org/10.6084/m9.figshare.12049314.v1 (Accessed: 31 March 2020).

RoRI (2025). *RoRI Atlas of Assessment (Norway)*. Available at: https://researchonresearch.org/observatory/norway/ (Accessed: 17 January 2026).

Rushforth, A. (2025a). 'Beyond impact factors? Lessons from the Dutch attempt to transform academic research assessment'. *Research Evaluation*, 34, rvaf035. Available at: https://doi.org/10.1093/reseval/rvaf035

Rushforth, A. (2025b). 'Research assessment reform as collective action problem: Contested framings of research system transformation'. *Minerva*. Available at: https://doi.org/10.1007/s11024-025-09573-3

POSTER SESSION

LATIN AMERICA & CARIBBEAN RETRACTION WATCH: A PROPOSAL WEB DIRECTORY

JOEL ALHUAY-QUISPE
Universidad Nacional Mayor de San Marcos, Perú
LOURDES BAUTISTA-YNOFUENTE
Universidad San Ignacio de Loyola, Perú
VICTORIA YANCE-YUPARI
Universidad Peruana de Ciencias Aplicadas, Perú

Keywords: Misconduct publishing, Scientific publishing, Retraction database, Scientific database, Questionable journals

The recent proliferation of publishing misconduct makes it necessary to study in-depth research ethics and integrity publications in the digital era. Preliminary studies from Latin America and the Caribbean (LAC) provide evidence of ethical misconduct and a lack of integrity in publications. The Retraction Watch Database, a global platform of retracted papers, and RetractBase, a Spanish database for retracted scientific publications sourced by Crossref and OpenAlex, inspired the proposal of a web application to serve as a directory of retractions from authors with affiliation to Latin America & Caribbean (LAC) countries. This proposal complements recent studies in LAC about research ethics and academic misconduct such as paper mills, authorship for sale, publishing in fraudulent and questionable journals, hijacked and cloned journals.

Introduction

Recent studies on Latin America & Caribbean (LAC) countries report an increasing evidence about misconduct ethics and publication integrity due digital context by misuse of artificial intelligence, continuous pressure to "publish papers", hegemony of selective/proprietary databases, and lack of policies on monitoring/assessment of research. An emerging black market of papers including farms of fake papers/paper mills and scientific authorship for sale is coming-up in South American countries with Brazil and Peru as cases documented (Santos-d'Amorim, et al, 2024; Alhuay-Quispe, & Yance-Yupari, 2025).

Rhodes & Linnenluecke (2025) defines "junkification research", as driven by the commodification of research following: academia intensifies pressures to produce and publish, intensification of 'publish or perish'. In recent years, new issues of anomalies in predatory publishing are defined as questionable journals which are delisted or excluded from selective scientific databases including cases in Latin American titles (Alhuay-Quispe, Bautista-Ynofuente, & Madero-Durán, 2024). Related cases of publishing in fraudulent sources reveals that more than twenty journals from Latin America were listed as hijacked journals (Alhuay-Quispe, et al., 2025a, 2025b).

Monitoring retractions in Social Sciences and Humanities for research evaluation, is a misconduct in these fields often involves plagiarism and the 'junkification' of theoretical discourse (Martínez-Rojas et al., 2025). Consequently, the proposed

observatory serves as a critical infrastructure for assessing how 'publish or perish' pressures distort the SSH bibliographic record in Latin America (Rhodes & Linnenluecke, 2025).

There are a limited background and research on retracted papers by Latin American authors. Santos-d'Amorim et al. (2021) investigated a sample of Brazilian retracted articles indexed in SciELO Brazil; furthermore, to 2024, 641 documents from the LAC region have been retracted (Damasio, 2024). Another study, about the reasons for retractions across different subject areas in Latin America concludes that misconduct is predominant cause for retraction, mainly in Business and Technology, while the unintentional errors are common in Biological and Health Sciences (Martínez-Rojas et al., 2025).

Retraction Watch started in 2010 as a journalism blog focused on creating a repository for retractions break stories, eventually leading to the creation of a comprehensive retraction database (Oransky, & Marcus, 2010). In September 2023, after more than a decade running the database at www.retractiondatabase.org, the organisation behind the Retraction Watch blog and database announced that the Retraction Watch Database (RWDB) was acquired by Crossref, making the entire raw dataset publicly available (Hendricks, & Lammey, 2023). More recently, in 2025 was launched RetractBase, a Spanish open database of scholarly publications retracted (retracted articles), retraction notices (retractions) or have been removed (withdrawals) which is sourced by Crossref and OpenAlex (Sánchez, et al., 2026).

Proposal and findings

The study presents the case of retracted papers in Latin America, and introduces a web-app proposal of directory or database of retractions from authors with affiliation to Latin America & Caribbean (LAC) countries.

The proposal of directory is contained in a prototype database of retractions coming from LAC countries is available at https://joelalhuayq.shinyapps.io/LACretractWatch/.

Figure 1 shows a preliminary view of the index web database that includes a single search box and advanced filters.

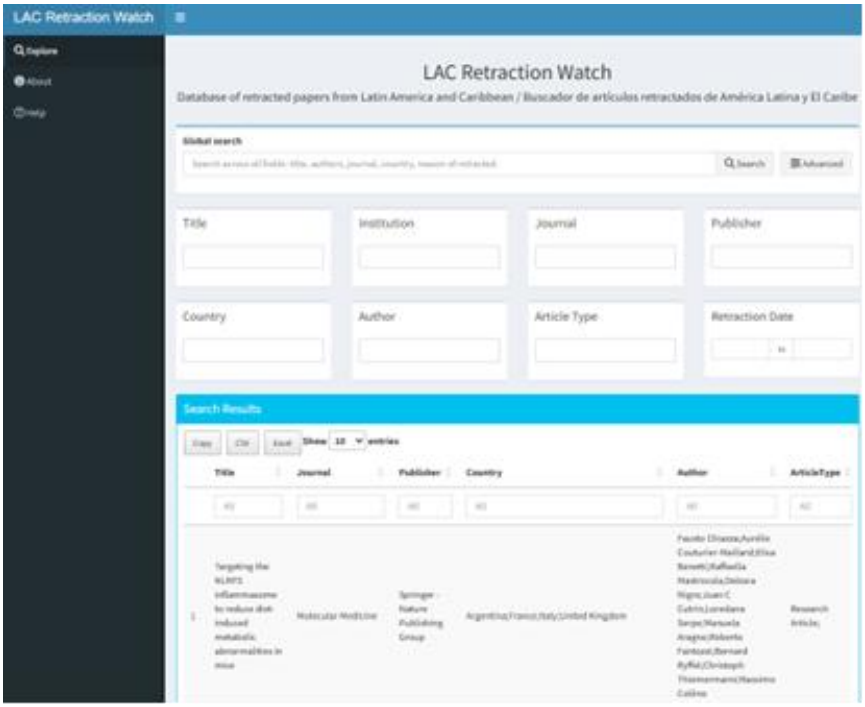

Figure 1. Latin America & Caribbean Retraction Watch Directory. *Source*: Own authors.

The application web was developed in R-language with shinydashboard library as framework web running at server shinyapps.io. On March 15 2026, a total of 967 retracted documents were retrieved in real time from GitHub CrossRef retracted database, which is filtered in the field "institution" that only documents with at least one country of Latin America and Caribbean.

By conclusion way

A prototype database serving as a directory of retracted papers complements recent studies in Latin America & Caribbean research ethics and publishing misconduct, such as paper mills, authorship for sale, and the rise of fraudulent and questionable journals, hijacked and cloned journals.

