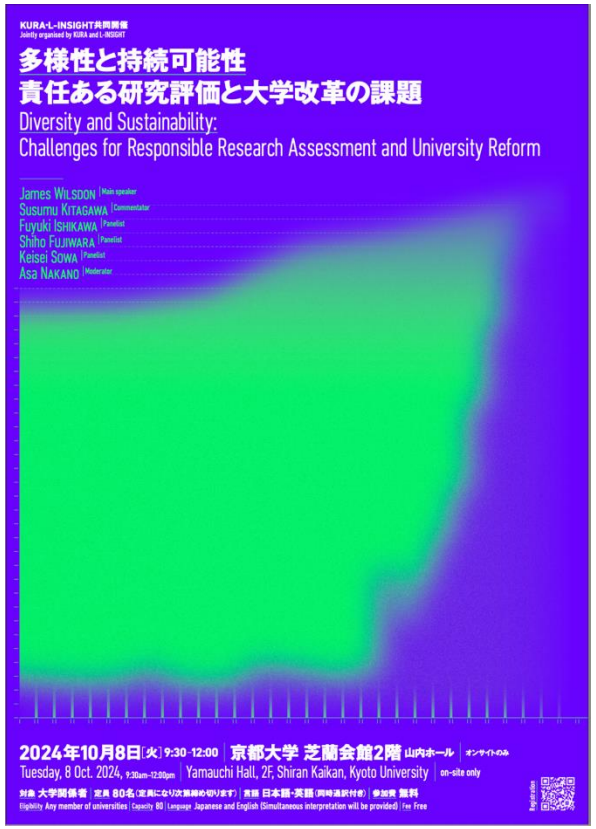
RESEARCH ON RESEARCH INSTITUTE

Diversity and sustainability: challenges for responsible research assessment & university reform

James Wilsdon, RoRI & UCL

Kyoto University, 8 October 2024



What I will aim to cover in ~45 mins:

1. A growing global movement for reform of research assessment systems
2. Defining RRA: responsible research assessment
3. A few lessons – good and bad! – from our experiences with these debates in the UK system
4. Five big trends shaping these debates
5. RoRI's Global Observatory of Responsible Research Assessment
6. Three questions for the Japanese system – and for you all!

本日カバーする予定の内容:

1. 研究評価システムの改革を求める世界的な動きの高まり
2. RRAの定義:責任ある研究評価
3. 英国のシステムにおけるこれらの議論から得た教訓(良いものも悪いものも含めて)
4. これらの議論を形成する5つの大きな傾向
5. RoRIの責任ある研究評価に関するグローバル・オブザーバトリー
6. 日本のシステムに対する3つの質問、そして皆さんにも!

A growing global movement for research assessment reform

研究評価改革を求める世界的な動きの高まり

May 2013

DORA

https://sfdora.org

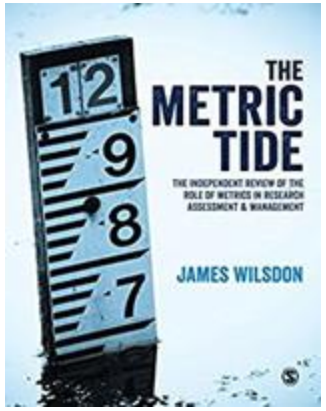
Nov 2013

Science in Transition

Mar 2015

Leiden Manifesto

Jul 2015



Metric Tide Report

Sept 2018

Plan S

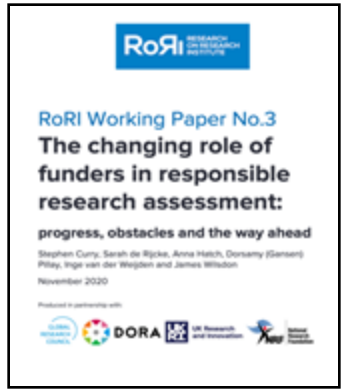
Nov 2019

Netherlands initiative

Jul 2020

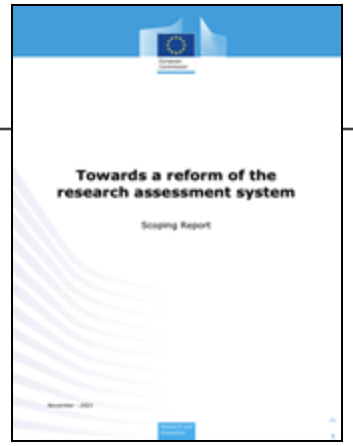
Hong Kong Principles

Nov 2020



Report for GRC meeting

Nov 2021



EC Scoping Report

July 2022

RoRI RESEARCH ON RESEARCH INSTITUTE

レビュー論文
「専門的な改革運動としての責任ある測定基準の台頭：コレクティブアクションとしての説明」

QSS

an open access journal

Check for updates

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

DOI: https://doi.org/10.1162/qss_a_00280

Peer Review: https://www.webofscience.com/api/gateway/wos/peer-review/10.1162/qss_a_00280

Supporting Information: https://doi.org/10.1162/qss_a_00280

Received: 9 December 2022
Accepted: 26 September 2023

Corresponding Author:
Alexander Rushforth
a.d.rushforth@cwts.leidenuniv.nl

REVIEW ARTICLE

The rise of responsible metrics as a professional reform movement: A collective action frames account

Alexander Rushforth[1] and Björn Hammarfelt[2]

[1]Centre for Science and Technology Studies (CWTS), Leiden University, Leiden, The Netherlands
[2]Swedish School of Library and Information Science, University of Borås, Borås, Sweden

Keywords: evaluative bibliometrics, research assessment reforms, responsible metrics, responsible research assessment, social movements

ABSTRACT

Recent years have seen a rise in awareness around "responsible metrics" and calls for research assessment reforms internationally. Yet within the field of quantitative science studies and in research policy contexts, concerns about the limitations of evaluative bibliometrics are almost as old as the tools themselves. Given that many of the concerns articulated in recent reform movements go back decades, why has momentum for change grown only in the past 10 years? In this paper, we draw on analytical insights from the sociology of social movements on *collective action frames* to chart the emergence, development, and expansion of "responsible metrics" as a *professional reform movement*. Through reviewing important texts that have shaped reform efforts, we argue that hitherto, three framings have underpinned the responsible metrics reform agenda: the *metrics skepticism* framing, the *professional-expert* framing, and the *reflexivity* framing. We suggest that although these three framings have coexisted within the responsible metrics movement to date, cohabitation between these framings may not last indefinitely, especially as the responsible metrics movement extends into wider research assessment reform movements.