By identifying patterns in retracted literature, this tool transitions from a technical directory to a diagnostic instrument for institutional integrity (Alhuay-Quispe & Yance-Yupari, 2025). This enables evaluators to distinguish between genuine scholarship and the output of emerging paper mills or fraudulent journals across the region.

References

Alhuay-Quispe, J., Yance-Yupari, V., Estrada-Cuzcano, A. and Bautista-Ynofuente, L. (2025a). 'Revistas fraudulentas, clonadas y secuestradas en América Latina'. *Palabra Clave (La Plata)*, 14(2), e253. Available at: https://doi.org/10.24215/18539912e253

Alhuay-Quispe, J., Yance-Yupari, V., Estrada-Cuzcano, A. and Bautista-Ynofuente, L. (2025b). 'Fraudulent, cloned and hijacked journals in Latin America: Dark side of traditional metrics', in *RESSH Conference 2025, Reforming Research Assessment — Book of Abstracts*, 19–21 May 2025, Helsinki, Finland, p. 41.

Alhuay-Quispe, J. and Yance-Yupari, V. (2025). 'The black market of publications in Peru: Paper mills and authorship for sale'. *Learned Publishing*, 38(3). Available at: https://doi.org/10.1002/leap.2014

Alhuay-Quispe, J., Bautista-Ynofuente, L. and Madero-Durán, S. (2024). 'Revistas médicas de América Latina descontinuadas de Scopus'. *Revista Cubana de Información en Ciencias de la Salud*, 35. Available at: http://scielo.sld.cu/scielo.php?pid=S2307-21132024000100003&script=sci_arttext

Damasio, E. (2024). 'A global study in the Retraction Watch Database RWDB', in *46th Society for Scholarly Publishing (SSP) Annual Meeting*, 29–31 May 2024, Boston, MA, USA. Available at: https://doi.org/10.14293/S2199-SSP-AM24-01009

Hendricks, G. and Lammey, R. (2023). 'Crossref acquires Retraction Watch data and opens it for the scientific community'. *Crossref Blog*. Available at: https://doi.org/10.13003/c23rw1d9

Martínez-Rojas, E., Zahn-Muñoz, C. and Espinaza-Solar, R. (2025). 'Retraction of scientific literature: Analyzing the reasons for retractions across different areas of knowledge in Latin America'. *Data and Metadata*, 4, 638. Available at: https://doi.org/10.56294/dm2025638

Oransky, I. and Marcus, A. (2010). 'Why write a blog about retractions?'. *Retraction Watch Blog*. Available at: https://retractionwatch.com/2010/08/03/why-write-a-blog-about-retractions/

Rhodes, C. and Linnenluecke, M.K. (2025). 'The junkification of research'. *Organization*, pp. 1–16. Available at: https://doi.org/10.1177/13505084251399576

Sánchez, C., Ortega, J.L., Delgado-Quirós, L., Becerra-Rodero, E., García, A. and Moreno, C. (2026). 'Incidence and evolution of retracted publications according to the sources used: A meta-analysis comparison with RetractBASE'. *Journal of Information Science*, pp. 1–12. Available at: https://doi.org/10.1177/01655515251398723

Santos-d'Amorim, K., Correia, A.E.G.C., Miranda, M.K.F.O. and Santa-Cruz, P. (2021). 'Reasons and implications of retracted articles in Brazil'. *Transinformação*, 33, e210001. Available at: https://doi.org/10.1080/10508422.2022.2141747

Santos-d'Amorim, K., Wang, T., Lund, B. and Macedo Dos Santos, R.N. (2024). 'From plagiarism to scientific paper mills: A profile of retracted articles within the SciELO Brazil collection'. *Ethics & Behavior*, 34(1), pp. 40–57. Available at: https://doi.org/10.1080/10508422.2022.2141747

INCLUSIVE CO-EVALUATION IN SOCIAL INNOVATION AND RESEARCH: CASE STUDY AND METHODOLOGICAL INNOVATION OF A TRAINING PROGRAM IN ENTREPRENEURIAL SKILLS WITH REFUGEE WOMEN

TAMARA BUENO DORAL, RUBÉN FUENTES-FERNÁNDEZ, NOELIA GARCÍA-CASTILLO
Complutense University of Madrid, Spain

Keywords: Co-evaluation, Diversity, Participatory methodologies, Forced migration, Social innovation

Current evaluation systems face substantial challenges in measuring societal impact and adequately assessing social innovation projects, as they continue to rely heavily on quantitative metrics, with criteria like societal relevance and social responsibility receiving less attention (Pontika et al, 2022). Therefore, we advocate that the social impact of projects be considered a substantive and decisive evaluation criterion, especially in cases where the final beneficiaries are involved in the research processes, as their real needs are taken into account. Given this gap, this case study seeks to answer: how a multi-stakeholder co-evaluation framework can be applied to effectively measure the societal impact of research and innovation projects for vulnerable populations?

Funded by the European Horizon 2020 program, RAISD[1] project aimed to identify highly vulnerable groups among the forcibly displaced people, analyze their specific needs, and find suitable practices to address them. It was based on a triangulation methodology that considered Responsible Research and Innovation (RRI), participatory action research, and socio-ecological models (Hänninen et al, 2022).

Among other aspects, this methodology requires that representatives of all actors directly participate in the co-design, co-implementation, and co-evaluation of the research. To implement this methodology, an Action-research Unit (ARU) was created in each of the participating countries (Italy, Finland, Turkey, Jordan, Hungary, Lebanon, and Spain) to co-design a Tailored Attention and Inclusion Strategy (TAIS) for a highly vulnerable group among the forcibly displaced people. For example, the Spanish ARU located at the Complutense University of Madrid (coordinator of the project) engaged more than 30 participants to co-create an intervention with the aim of promoting the inclusion of female refugees from sub-Saharan origin by enhancing their entrepreneurial skills.

This poster presents a case study on the evaluation of the Spanish TAIS, which reflects this methodological innovation regarding the co-evaluation of research on social intervention actions from the perspectives of different social actors, including the final beneficiaries themselves. The co-evaluation was carried out using a combination of predominantly qualitative techniques. Apart from gathering information from the organizing team and conducting participatory observation, the Spanish team

[1] See https://raisd-h2020.eu/

conducted the outcome evaluation through in-depth interviews with key service providers and beneficiaries of the training program. This research not only presents the results but also the processes that gave rise to this co-evaluation from the perspective of diversity, as well as the limitations and lessons learned so that this approach can be transferred to other research and innovation initiatives.

References

Hänninen, L., García-Castillo, N., Fuentes-Fernández, R., Bueno Doral, T., Pérez, C. and Verdejo, V. (2022). *Tailored Attention and Inclusion Practices — Methodology*. Available at: https://raisd-h2020.eu/resources/scientific-outputs/

Pontika, N., Klebel, T., Correia, A., Metzler, H., Knoth, P. and Ross-Hellauer, T. (2022). 'Indicators of research quality, quantity, openness, and responsibility in institutional review, promotion, and tenure policies across seven countries'. *Quantitative Science Studies*, 3(4), pp. 888–911. Available at: https://doi.org/10.1162/qss_a_00224

THE COMMITMENT OF THE UNIVERSITY OF FLORENCE LIBRARY SYSTEM TO OPEN SCIENCE AND EXCELLENCE IN RESEARCH

BENEDETTA CALONACI, SILVIA MOLTENI
University of Florence, Italy

Keywords: Open access, Academic libraries, Institutional repositories, Research evaluation

For over twenty years, the University of Florence has supported the Open Science philosophy, disseminating its principles within its academic community and promoting their adoption in daily practice. In fact, Florence was among the universities that signed the Messina Declaration (2004), and was the first in Italy to issue, in 2012, a specific policy on open access to research outputs. In the following years, the importance of Open Science at the University grew, culminating in its inclusion in the new Statute (2019) as a core principle for research. In this process, the University Library System has taken on a role of coordination, training, and support. New and cross-disciplinary skills have developed within the libraries, which have proven essential for supporting genuine openness in research and open, innovative teaching. The University of Florence thus serves as a concrete example of how academic libraries have been able to harness the potential of Open Science and have assumed a leading role within the University. In particular, since 2010, an ad hoc working group called "Support for Open Access and Research Evaluation" has been established. Focusing on the FloRE Repository, this group assists researchers in the process of publishing their work, encouraging the adoption of open licenses and proper metadata creation. Support for open access began with the green road (self-archiving) and took shape through the implementation of the institutional repository (E-prints, since 2001, and then, starting in 2014, FloRe), managed in collaboration between the Libraries, Research area, and the University IT System. The repository also serves as the basis for research evaluation. Periodic evaluation exercises (VQR) and national scientific qualifications (ASN) have provided the OA-SBA Working Group with valuable opportunities to improve the quality of the catalog and, at the same time, to disseminate knowledge of the principles of Open Science among researchers, especially those at the beginning of their academic careers. In this regard, the Libraries have also been called upon to provide support to the University of Florence Research Observatory for monitoring the University's scientific output and its integrity, as well the proper management of data through DMP. Since 2020, the University Library System has assumed the managerial and financial responsibility for "transformative agreements" for open-access publication of research outputs. Finally, joining the European Euniwell Alliance has provided a boost to open science, with the goal of harmonizing different habits towards shared best practices.

TOWARDS A STRATEGIC VIEW OF RESEARCH ASSESSMENT: THE COARA NATIONAL CHAPTER OF ITALY 2026-27

MONICA DI LUCA, PAOLA GALIMBERTI
University of Milan, Italy
GIULIANA RUBBIA, ALESSANDRO CAROSI
National Institute of Geophysics and Volcanology, Italy

Keywords: Open science, Research evaluation, CoARA Italian National Chapter

Over the past three years, a network of universities, research centres, learned societies, and associations has developed through active engagement in the National Chapter of the Coalition for Advancing Research Assessment (CoARA NC ITA). Efforts to plan the implementation of the shared principles and values of the Agreement on Reforming Research Assessment (ARRA) have benefited from diverse mutual learning experiences. However, CoARA NC ITA also faced a series of difficulties and obstacles, for which collaboration and joint action with the ministries that fund research and with the evaluation agency are seen as possible solutions.

Over the next two years, CoARA NC ITA aims to build on the results achieved during the 2023-2025 period, drawing on the existing network of relationships, activities, and documents, and to strengthen relationships with Ministries, Working Groups, and National Chapters of other countries, to share best practices adopted by NC members.

The planned activities for CoARA NC ITA are:

- Consolidate the active network among Italian institutions, promoting the monitoring of the Institutional Action Plans;
- Maintaining stable and bidirectional connections with the thematic COARA Working Groups, via periodic updates, to explore the impact of their results on national practices and provide a feedback based on the experiences of the Italian communities on more specific topics: examples are TIER-inclusive evaluation, multilingualism, peer review, biomedical disciplines, ethics, etc…)
- Strengthen interaction and collaboration with the other National Chapters, by organizing joint thematic workshops and webinars on cross-cutting topics and developing shared tools, such as templates for reporting and monitoring progress on CoARA commitments, with at least two NCs;
- Engage with the Ministries that primarily conduct research evaluations at various levels, MUR and the Ministry of Health, for the implementation of the Agreement's principles in our country. Promote institutional exchanges with the two ministries through regular communications and targeted meetings.
- Establish connections with other European initiatives related to research evaluation (e.g. Barcelona Declaration on Open Research Information), since transparency and openness of research processes and data are essential for fair, cross-disciplinary, understandable, and reproducible evaluation.

The work is split into three work packages: WP1, for coordination, engagement and reinforcement of the NC, resulting in Policy Briefs; WP2, for coordination with the CoARA WGs and the other NCs, activities include organizing events and providing reports; WP3 dedicated to open and qualitative research information for a more transparent evaluation, producing dissemination events on the principles of open research information, recommendations and guidelines.

PERSPECTI: A COMPREHENSIVE RESEARCH AND CITATION DATABASE FOR SCHOLARLY WORKS IN THE HUMANITIES AND THE ARTS

JOSHUA EYKENS
Brepols Publishers, Belgium

Keywords: Arts and Humanities, Citation index, Bibliographic coverage, Bibliodiversity, Multilingualism

The need for representative research and citation data for the arts and humanities (A&H) remains largely unmet. Existing indexing services primarily focus on English-language journal content published by large commercial publishers (Colavizza, Peroni, & Romanello, 2023). While these services have enabled important analyses of the global research landscape, they neglect a substantial share of the diverse book-based and multilingual scholarship that is central to research in the A&H. This limitation particularly affects disciplines in which monographs and edited volumes play a key scholarly role (see, for example, Vandewalle & Arhiliuc, 2025). More recently developed indexing services (e.g. OpenAlex, Research Commons, Dimensions) offer broader coverage, but their citation linking remains limited (ibid.). These structural biases undermine the discoverability, visibility, and recognition of A&H research.

Perspecti (www.perspecti.net) addresses this gap by developing a database that captures the publication and citation fingerprint of research in the A&H. It aims to provide reliable citation information for researchers, librarians, policymakers, and funders. Central to Perspecti's approach is a commitment to foster bibliodiversity: beyond major academic publishers, the project actively collaborates with medium-sized and smaller presses that play a crucial role in sustaining humanities scholarship.

This poster presents Perspecti's inclusive approach to publication types (e.g. journal articles, monographs, text editions, conference proceedings, collected essays, festschrifts, and book chapters), languages (including Romance, Germanic, Slavic, other Indo-European, and ancient languages), and disciplinary fields. We outline how citation data are collected and discuss the main bibliographic sources that feed into Perspecti.

In addition, the poster provides a visual overview of the Perspecti user interface, illustrating how citation and reference data are presented alongside author affiliations, open access status, and funder information. The overview of a development timeline informs RESSH participants about future plans. While Perspecti will initially focus on scholarly publications, coverage may later be expanded to other cited materials of importance to the A&H. By embracing bibliodiversity and multilingualism as core principles, Perspecti aims to comprehensively capture the A&H research landscape and enable more meaningful data analysis and information retrieval across these disciplines.

References

Colavizza, G., Peroni, S. and Romanello, M. (2023). 'The case for the Humanities Citation Index (HuCI): A citation index by the humanities, for the humanities'.

International Journal on Digital Libraries, 24, pp. 191–204. Available at: https://doi.org/10.1007/s00799-022-00327-0

Vandewalle, E. and Arhiliuc, C. (2025). 'How well does OpenAlex cover the Flemish social sciences and humanities?', in *ISSI 2025: 20th International Conference on Scientometrics and Informetrics*, 23–27 June 2025, Yerevan, Armenia. International Society for Scientometrics and Informetrics, pp. 128–141. Available at: https://doi.org/10.51408/ISSI2025_074

COARA EXPECT: VIDEOS TO PROMOTE RESPONSIBLE RESEARCHER ASSESSMENT

MIKI KALLIO, MARIKA RAUHALA
University of Oulu, Finland

Keywords: Responsible researcher assessment, CoARA, Cultural change in assessment, Instructional videos

CoARA ExPECT was a collaborative initiative involving Ruhr University Bochum, Tampere University, and the University of Oulu, concluding in December 2025. Its primary objective was to advance CoARA's vision for reforming research assessment practices. Supported by one year of CoARA Boost funding, the project produced 19 concise instructional videos designed to assist researchers, reviewers, university leadership, funders, and other stakeholders in implementing responsible researcher assessment at various career stages.