日本学術会議 科学者委員会研究評価分科会提言
「学術の振興に寄与する研究評価を目指して －望ましい研究評価に向けた課題と展望－」

RECOMMENDATION

Toward Research Evaluation for the Advancement of Science: Challenges and Prospects for Desirable Research Evaluation

November 25, 2021
Science Council of Japan
Committee for Scientific Community
Subcommittee on Research Evaluation

DORA

About Communities Meetings Grants Conta

Reformscape The Declaration Project TARA DORA Reports News and Resources Sign DORA

25,450 individuals and organizations in 166 countries have signed DORA to date.

Including 19 organisations in Japan!

研究評価に関するサンフランシスコ宣言(DORA)： 25450人、166か国の組織がDORAに署名

日本からも19機関が署名済み！

- The Japanese Society of Applied Glycoscience
- Japanese Association for the Advancement of Science
- The University of Tokyo
- Nartional Institutes for Quantum Science and Technology
- RIKEN
- Institute for Quantitative Biosciences
- Japan Science and Technology Agency
- The Molecular Biology Society of Japan
- GN Corporation Co.Ltd
- The Botanical Society of Japan

- CYTOLOGIA
- The Japan Mendel Society
- Genes to Cells
- The Union of Japanese Societies for Biological Science (SEIKAREN)
- The Japan Neuroscience Society
- National Institutes of Biomedical Innovation , Health and Nutrition
- Journal of Illusion Japan
- The Japanese Biochemical Society Japan
- Cell Structure and Function (a journal published by Japanese Society of Cell Bio

RoRI RESEARCH ON RESEARCH INSTITUTE

As of 4 October 2024, 792 organisations from 55 countries have signed the CoARA agreement (none – so far! – from Japan…)

「研究評価改革の推進のための有志連合（Coalition for Advancing Research Assessment, COARA）」は、2024年10月4日現在、55カ国から792の機関が署名済み（日本からはまだ）


nature

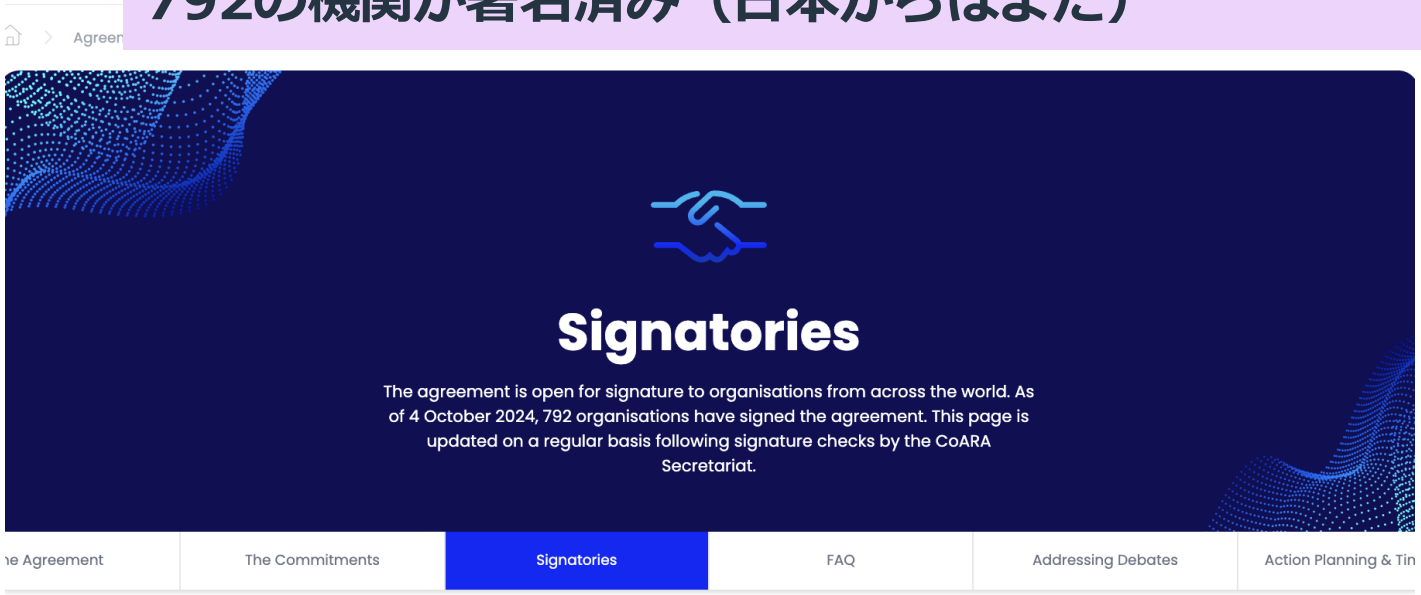
Explore content | About the journal | Publish with us | Subscribe

nature > editorials > article

EDITORIAL | 27 July 2022

Support Europe's bold vision for responsible research assessment

There have been many initiatives to combat the distorting effect of research assessment exercises. The latest looks like it might work


Agreen

Signatories

The agreement is open for signature to organisations from across the world. As of 4 October 2024, 792 organisations have signed the agreement. This page is updated on a regular basis following signature checks by the CoARA Secretariat.

he Agreement | The Commitments | Signatories | FAQ | Addressing Debates | Action Planning & Tin

All | Academies, learned societies, and their associations, and associations of researchers | Albania | Andorra | Australia | Austria

Belgium | Bosnia and Herzegovina | Brazil | Bulgaria | Canada | Chile | Costa Rica | Country/scope | Croatia | Cyprus

Czechia | Denmark | Estonia | Europe | Finland | France | Georgia | Germany | Global | Greece | Hungary

Iceland | Ireland | Italy | Kenya | Kosovo | Latvia | Lithuania | Luxembourg | Mexico | Moldova | Montenegro

「研究評価の未来ー討議資料」
過去10年の文献に基づくと5つの結論に

Building on the past decade of scientific literature and advocacy work, there are five main conclusions.