Throughout the project, academic staff and researchers from each institution actively contributed to topic selection and script development for the videos. Insights and feedback were gathered during numerous events and conferences, with RESSH2025 in Helsinki proving particularly significant.

Recognising that responsible researcher assessment cannot rely on a universal model, the project emphasised the need to adapt evaluation processes to reflect the distinct characteristics of each discipline. The ExPECT videos address a wide range of issues in researcher assessment, including affecting cultural change, promoting process transparency, careful evaluator selection, avoiding unconscious biases, career-stage-specific assessment, teaching-focused positions, narrative CVs, career assessment matrices, assessing research quality and impact, rewarding research integrity and open science, and the misuse of metrics. Special attention was also given to the unique context of SSH disciplines.

At the RESSH2025 conference, participant feedback on video scripts provided valuable perspectives, allowing for stronger representation of researchers' viewpoints in the final materials. With this presentation, we aim both to showcase the completed videos to the research community and to express our gratitude to all contributors involved in their development.

The instructional videos, each lasting between two and six minutes, are freely accessible under the CC BY-SA license for the benefit of the wider research community.

References

Kallio, M., Riedel, S., Ihonen, M., Sassali, J., Rauhala, M., Kanstrén, K., Lukkari, E., Hämeenkorpi, H., Kalliorinne, K. and Jantunen, J. (2025). *CoARA ExPECT — Video Catalog*. Zenodo. Available at: https://doi.org/10.5281/zenodo.17648971

AIBOX – DASHBOARDING ALGORITHMIC ANALYSIS TOOLS FOR APPLICATION PROCESSING IN A RESEARCH FUNDING ORGANISATION

VILMA LEHTINEN
Research Council of Finland, Finland

Keywords: Algorithms, Application processing, Research funding organisation

The Research Council of Finland (RCF) grants more than EUR 500 million in research funding each year and opens around 40 funding calls, which attract nearly 6,000 applications.

RCF has developed in-house several algorithmic analysis tools to support application processing. For example, computer-assisted text analysis is used to identify thematic similarities across applications and to anticipate the expertise required in the review process. Until now, these analyses have been developed and deployed separately, without integrating them into a common platform.

The AIbox project creates a solution that gives RCF staff easy access to existing tools and streamlines the processing of applications by pooling the tools together onto a dashboard. At the same time, the tools are being developed to support the principles of responsible research evaluation that RCF is committed to, such as transparency and fairness.

AIbox has received funding from The Finnish Innovation Fund (Sitra), from the Public Sector Productivity Programme. The project ran from June 2025 to February 2026, and the development continues within RCF.

With this poster, we share our insights from the project and provide a platform for discussion on how to address concerns about reliability, transparency and fairness regarding application processing supported by algorithms.

DIAMOND OPEN ACCESS IN ITALY: EVIDENCE AND IMPLICATIONS FOR RESEARCH EVALUATION

GINA PAVONE, SIMONE ANGIONI, MIRIAM BAGLIONI, ALESSIA BARDI, ANDREA MANNOCCI
National Research Council – Institute of Information Science and Technologies "Alessandro Faedo" (CNR-ISTI), Italy

Keywords: Diamond open access, Open research information, Open science, Open scientometrics, Research evaluation

Diamond Open Access (DOA) is increasingly seen by the scientific community as a promising model for a more equitable and sustainable scholarly communication system. Understanding how DOA is positioned in research evaluation is essential, as evaluation practices directly affect its recognition, sustainability, and transformative potential.

This abstract summarizes the work accepted at Quantitative Science Studies (Angioni et al., 2025). We investigate the presence and impact of DOA in Italy. Since the absence of APCs is not sufficient to identify DOA journals (Simard et al., 2024), and existing authoritative sources don't capture all relevant dimensions, we constructed a dedicated dataset. Italian DOA journals were identified according to the "Operational Diamond OA Criteria for Journals" (Armengou et al., 2024), developed within the DIAMAS and CRAFT-OA projects.

The dataset

To meet the operational Diamond criteria, it was necessary to combine and cross-integrate multiple data sources — EZB, ROAD, DOAJ, and OpenAIRE — to build a dataset covering all relevant Diamond dimensions.

Overview

We identified 168 DOA journals, 118 active in 2024 compared with 14 in 2000 - an overall twelvefold increase. The strongest presence is in the Social Sciences, Humanities, and Arts, while Engineering and Technology remain marginal. Italian DOA journals are smallscale, publishing 10-24 articles per year. Recent years show a shift toward 25-49. DOA accounts for 9.4% of journals published in Italy and indexed in OpenAIRE, in line with international estimates.

Impact, citations and evaluation

Although DOA publications are increasing, citations show a recent decline, possibly due to the time needed for citation accumulation. DOA journals demonstrate recognition within evaluation systems: 128 journals are classified as scientific in the ANVUR list, and 72 are top-rated in at least one disciplinary area. Internationally, 81 Italian DOA journals are indexed in SCImago, with 12 ranked in Q1, confirming that high-quality and highimpact publishing is compatible with the DOA model.

Rights and publishing systems

Technical features show mixed results: 57.7% of publishers are institutional (e.g., universities) and 26.2% nonprofit (e.g., academic societies) 81.8% support copyright retention, while 17% impose restrictions. Licences include CCBY (42.9%) and

CCBYNCND (20.3%). ORCID integration remains limited (26.7%), and preservation data are available for 35% of journals.

Final remarks

Italian DOA is expanding and recognised by evaluation systems, but challenges persist at the technical and policy levels. Strengthening DOA means rethinking evaluation practices so they reinforce autonomy, integrity, and bibliodiversity rather than reproducing systemic imbalances.

References

Angioni, S., Baglioni, M., Bardi, A., Mannocci, A., Manghi, P. and Pavone, G. (2025). *Charting the Landscape of Italian Diamond Open Access Publishing*. Zenodo. Available at: https://doi.org/10.5281/zenodo.17484911

Armengou, C., Bargheer, M., Gingold, A., Holsinger, S., Laakso, M., Mitchell, D., Mounier, P., Pölönen, J., Rooryck, J., Ševkušić, M., Souyioultzoglou, I. and Varachkina, H. (2024). *Operational Diamond OA Criteria for Journals*. Zenodo. Available at: https://doi.org/10.5281/zenodo.12721408

Simard, M.A., Butler, L.A., Alperin, J.P. and Haustein, S. (2024). 'We need to rethink the way we identify diamond open access journals in quantitative science studies'. *Quantitative Science Studies*, 5(4), pp. 1042–1046. Available at: https://doi.org/10.1162/qss_c_00331

CULTURAL HERITAGE AS RESEARCH OUTPUT: GLAM WIKIMEDIA PROJECTS AT THE CNR CENTRAL LIBRARY

EMANUELA CILIONE, GIORGIA MIGLIORELLI, MARIA ADELAIDE RANCHINO, SARA SANTORSA
National Research Council - Planning, Programming and Central Library Unit, Italy
ISABELLA FLORIO
National Research Council - Cosenza Research Area, Italy

Keywords: GLAM Wikimedia, Open access, Research libraries, Knowledge sovereignty, Digital cultural heritage

In a context in which scientific autonomy is increasingly conditioned by commercial publishing models, proprietary infrastructures, and evaluation systems grounded in quantitative metrics (SIBI CNR, 2022), GLAM Wikimedia projects can be understood as institutional strategies for research sovereignty and the ethical dissemination of knowledge (Bailey, 2018; Wikimedia Foundation, 2009).

This contribution presents the Central Library of the Consiglio Nazionale delle Ricerche as an emblematic case study of the role that a public library can assume within the digital research ecosystem (ARL, 2019), through three interconnected GLAM Wikimedia Italia projects.

Developed within the institutional framework of the CNR, the three experiences concern: the digitization and open-access publication of the institution's scientific and historical cartographic heritage preserved by the Central Library; the valorization and dissemination of materials related to the history, education, and culture of Deaf Community — including content in Italian Sign Language (LIS) — belonging to the heritage of CNR ISTC; the digitization and open dissemination of the musical collections of the Library of the Conservatorio di Santa Cecilia, including scores, librettos, and historical documentation.

Taken together, these projects show how a research library can move beyond a merely custodial role and operate as a hub for the production and circulation of open scientific knowledge, transforming heritage materials into reusable outputs for research (Kelly, 2019).

Through the systematic integration of primary sources, images, datasets, and structured metadata in Wikimedia Commons, Wikipedia, and Wikidata, the Central Library adopts a publication model that reduces dependence on commercial publishing platforms and evaluation systems (Candela et al., 2024). At the same time, these projects act on centre-periphery dynamics by repositioning forms of knowledge traditionally domain-specific — such as thematic cartography, sign language, and specialized musical resources — at the centre of global knowledge circulation.

In conclusion, the contribution uses GLAM Wikimedia projects as an interpretative lens to interrogate the role of public research libraries in contemporary processes of knowledge production, legitimation, and circulation (IFLA, 2011).

References

Association of Research Libraries (2019). *ARL White Paper on Wikidata*. ARL, Washington, DC. Available at: https://www.arl.org/resources/arl-whitepaper-on-wikidata/ (Accessed: 27 February 2026).

Bailey, J. (2018). 'The changing orientation in global GLAMWiki'. *JLIS.it*, 9(3), pp. 1–19.

Candela, G. et al. (2024). 'A systematic review of Wikidata in GLAM institutions: A labs approach', in *Linking Theory and Practice of Digital Libraries, TPDL 2024*. Springer, Cham, pp. 34–50.

IFLA (2011). *IFLA Statement on Open Access*. Available at: https://www.ifla.org/wp-content/uploads/2019/05/assets/hq/news/documents/ifla-statement-on-open-access.pdf (Accessed: 27 February 2026).

Kelly, E.J. (2019). 'Reuse of Wikimedia Commons cultural heritage images on the web'. *Evidence Based Library and Information Practice*, 14(3), pp. 52–76.

SiBiCNR (2022). *Valutazione della ricerca*. Available at: https://sibi.cnr.it/open-science-2/valutazione-della-ricerca/ (Accessed: 27 February 2026).

RETRACTIONS IN THE SOCIAL SCIENCES AND HUMANITIES: DESCRIPTIVE EVIDENCE FROM RETRACTION WATCH NOTICES

AYÇA NUR SEZEN
The Scientific and Technological Research Council of Türkiye (TÜBİTAK), National Academic Network and Information Center (ULAKBİM), Turkey
ESMA BAĞCI
Directorate General for Libraries and Publications, Turkey
CEREN BILGE SEFEROĞLU
Bartın University, Department of Information and Records Management, Turkey
GÜLEDA DOĞAN
Hacettepe University, Department of Information Management, Turkey

Keywords: Research integrity, Retraction notices, Social Sciences and Humanities, Editorial actions, Retraction Watch Database

Retractions and related editorial actions such as corrections and expressions of concern play an important role in protecting the scholarly record by signaling unreliable or problematic publications (COPE, 2025). Although research integrity concerns affect all disciplines, retractions remain considerably less visible in the Social Sciences and Humanities (SSH) than in STEM fields.

This study provides a descriptive overview of retraction notices with a focus on SSH. Data were drawn from the Retraction Watch Database downloaded on 31 October 2025. The original dataset contained 67,190 records. Because individual publications may generate multiple notice entries, records were consolidated so that each work appears once in the analytic dataset. The final dataset contains 64,721 unique works. Records were grouped into STEM, SSH, and multidisciplinary categories using subject tags provided in the database.

The analysis compares fields in terms of editorial action types, stated reasons, and publication to action timing, and profiles SSH cases by publication type, region, access status, and collaboration structure. The study does not estimate retraction risk because the database includes only notice bearing records. Instead, it describes patterns within recorded cases.

STEM dominates the registry with about 82 percent of records, while multidisciplinary publications account for 14 percent and SSH approximately 4 percent. Across all fields, full retractions represent more than 95 percent of editorial actions, with corrections and expressions of concern forming a small minority. Reason patterns differ across fields. SSH notices are more frequently associated with plagiarism and administrative or legal issues, whereas fabrication or manipulation related cases are more prominent in STEM. Timing distributions show broadly similar median delays across fields but greater dispersion in SSH, including longer upper delay tails. SSH records also tend to involve smaller collaboration structures, with a higher share of single author, single

institution, and single country publications, reflecting common publication patterns in these fields.

Overall, the results suggest that although SSH retractions are relatively rare, their documented patterns differ from those observed in STEM contexts.

References

Committee on Publication Ethics (COPE) (2025). *Retraction Guidelines*. 29 August. Available at: https://publicationethics.org/guidance/guideline/retraction-guidelines

FINNISH CAREER ASSESSMENT MATRIX (FIN-CAM): SUPPORTING RESPONSIBLE RESEARCHER ASSESSMENT

MIRA SÖDERMAN, JANNE PÖLÖNEN
Federation of Finnish Learned Societies, Finland
LAURA NIEMI
University of Turku, Finland
ANNE-MARI JÄRVELIN
Lapland University of Applied Sciences, Finland

Keywords: FIN-CAM, Career assessment matrix, Responsible research assessment, Balancing qualitative and quantitative assessments, Recognition of diverse research contributions and career paths

Career Assessment Matrices (CAMs) have emerged as part of international efforts to reform research assessment, aimed at promoting transparent, holistic and responsible assessment practices. CAMs provide a structured yet flexible approach to assessing academic contributions. They aim to capture the full breadth of academic contributions, competencies and professional trajectories.

FIN-CAM (Finnish Career Assessment Matrix) represents a national-level implementation of this approach (Söderman et al. 2026). It is designed to support the implementation of the National Recommendation for the Responsible Evaluation of a Researcher in Finland (2020) and to be used alongside it.

FIN-CAM is structured across six competence areas that reflect the diversity of academic work:

- research, development and innovation
- education and supervision
- societal interaction
- participation in the art and research community
- general researcher skills
- background and formal qualifications

Within each area, competences, skills and merits, as well as illustrative examples of how the competences can be described both qualitatively and quantitatively, are outlined. Aligned with the National Recommendation and CoARA commitments, assessment should focus primarily on qualitative evaluation, supported by quantitative indicators where relevant and necessary. FIN-CAM serves as a tool for both evaluators and researchers.

FIN-CAM has been developed through a participatory process that included a survey to researchers on the diversity of career assessment criteria (Pölönen et al. 2024) and a public consultation on the draft version of FIN-CAM. The development of FIN-CAM has received CoARA Boost Cascade Funding, which enabled the organisation of three workshops to pilot the tool's usability in different organisational contexts and

assessment situations. These consultations have informed both the structure and content of the matrix, as well as the forthcoming online tool.

The aim of this poster is to present FIN-CAM as a practical tool for promoting responsible researcher assessment practices. The poster introduces key components of FIN-CAM and potential use cases. In addition, the poster presents ongoing work towards a web-based version of FIN-CAM and its anticipated functionalities, which is intended to enhance accessibility and uptake and to support its reflective use.