1. The imperative to rethink the way in which research individuals, institutions and outputs are evaluated is clear and urgent. Maintaining research integrity and quality, maximizing diverse, inclusive and non-discriminatory science, and optimizing science for the global public good are major drivers, set in the context of a fast-changing world.

2. The way in which research is commissioned, funded, delivered and communicated is evolving at pace. Moves towards mission-oriented and transdisciplinary science, open science frameworks, evolving models of peer review, the use of AI and machine learning and the rapid rise of social media are changing traditional ways of doing and communicating research, requiring new thinking on research evaluation systems and the metrics and peer review processes underpinning it. More, and urgent, research is needed to future-proof these systems.

3. There is an imperative for more balanced research evaluation systems with both quantitative and qualitative indicators that value multiple forms of research output, processes and activities. However, stating that qualitative peer review processes are at least as important as bibliometrics is not straightforward and is further complicated by different parts of the world being at different stages in developing their assessment systems: in some, debates on research evaluation reform are quite advanced, in others they are nascent or absent.

Responsible Research Assessment Working Group

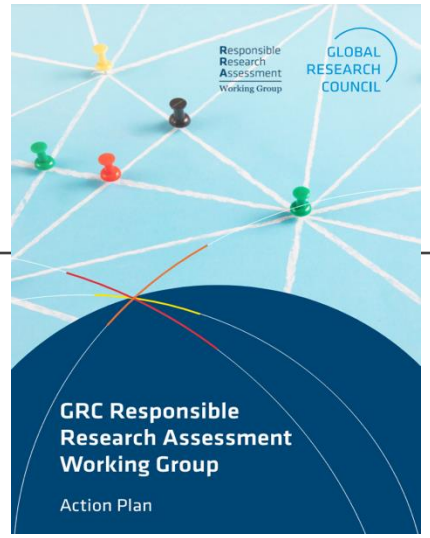
研究資金配分機関のグローバルなネットワーク、Global Research Councilの中の責任ある研究評価ワーキンググループの活動の軌跡

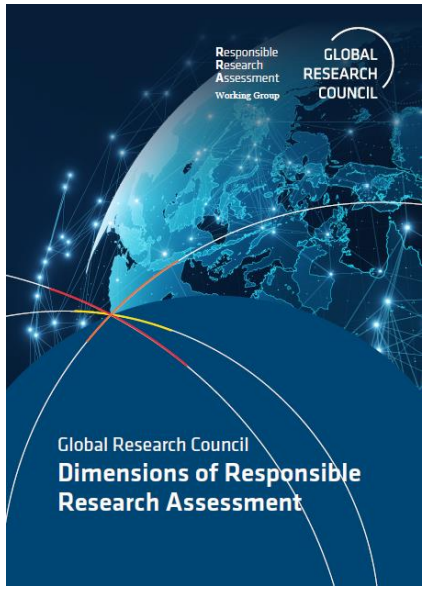
October 2020
Survey of GRC members
↓
RoRI report

November 2020
GRC – RRA Conference
↓
over 1000 attendees

May 2021
GRC Annual Meeting
↓
GRC Endorsed Call to Action

September 2021
↓
Established GRC RRA Working Group





RRA Working group set up (2022)
↓
Action Plan

May 2024
↓
Publish 11 Dimensions of RRA

May 2024
Launch survey of GRC participants (with RoRI)

May 2024
Launch a framework for RRA case studies

Responsible metrics can be understood in terms of:

責任ある研究評価・測定の理解「「指標は真の力を秘めている。指標は価値観、アイデンティティ、生活を構成する要素である」

“Metrics hold real power: they are constitutive of values, identities and livelihoods.”

Robustness: basing metrics on the best possible data in terms of accuracy and scope;

Humility: recognizing that quantitative evaluation should support – but not supplant – qualitative, expert assessment;

Transparency: keeping data collection and analytical processes open and transparent, so that those being evaluated can test and verify the results;

Diversity: accounting for variation by field, using a variety of indicators to reflect and support a plurality of research & researcher career paths;

Reflexivity: recognizing the potential & systemic effects of indicators and updating them in response.

「頑健性」

Robustness: データの正確性と範囲

「謙虚さ」

Humility: 定量的評価は定性・専門家評価を支援すべきであること

「透明性」

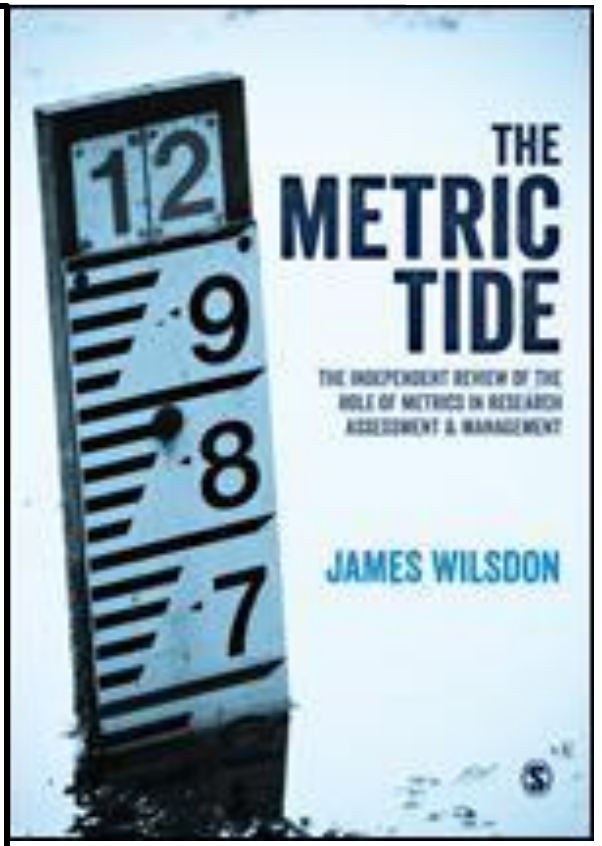
Transparency: データと分析プロセスの公開性・透明性が確保され、その結果の妥当性検証が可能であること