References

Pölönen, J., Himanen, L., Hyrkkänen, A.-K., Kallio, M., Koivisto, E., Lahdenperä, H., Muhonen, R., Niemi, L., Pietilä, M., Sipola, T. and Söderman, M. (2024) *Researchers' views on diversity of career assessment criteria in Finland: a survey report*. Zenodo. Available at: https://doi.org/10.5281/zenodo.11612535

Söderman, M., Himanen, L., Hyrkkänen, A.-K., Kallio, M., Koivisto, E., Mustajoki, H., Niemi, L., Pietilä, M., Pölönen, J. and Sipola, T. (2026) *FIN-CAM Tool for Researcher Assessment*. Zenodo. Available at: https://doi.org/10.5281/zenodo.18345588

Working group for responsible evaluation of a researcher (2020) *Good Practice in Researcher Evaluation in Finland: Recommendation for the responsible evaluation of a researcher in Finland*. Responsible Research Series 7:2020. Available at: https://doi.org/10.23847/isbn.9789525995282

COLLATERAL EVENT - BOOK LAUNCH

"THE KNOWLEDGE OF HUMANITIES" BY ANDREA BONACCORSI (2026): FIRST MONOGRAPH IN THE BREPOLS' SERIES "RESEARCH ON RESEARCH IN THE ARTS, HUMANITIES AND SOCIAL SCIENCES"

ANDREA BONACCORSI
University of Pisa, Italy
GUNNAR SIVERTSEN
NIFU, Norway
TIM C.E. ENGELS
University of Antwerp, Belgium
EMANUEL KULCZYCKI
Humboldt University and DZHW, Germany

Keywords: Epistemology, Historiography, Literary criticism, History of Art, History of Architecture

The new book series *Research on Research in the Arts, Humanities and Social Sciences* (RoRAHSS - https://www.brepols.net/series/rrahss) is published by Brepols Publishers to provide a dedicated long-form outlet for in depth studies of how research works within the arts, humanities, and social sciences (AHSS). Tim Engels and Emanuel Kulczycki (Engels & Kulczycki, 2025) initiated the series as a platform for comprehensive, book-length treatments for scholars who investigate the processes, communication, evaluation, and impact of scholarship in AHSS – as well as those studying science and technology systems more broadly from AHSS perspectives, such as the sociology of science or critical science studies. The editorial board of the series consists of internationally recognized experts: Andrea Bonaccorsi (University of Pisa, Italy), Noela Invernizzi (Universidade Federal do Parana, Brazil), Lai Ma (University College Dublin, Ireland), Michael Ochsner (FORS, Switzerland), and Cassidy Sugimoto (Georgia Institute of Technology, USA). Together they bring a broad range of perspectives and deep experience in the field and will help us ensure the series reflects diversity and scholarly excellence.

With RoRAHSS we welcome book proposals from a wide range of disciplinary and interdisciplinary backgrounds. By doing so, we seek to establish the meta-research on AHSS disciplines more prominently as a recognized branch of science studies. The series explicitly positions itself with a global outlook, building on long-standing European and international collaborations in the field.

During this session, we will first introduce the RoRAHSS book series, the need for it, and the community momentum that lead to the initiation of the series. Then, Andera Bonaccorsi will introduce his monograph *The knowledge of humanities*, which will be available for purchase in print and as a free PDF via https://www.brepols.net/products/IS-9782503618746-1 from mid-February 2026 onwards. Andrea worked on the monograph for more then 10 years, bringing together a wide range of perspectives and reflections.

Summary of the monograph from the website:

"Are humanities scientific disciplines? This book offers compelling arguments for reopening the issue of the relations between humanities and natural sciences. It criticizes the current trend of defending humanities on the basis of their expected social or economic impact, in terms of cultural heritage, wellness, mass culture or democratic habits.

The book vigorously calls for a new epistemological reflection, which addresses the issues left open by the XX century: Do humanities produce valid knowledge or just culture and informed opinion? How do researchers in humanities advance explanations and causal statements? How do they build up the proofs for their statements, given that they deal with objects and events buried in the past? Do we have science for individual cases? The author suggests a novel approach, based on the comparative analysis of various disciplines in humanities. They deal with historiography in the first place and with the main historical objects of investigation: the text (literary criticism), the image (history of art), the building (history of architecture).

The book dives deep into some of the most authoritative works of researchers of the XX century in these fields, reading them from the perspective of epistemology. It then identifies a number of scientific principles and procedures that meet the highest requirements for reliability and validity. Recent epistemological developments in theories of causality, nature of scientific laws, and history-dependent natural sciences suggest that a new type of dialogue between humanities and hard sciences is possible. A dialogue between peers."

Following the presentation of *The Knowledge of Humanities* (Bonaccorsi, 2026), Gunnar Sivertsen will reflect on the book. Sivertsen researches and advises the development of research evaluation and funding systems in several countries in direct interaction with governments, institutions, and academic communities. He received the Derek de Solla Price Memorial Medal in 2025 for "outstanding contributions to the fields of quantitative studies of science". Trained in philology and specialising in the study of the Dano-Norwegian writer of the Enlightenment, Ludvig Holberg (1684-1754), Sivertsen defended his doctoral thesis "Kilden til Jeppe paa Bierget" in 2006.

References

Bonaccorsi, A. (2026) *The knowledge of humanities: a comparative epistemology of historiography, literary criticism, history of art, and history of architecture*. Turnhout: Brepols. Available at: https://www.brepols.net/products/IS-9782503618746-1 (Accessed: 16 April 2026).

Engels, T.C.E. and Kulczycki, E. (2025) 'Introducing the research on research in the arts, humanities and social sciences book series', *ISSI Newsletter*, pp. 16–19.

INDEX OF KEYWORDS

Keywords

Keywords are based on authors' submissions and have been harmonised for readability; page numbers indicate where each topic is addressed

Academia; 138
Academic autonomy; 81; 145
Academic career assessment; 122
Academic libraries; 87; 262
Academic promotion; 125
Academic publishing; 43
Academic recruitment; 141
Academic staff perceptions; 69
Academic work; 118
Academic-societal collaboration; 48
AI-assisted peer review; 25
Algorithms; 268
Altmetrics; 87
Application processing; 268
Artificial Intelligence; 18; 35; 251
Arts and Humanities; 232; 265
Arts and Humanities scholarship; 53
Assetisation; 21
Balancing qualitative and quantitative assessments; 275
Barcelona Declaration; 104; 114; 188
Becher's typology; 221
Bibliodiversity; 265
Bibliographic coverage; 265
Bibliographies; 232
Bibliometrics; 83
Book bibliometrics; 232
Books and book chapters; 232
Brazil; 158
Canadian journals; 192
Career assessment matrix; 275
Career progression; 58
Career stage and evaluation; 128
Citation index; 265
Citizen science; 177
CoARA; 12; 58; 132; 198; 267
CoARA action plan; 62
CoARA Italian National Chapter; 263
CoARA principles; 248
Co-creation; 81
Co-evaluation; 260
Collective capabilities; 48
Community engagement; 92; 145
Community spaces; 164
Comparative legal analysis; 141
Copyright law; 40
Cross-border research; 40
Cultural change in assessment; 195; 267
Data Act; 40
Data disambiguation; 228
Database coverage and comparability; 228
Definition(s) of quality; 50
Diamond open access; 269
Digital cultural heritage; 271
Disciplinary differences; 221
Disciplinary heterogeneity; 149
Discoverability; 87
Diversification of career paths; 122
Diversification of contributions; 122
Diversity; 260
Diversity of research outputs; 128
Diversity of scientific contributions; 62
Editorial actions; 273
Editorial office; 18
Effects; 118
English dominance; 109
Epistemology; 29; 278
Equality; 138
ERA Act; 40
European research funding; 12
European Union; 12
Evaluative frameworks; 248
Fair assessment of research; 232
Family resemblances; 50
Financialization; 43
FIN-CAM; 275
Forced migration; 260
FP10 research policy; 12
Funding priorities and evaluation criteria; 158
GenAI; 21
Gender; 138
Generativity; 21
Geopolitics; 35

GLAM Wikimedia; 271
Grey publishers; 78
Historiography; 278
History of Architecture; 278
History of Art; 278
Humanities; 128
Impact evaluation; 174
Impact of SSAH; 35
Inactive researcher; 132
Inclusive excellence; 138
Indicators; 251
Influence on research topics; 98
Informativity; 50
Institutional repositories; 262
Institutional support; 69
Instructional videos; 267
Interdisciplinarity; 182; 228
International collaboration; 64
Internationalization of science; 64
Intuitionism; 50
ISO standards; 162
Japan; 149
Journal reputation; 78
Judicial review; 141
Key value indicators; 73
Knowledge co-production; 83
Knowledge sovereignty; 271
Knowledge valorisation; 92
Language diversity; 128
Large Language Models; 208; 215
Latin America; 64
Literary criticism; 278
Literature review; 192
Machine learning; 221
Metadata curation; 188
Metric-based governance; 125
Metrics; 29; 43
Misconduct publishing; 256
Mode 1/Mode 2 knowledge production; 158
Montevideo Group; 64
Multilingual research assessment; 114
Multilingualism; 98; 109; 265
Narrative assessment; 81
New public management; 224
Non-scholarly publications; 170
Novel quantitative indicators; 53
Objectivity; 29
Ontologies; 251
Open; 248
Open access; 87; 262; 271
Open access publishing; 78
Open data; 203
Open peer review; 192
Open research assessment; 198
Open research information; 104; 269
Open research infrastructures; 195
Open science; 35; 40; 114; 188; 195; 198; 203; 263; 269
Open scientometrics; 269
OpenAlex; 104
Paper mills; 224
Participatory methodologies; 260
Peer identification; 215
Peer review; 18; 25; 29
Peripheral universities; 69
Policy implementation; 58
Principles; 182
Professional communication; 170
Public engagement; 177
Public libraries; 164
Public value; 73
Publication dynamics; 125
Publication formats; 98
Publishing practices; 78
Publishing research; 40
Qualitative research; 53
Quality management; 162
Questionable journals; 256
Recognition of diverse research contributions and career paths; 275
Reducing linguistic biases; 109
Research assessment reform; 114; 122; 188; 198; 248
Research evaluation; 25; 35; 43; 48; 53; 69; 73; 78; 83; 87; 92; 98; 118; 125; 128; 132; 138; 141; 149; 154; 158; 164; 195; 221; 224; 251; 262; 263; 269
Research exceptions; 40
Research funding; 174
Research funding organisation; 268
Research funding portfolios; 208
Research governance; 154
Research information infrastructures; 21
Research information platforms; 188
Research information systems; 114
Research integrity; 25; 273
Research libraries; 271
Research output; 40
Research performing organisation; 58

Research policies; 34; 35
Research quality; 50
Research security; 203
Research societal impact; 174
Researchers' freedom; 203
Responsible innovation; 73
Responsible metric; 132
Responsible research and innovation; 162
Responsible research assessment; 58; 64; 81; 118; 128; 177; 228; 275
Responsible researcher assessment; 267
Responsible scientific career evaluation; 62
Retraction database; 256
Retraction notices; 273
Retraction Watch Database; 273
Reviewer matching; 221
Reviewer training; 81
Scholarly benchmarking; 215
Scholarly publishing; 18; 192
Science and law; 203
Science, technology and innovation policy; 158
Sciences; 177
Scientific autonomy; 21; 43; 78; 195
Scientific communication; 109
Scientific database; 256
Scientific publishing; 18; 256
Scientific societies; 34
Social inclusion; 164
Social innovation; 260
Social Sciences and Humanities; 92; 98; 104; 118; 125; 132; 149; 158; 170; 208; 221; 248; 251; 273
Social sciences research; 69
Societal impact; 83; 92; 162; 170; 177; 182
Stratified evaluation regime; 154
Subject classification; 228
Survey; 192
Technoscientific capitalism; 21
Text and data mining (TDM); 40
Theory; 29
Third mission; 145
Topic concentration; 208
Transdisciplinarity; 182
Transformative niches; 48
Trust in peer review; 25
Trusted media; 170
Ukraine; 154
University; 34
University autonomy; 141
University journals; 154
University-government collaboration; 83
Unwanted loneliness; 164
Value capture; 48
Value-based evaluation; 73
Values-based scholarship array; 53

INDEX OF AUTHORS

Authors

Each author is followed by the session(s) in which their contribution appears

Agnoloni, Tommaso (2.3A)
Alhuay-Quispe, Joel (Poster)
Allnutt, Vanessa (3.1)
Angioni, Simone (Poster)
Antonioli, Davide (2.3B)
Anzidei, Marco (2.2A)
Audet, René (3.1)
Aviv-Reuven, Shir (3.2)

Bağcı, Esma (Poster)
Baglioni, Miriam (Poster)
Bardi, Alessia (Poster)
Bautista-Puig, Núria (2.4B; 2.5)
Bautista-Ynofuente, Lourdes (Poster)
Biernaux, Judith (1.2)
Birch, Kean (2.3A)
Blasi, Brigida (2.2B)
Bonaccorsi, Andrea (3.3)
Bonfanti, Pietro (2.2A)
Bornmann, Lutz (3.2)
Brasil, André (2.2A)
Bueno Doral, Tamara (Poster)

Calonaci, Benedetta (Poster; 2.2B)
Carazzolo, Valentina (2.4A)
Carosi, Alessandro (Poster)
Céspedes, Lucía (3.1)
Chan, Christopher (2.1)
Chen, Xiaoting (2.2B)
Christie, Stuart (2.1)
Cilione, Emanuela (Poster)
Clemente, Augusto (2.2A)
Coniglione, Maura (2.4A)
Criveller, Margherita (3.1)
Cupar, Drahomira (1.1)

De Angelis, Deborah (1.2)
De Vita, Sandro (2.2A)
Decadt, Brigitte (1.2)
Delzenne, Eugénie (2.3A)
Di Donato, Francesca (Keynote, 2.2A)
Di Luca, Monica (Poster)
Di Renzoni, Andrea (3.3)
Ding, Alvin (2.1)
Dobre, Silvia (2.1)
Doğan, Güleda (Poster; 2.3B)
Dorsch, Isabelle (2.2B)
Du, Luyu (2.2B)

Engels, Tim (2.3A)
Estradivari (2.5)
Eykens, Joshua (Poster)

Falzea, Luca (2.4A)
Faro, Sebastiano (2.3A; 2.4A)
Florio, Isabella (Poster)
Fortier, Anne-Marie (3.1)
Frattini, Federico (2.3B)
Fuentes-Fernández, Rubén (Poster)

Galimberti, Paola (Poster; 3.1)
García-Castillo, Noelia (Poster)
Gedutis, Aldis (2.1)
Gerland, Eva (1.2)
Giglia, Elena (3.1)
Giménez Toledo, Elea (2.5)
Giovinazzi, Laura (3.2)
Gracy, Noelle (2.1)
Gruppo di lavoro CoARA@INGV (2.2A)

Hara, Yasushi (2.4A)
Hashizume, Hiroshi (2.2B)
Hayashi, Takayuki (2.3A; 2.3B)
Herrera-Torres, Lucía (2.2A)
Holm, Jon (3.3)

Järvelin, Anne-Mari (Poster)
Jermen, Nataša (2.5)
Jia, Arilys (3.1)