「多様性」

Diversity: 分野に応じた説明、研究者や研究者のキャリアパスの多層性を反映・支援するような指標の使用

「省察性」

Reflexivity: 指標がもつ潜在的かつシステム上の効果に応じた更新



訳文は以下による。日本学術会議「提言：学術の振興に寄与する研究評価を目指して～望ましい研究評価に向けた課題と展望」2021年9月（pp.71-72） https://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-25-t312-1.pdf

A frame extension from metrics to RRA…

「責任ある指標」から「責任ある研究評価」への枠組みの拡張

QSS

an open access journal

Check for updates

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

DOI: https://doi.org/10.1162/qss_a_00280

Peer Review: https://www.webofscience.com/api/gateway/wos/peer-review/10.1162/qss_a_00280

Supporting Information: https://doi.org/10.1162/qss_a_00280

Received: 9 December 2022
Accepted: 26 September 2023

Corresponding Author:
Alexander Rushforth
a.d.rushforth@cwts.leidenuniv.nl

REVIEW ARTICLE

The rise of responsible metrics as a professional reform movement: A collective action frames account

Alexander Rushforth[1] and Björn Hammarfelt[2]

[1]Centre for Science and Technology Studies (CWTS), Leiden University, Leiden, The Netherlands
[2]Swedish School of Library and Information Science, University of Borås, Borås, Sweden

Keywords: evaluative bibliometrics, research assessment reforms, responsible metrics, responsible research assessment, social movements

ABSTRACT

Recent years have seen a rise in awareness around "responsible metrics" and calls for research assessment reforms internationally. Yet within the field of quantitative science studies and in research policy contexts, concerns about the limitations of evaluative bibliometrics are almost as old as the tools themselves. Given that many of the concerns articulated in recent reform movements go back decades, why has momentum for change grown only in the past 10 years? In this paper, we draw on analytical insights from the sociology of social movements on *collective action frames* to chart the emergence, development, and expansion of "responsible metrics" as a *professional reform movement*. Through reviewing important texts that have shaped reform efforts, we argue that hitherto, three framings have underpinned the responsible metrics reform agenda: the *metrics skepticism* framing, the *professional-expert* framing, and the *reflexivity* framing. We suggest that although these three framings have coexisted within the responsible metrics movement to date, cohabitation between these framings may not last indefinitely, especially as the responsible metrics movement extends into wider research assessment reform movements.

"Since the late 2010s, a notable *frame extension* (Benford & Snow, 2000) of the responsible metrics reform movement has occurred, from the more specific focus on appropriate uses of bibliometrics into a widened framing of "responsible research assessment."

A 2020 report by members of the Research on Research Institute (including authors of DORA, the Leiden Manifesto and the Metric Tide texts) defined responsible research assessment as "an umbrella term for approaches to assessment which incentivize, reflect and reward the plural characteristics of high-quality research, in support of diverse and inclusive research cultures" (Curry et al., 2020, p. 7). A large number of texts have emerged supporting this widened agenda (CoARA, 2022; EC, 2017; EU, 2022; LERU, 2022; UNESCO, 2021)."

Defining RRA

Responsible research assessment (RRA) is an umbrella term for approaches to assessment which incentivise, reflect and reward the plural characteristics of high-quality research, in support of diverse and inclusive research cultures.

While notions of 'responsible metrics' can be applied at a micro level to indicators themselves, the idea of RRA encourages funders, research institutions, publishers and others to focus attention on the methodologies, systems and cultures of research assessment.

責任ある研究評価（RRA）とは、多様で包括的な研究文化を支援するために、質の高い研究の多様な特性を奨励し、反映し、評価するアプローチを包括する用語

「責任ある指標」という概念は、指標そのものにミクロレベルで適用することができますが、RRAの考え方は、資金提供者、研究機関、出版社、その他の関係者に研究評価の方法、システム、文化に注目するよう促すもの

RoRI RESEARCH ON RESEARCH INSTITUTE

RoRI Working Paper No.3

The changing role of funders in responsible research assessment:

progress, obstacles and the way ahead

Stephen Curry, Sarah de Rijcke, Anna Hatch, Dorsamy (Gansen) Pillay, Inge van der Weijden and James Wilsdon

November 2020

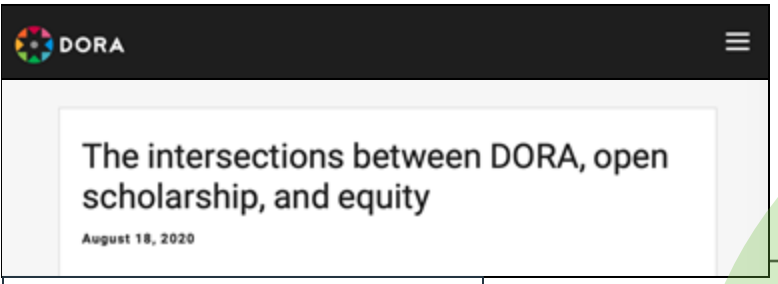
Produced in partnership with:

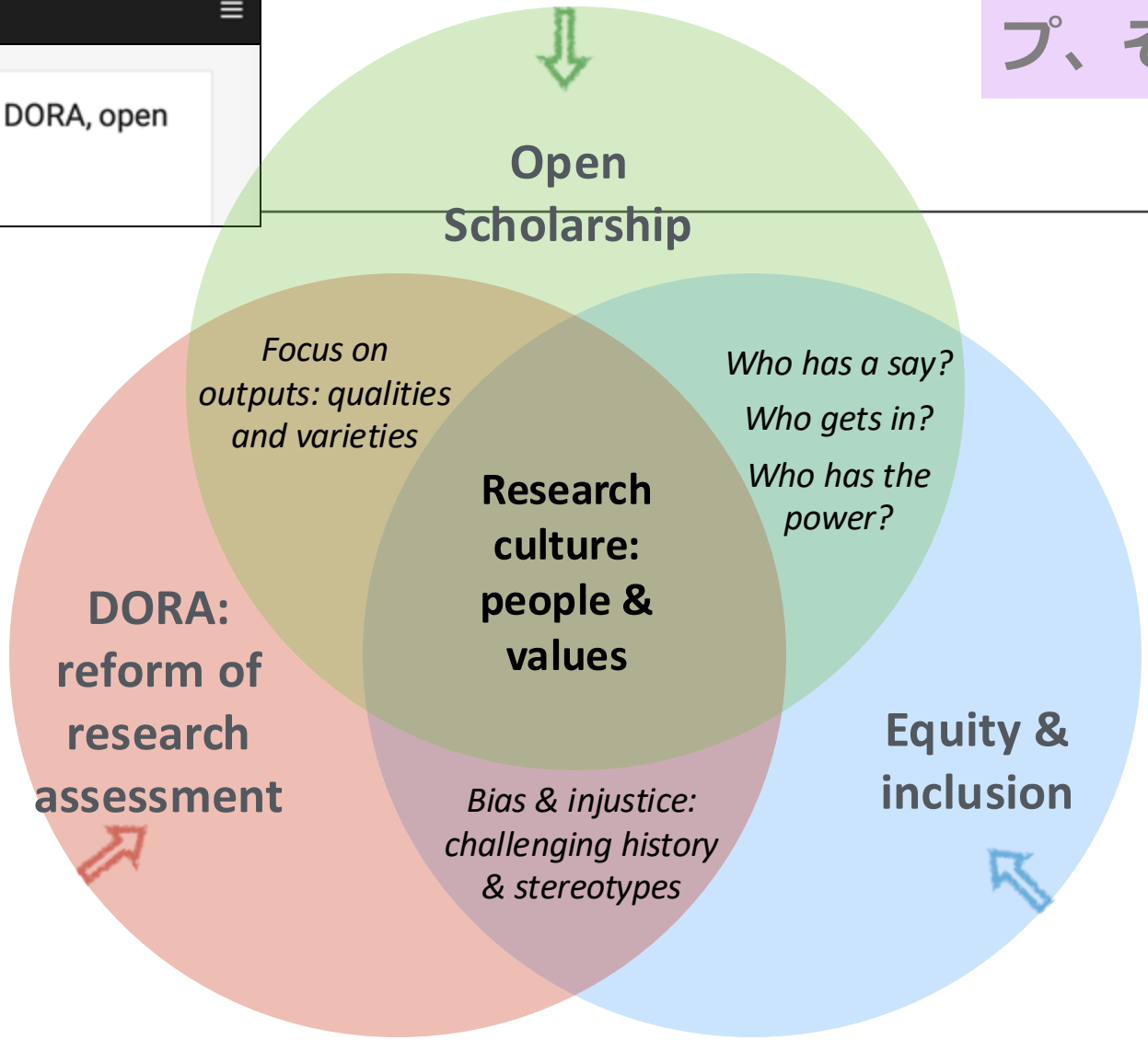
GLOBAL RESEARCH COUNCIL · DORA · UK Research and Innovation · NRF National Research Foundation

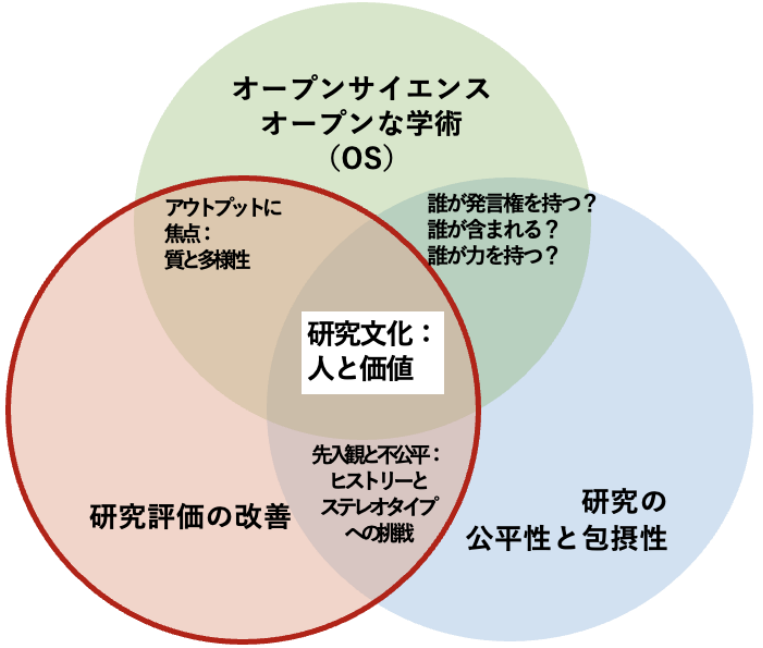
Turning tides: RRA and research cultures

潮流の変わり目:RRAと研究文化

DORA、オープン・スカラーシップ、そして公平性の交差点







RoRI RESEARCH ON RESEARCH INSTITUTE

日本と責任ある研究評価との関わり

nature

Explore content ∨ About the journal ∨ Publish with us ∨ Subscribe

nature > nature index > article

NATURE INDEX | 29 August 2024

Japan moves to halt long-term postgraduate decline by tripling number of PhD graduates

Plan aims to elevate the status of PhD holders and give them greater career mobility.