Kallio, Miki (Poster)
Karube, Masaru (2.4A)
Kreth, Quintin (2.3A)

Laakso, Mikael (1.1)
Larivière, Vincent (3.1)
Lecocq, Sharon (1.2)
Lehtinen, Vilma (Poster)

Lombardo, Gabi (Keynote)
Lorenzo-Quiles, Oswaldo (2.2A)
Lozano, Rodrigo (2.5)

Maddi, Abdelghani (2.2B)
Malgarini, Marco (2.2B)
Manfredini, Carolina (3.1)
Mannocci, Andrea (Poster; 2.2A)
Martínez-Cardama, Sara (2.4B)
Martorell Camps, Marta (2.2A)
Mazzanti, Massimiliano (2.3B)
Melander, Sakari (2.4A)
Melinščak Zlodi, Iva (3.3)
Melkers, Julia (2.2B)
Michaud, Alexandra (3.1)
Migliorelli, Giorgia (Poster)
Miliani, Costanza (3.3)
Missiroli, Loretta (2.3B)
Miyazawa, Yuki (2.4A)
Molteni, Silvia (Poster)
Montagna, Chiara (2.2A)
Montorsi, Arianna (2.4A)
Morais, Rita (2.3B)
Mu, Jingwen (2.1)

Nakano, Asa (2.2B)
Napolitano Sanchez, Alda Maria (2.4B)
Nazarovets, Maryna (2.4B)
Niemi, Janne (3.2)
Niemi, Laura (Poster)
Numajiri, Honami (2.3A)

Ochsner, Michael (1.1; 2.5)
Okamoto, Shohei (2.3A)
Orduña-Malea, Enrique (2.5)
Oshiumi, Keiichi (2.3B)
Otero Borges, Lisandra (2.4B)

Pavlović, Nikolina Peša (1.1)
Pavone, Gina (Poster)
Pecoraro, Fabrizio (2.2A)
Pensec, Emmanuelle (2.4B)
Perez Clausen, Viviana (2.2A)
Peruginelli, Ginevra (2.3A; 2.4A)
Pianta, Mario (1.2)
Piergentili, Irene (2.2B)
Pietarinen, Ahti-Viekko (2.1)
Pietilä, Maria (2.4A)
Plume, Andrew (2.1)
Pölönen, Janne (Poster; 1.1; 2.4A)
Pourbaix, Jérôme (1.2)
Provost, Lottie (2.2A)

Ranchino, Maria Adelaide (Poster; 2.2B)
Rauhala, Marika (Poster)
Reale, Emanuela (2.4A; 2.5)
Richebourg, Aube (3.1)
Robinson-Garcia, Nicolas (2.2B)
Roelants, Nienke (1.2)
Romagnosi, Sandra (2.2B)
Rosenfeld, Ariel (3.2)
Rubbia, Giuliana (Poster; 2.2A; 2.4a)

Santorsa, Sara (Poster)
Sasaki, Yu (2.2B)
Savard-Arseneault, Adrien (3.1)
Schiffbänker, Helene (2.4A)
Seferoğlu, Ceren Bilge (Poster)
Sezen, Ayça Nur (Poster)
Sigurðarson, Eiríkur Smári (2.1)
Simon, Madeline (1.2)
Singh, Pragya (2.1)
Sivertsen, Gunnar (2.5)
Söderman, Mira (Poster)
Soycan, Nihan (1.2)
Spinello, Andrea Orazio (2.4A)

Taşkın, Zehra (1.2; 2.3A; 2.3B)
Teunens, Titouan (1.2)
Tondi, Rosaria (2.2A)
Torny, Didier (3.1)

Valente, Paolo (3.1)
Vandewalle, Eline (2.3A)
Vanholsbeeck, Marc (1.2)
Vas, Virag (2.3A)
Vigni, Roberta (3.1)

Wahlfors, Leena (1.1)
West, Josie R. (1.1)
Williams, Geoffrey (2.4B)
Wyber, Stephen (1.2)

Yance-Yupari, Victoria (Poster)
Yoshizawa, Go (2.3A)
Yu, Lei (2.4A)

Zuccala, Alesia A. (3.2)

The **RESSH Conference** is the regular event of the international association ENRESSH – European Network for Research Evaluation in the Social Sciences and Humanities. It brings together specialists in research evaluation and policy, with a particular focus on the social sciences, humanities, and the arts. The 2026 edition, held in Florence from 22 to 24 April 2026, is organized by the Institute of Legal Informatics and Judicial Systems of the National Research Council, Italy (CNR-IGSG) and ENRESSH.

Research Evaluation and Scientific Autonomy under Pressure

Research evaluation is being reshaped by shifting global powers, rising security concerns and rapid technological developments such as artificial intelligence. At the same time, science faces growing public scrutiny, as misconduct, plagiarism, data fabrication and predatory journals challenge its credibility and ethics. Debates on scientific integrity are increasingly entangled with geopolitical interests and technological change, revealing evaluation as a politicised tool. As political pressures constrain international collaboration, open science and academic mobility, evaluation systems become key arenas where diverse actors negotiate research sovereignty and academic freedom. In parallel, technological advances are transforming evaluation procedures, with AI-driven tools – algorithmic indicators, automated peer review, and large-scale data analytics – offering efficiency while raising concerns about reliability, transparency and fairness.
In this scenario, the academic community needs to critically reassess how evaluation affects not only excellence and impact, but also individual academic freedom, institutional autonomy, ethical responsibility, inclusivity and trust in research.

La **Conferenza RESSH** è l'evento di riferimento dell'Associazione internazionale ENRESSH – European Network for Research Evaluation in the Social Sciences and Humanities. Riunisce esperti nella valutazione della ricerca e nelle politiche scientifiche, con particolare attenzione alle scienze sociali, umanistiche e alle arti. L'edizione 2026, ospitata a Firenze dal 22 al 24 aprile 2026, è organizzata dall'Istituto di Informatica Giuridica e Sistemi Giudiziari del Consiglio Nazionale delle Ricerche (CNR-IGSG) e da ENRESSH.

Valutazione della ricerca e autonomia scientifica sotto assedio

La valutazione della ricerca è investita da profondi mutamenti: gli equilibri globali si spostano, le preoccupazioni per la sicurezza crescono e lo sviluppo tecnologico, in particolare l'intelligenza artificiale, avanza a ritmo accelerato. Al tempo stesso, la scienza è sottoposta a un controllo pubblico sempre più stringente: casi di cattiva condotta scientifica, plagio, fabbricazione di dati e riviste predatorie ne mettono a rischio la credibilità e l'integrità etica. Il dibattito sull'integrità scientifica si intreccia sempre più con gli interessi geopolitici e il cambiamento tecnologico, rivelando come la valutazione possa diventare uno strumento politicamente orientato. Mentre le pressioni politiche limitano la collaborazione internazionale, la scienza aperta e la mobilità accademica, i sistemi di valutazione si trasformano in arene in cui diversi attori negoziano sovranità della ricerca e libertà accademica. In parallelo, i progressi tecnologici stanno ridefinendo le procedure valutative: strumenti basati sull'IA offrono nuove straordinarie potenzialità, sollevando al tempo stesso interrogativi su affidabilità, trasparenza ed equità.
In questo scenario, la comunità scientifica è chiamata a ripensare criticamente il modo in cui la valutazione incide non solo sull'eccellenza e sull'impatto della ricerca, ma anche sulla libertà individuale dei ricercatori, sull'autonomia istituzionale, sulla responsabilità etica, sull'inclusività e sulla fiducia nella scienza.

ISBN: 9788890576485
Doi: https://doi.org/10.32091/RESSH2026