By Tim Hornyak

nature

Explore content ∨ About the journal ∨ Publish with us ∨ Subscribe

nature > news > article

NEWS | 19 July 2022

'I feel disposable': Thousands of scientists' jobs at risk in Japan

Universities are terminating workers on fixed-term contracts – but unions say the actions do not reflect the purpose of employment laws introduced ten years ago.

By Tim Hornyak

Dozens of research teams at RIKEN's network of basic research laboratories will be closed down early next year. Credit: Kyodo News Stills/Getty

RECOMMENDATION

Toward Research Evaluation for the Advancement of Science: Challenges and Prospects for Desirable Research Evaluation

November 25, 2021
Science Council of Japan
Committee for Scientific Community
Subcommittee on Research Evaluation

Peer reviewed analysis from world leading experts

Subscribe

EASTASIAFORUM

About Contributors The Quarterly Submissions

NORTHEAST ASIA

Japan's University Fund is ill-equipped to stem decline in research performance

Published: 29 December 2023
Reading Time: 5 mins

Japan
index

Let's free up PhD potential

Not enough prospects in universities and industry means Japan isn't making the most of valuable researchers.

Ayuko Hoshino calls for a new view to widen career pathways for PhD holders.

Japanese science has a problem: there are too many PhD holders and not enough senior roles in universities for them to move into. This is partly caused by a well-meaning, but flawed policy to promote Japanese research that dates back almost three decades.

In 1996, Japan began a plan to boost the number of its academic researchers with a PhD but who are not yet in permanent faculty positions. The country aimed to produce 10,000 of these postdoctoral roles and by 2006 it had exceeded this goal, creating more than 16,000 positions. This leaves a fairly obvious question: what happens to a researcher after they've completed a postdoc? There hasn't been a serious enough effort to create a career pathway for these researchers in academia. Employment in industry is also an uphill battle for them because – although progress has been made – Japanese businesses on the whole still don't fully appreciate PhDs as a qualification.

"Japanese businesses on the whole still don't fully appreciate PhDs."

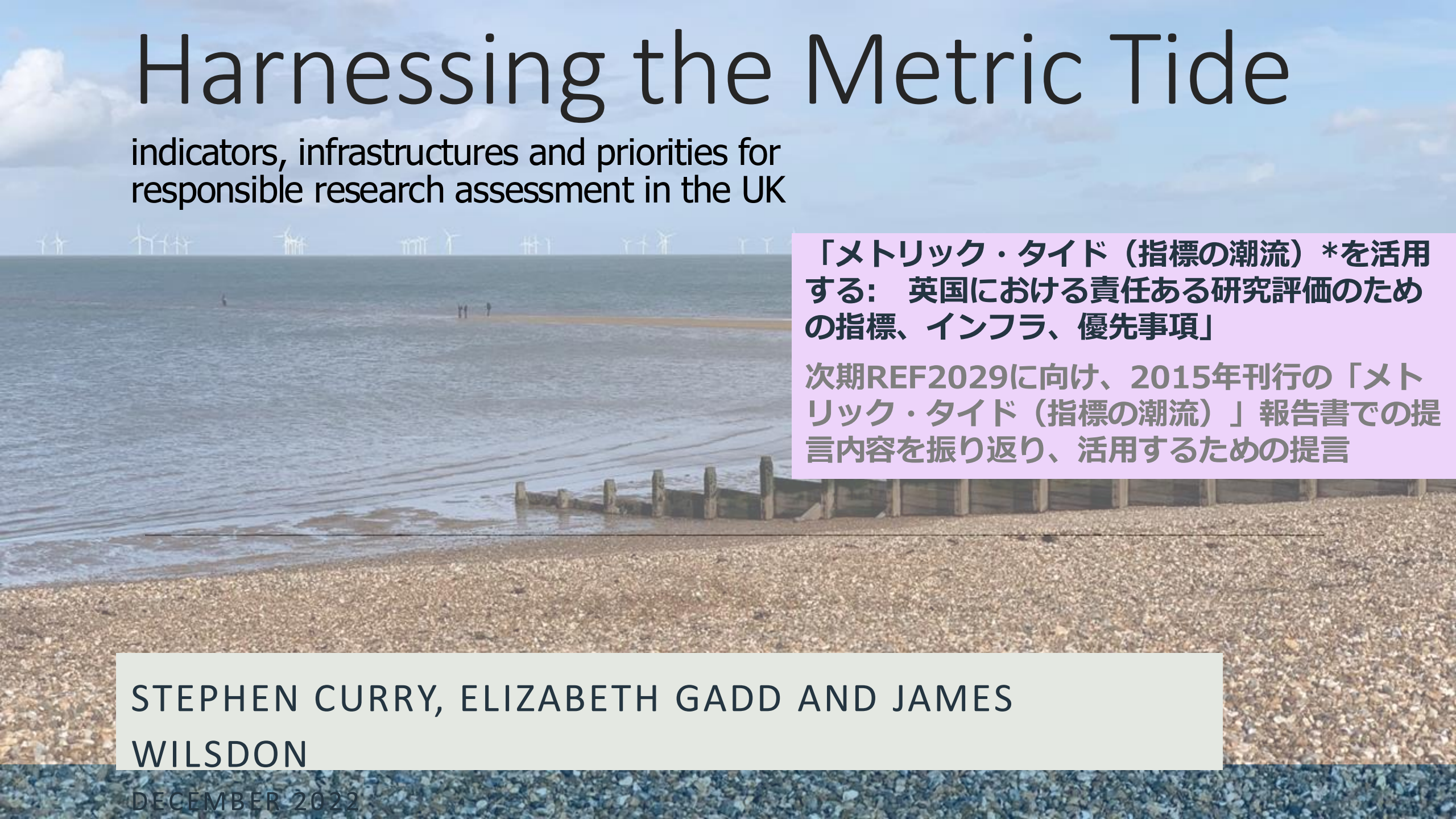
Many students here in Japan increasingly believe that now so many postdocs aged 40 and above that they're now competing with their younger colleagues for increasingly narrow career pathways. In 2018, for instance, 28.2% of postdocs were in this age group compared with 16.4% in 2012.

Instead of benefiting from these educated and experienced PhD graduates, Japan is squandering them. This needs to change – not least because it is undermining efforts to improve the gender balance in science, technology, engineering and maths (STEM) subjects, something

RoRI RESEARCH ON RESEARCH INSTITUTE

Harnessing the Metric Tide

indicators, infrastructures and priorities for responsible research assessment in the UK



「メトリック・タイド（指標の潮流）*を活用する:　英国における責任ある研究評価のための指標、インフラ、優先事項」

次期REF2029に向け、2015年刊行の「メトリック・タイド（指標の潮流）」報告書での提言内容を振り返り、活用するための提言

STEPHEN CURRY, ELIZABETH GADD AND JAMES WILSDON

DECEMBER 2022

REFの歴史、目的、審査パネル、専門家の審査

History 歴史

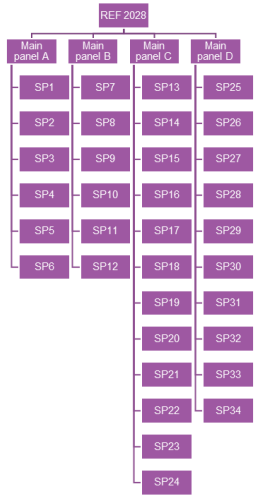
RSE 1986 → RSE 1989 → RAE 1992 → RAE 1996 → RAE 2001 → RAE 2008 → REF 2014 → REF 2021 → REF 2029

Purposes 目的

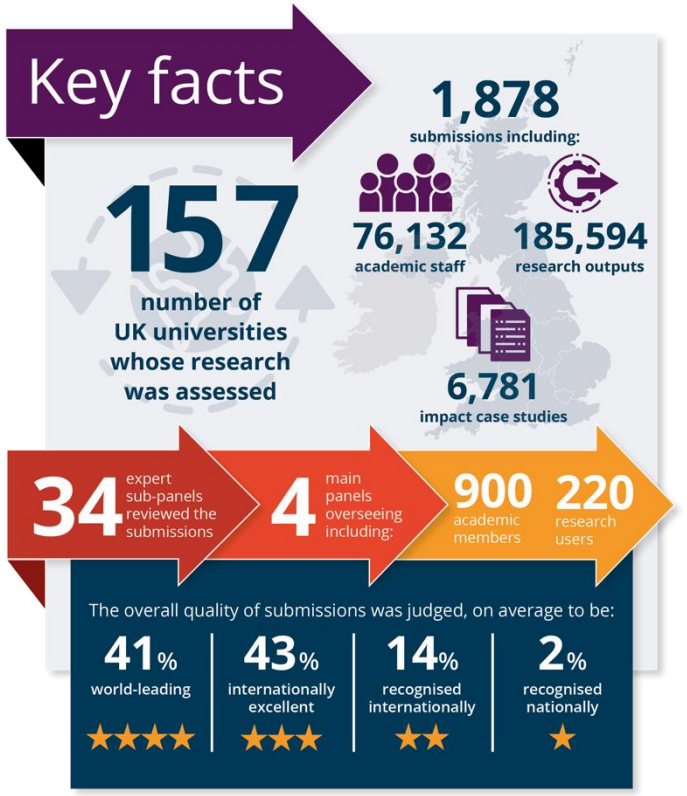
- Inform the allocation of block-grant research funding to HEIs based on research quality
- Provide accountability for public investment in research and produce evidence of the benefits of this investment
- Provide insights into the health of research in HEIs in the UK

Panels 審査パネル

- 34 Disciplinary Panels
- 4 Main Panels
  - A: Health and Life Sciences
  - B: Physical Sciences and Engineering
  - Social Sciences
  - Arts and Humanities



Key facts



Expert Review

- Quality judgements by academic and non-academic experts
- Informed by metrics where appropriate
- Responsible use of metrics

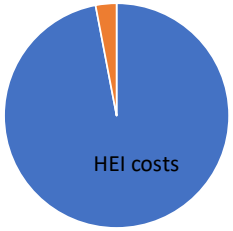
専門家による（ピア）審査

- アカデミア内外の専門家による質的判断
- 必要に応じ指標を参考
- 指標の責任ある利用

Elements of Assessment 評価の要素

People, culture and environment	Contribution to knowledge and understanding	Engagement and impact
• Institution-level and disciplinary-level evidence statements	• Research outputs (2.5x volume) • Disciplinary-level evidence statement	• Impact case study/ies • Disciplinary-level evidence statement
25%	50%	25%
人・文化・環境 25%	知識と理解への貢献 50%	エンゲージメントとインパクト 25%

Costs コスト



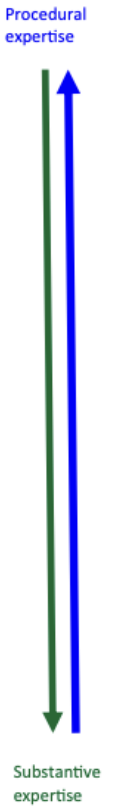
- Cost of REF 2021 estimated at £471M
- 3-4% of funding allocated
- 97% costs incurred by HEIs

RoRI RESEARCH ON RESEARCH INSTITUTE

機関間、機関、内部組織等、レベルごとの審査
形態：目標、ステイクホルダー、方法

Table 1 Stratified assessments and forms of expertise (adapted from Oa

Strata of research assessment	Scope of assessment	Purposes	Stakeholders	Methods and measures
Supra- organisational	International, national, multidisciplinary and disciplinary	Policy and strategic decisions; public resource allocation; political debate; construction of field identity and status; system performance; performance-based funding	International organisations; national government; funding bodies; national strategic bodies; interest groups and sector representative bodies	Performance-based research funding systems; system ratings; economic indicators; bibliometrics; cultural indicators; impact assessments; expert descriptions; scenarios; peer review; Delphi panels; consultation; public debate
Organisational	Research organisations; funding schemes; publishing investments; research units; large-scale programmes and partnerships	Benchmarking; accreditation; social accountability; organisational mission and strategy; reputation/ image management; allocation of funds within organisations; human resources management; capacity building	Funding bodies; quality assurance and audit bodies; research organisations; professional/ external evaluators; publishers; industry/ user bodies; third sector; professional associations; media	Formal evaluations; ratings; peer review; bibliometrics; economic metrics; cultural indicators; inter/national standards; impact, engagement and use studies; benchmarking; quality management; consensus conferences; network studies; case studies
Sub-organisational and para-organisational	Teams; individuals; projects; outputs; outcomes	Access to funds; publication; personnel decisions and workload allocations; awards and recognition; training and development; research decisions (substantive, methodological, practical); decisions on: research synthesis, dissemination, brokerage, KE, PER, open scholarship etc.	Peers; human resources departments; managers; grant awarding bodies; editors and referees; users and partners; brokers; professional associations; etc	Peer review; systematic review; network maps; case studies; public debates; open appraisal; bibliometric counts; alternative metrics etc



Evaluate with the evaluated

評価対象者と評価する

Stakeholders involved in research evaluation should enable and incentivise the **co-design and co-interpretation** of research assessments with research-active and research-enabling staff.

研究評価に関わるステークホルダーは、研究活動に従事するスタッフや研究を可能にするスタッフと協力して、研究評価の**共同設計と共同解釈**を可能にし、奨励すべきである。

Use data for good

DATA FOR GOOD

- Gender pay gaps for research staff;
- % of research staff on short term contracts;
- Measures of research staff wellbeing and contentment in surveys of workplace culture
- Volume of teamwork; collaborations; co-produced research (with users);
- Open research indicators;
- Policy impacts e.g. via citations in policy literatures;
- Peer review work;
- Citizenship contributions (from workload models);
- Measures of support for EDI;
- Effective measures for dealing with bullying and harassment.

良いことのためのデータ

- 研究スタッフの男女間賃金格差
- 短期契約の研究スタッフの割合
- 職場文化調査における研究スタッフの幸福度と満足度の測定
- チームワークの量、共同作業、（ユーザーとの）共同生産の研究
- オープンリサーチ指標
- 政策へのインパクト、政策文書への引用
- ピアレビュー作業
- 市民的貢献
- EDIへのサポート手段
- いじめやハラスメントへの効果的な対策。

THE Times Higher Education

Home News Rankings Jobs Students Events Resources

News

Home Latest Opinion In-depth Leadership Digi

Russell Group ‘nervous’ over downgrading of outputs for REF 2028

Elite universities harbour doubts about shift away from rewarding proven research excellence, says Manchester vice-president, though sector broadly welcomes changes for next assessment exercise

June 21, 2023

Jack Grove

Twitter: @jgro_the

Source: Getty Images

Moving away from “tried-and-tested” methods of evaluating research excellence via outputs-based peer review towards more experimental ways of judging academic culture is causing “nervousness” and “uncertainty” at leading UK universities, a Russell Group leader has admitted.

Under proposed rules for the 2028 Research Excellence Framework, just 50 per cent of an institution’s score will be judged on its research outputs, down from 60 per cent in 2021 and 65 per cent in 2014. Of this 50 per cent, at least 10 per cent will be decided on an institution’s self-described disciplinary contribution, potentially pushing the direct weighting of outputs close to 40 per cent. Under the revamp announced on 15 June, “people, culture and environment” will be worth 25 per cent of an evaluation, the same as “engagement and impact”.

Downgrading research outputs, however, had been raised as a concern by senior

ラッセルグループ（英国のエリート大学）は、次期2028年の研究評価REFにおいて、出版物（Output）の評価割合が相対的に減ったことに「神経質」になっている

ResearchProfessional News

REF2028: 次期研究評価（REF2028）に向けた協議においては、いくつかの困難な課題がある

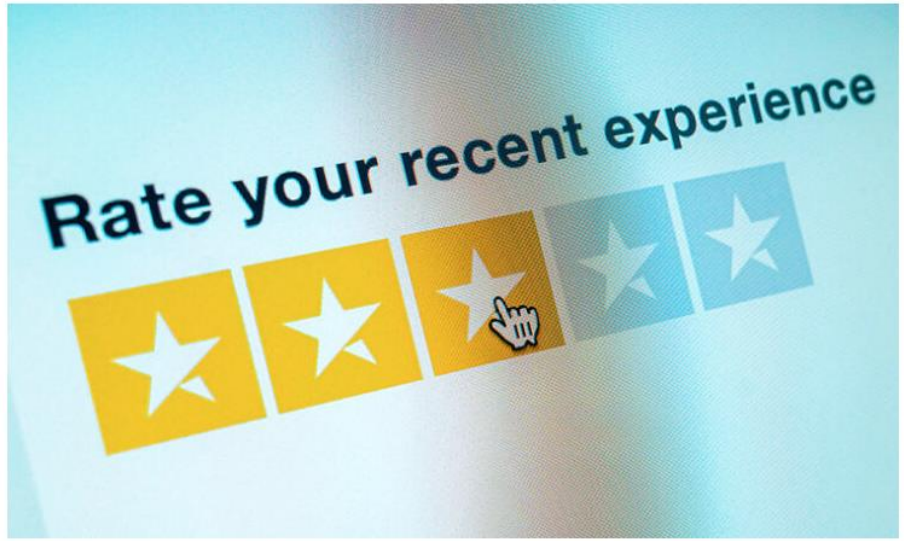
< Go back

OPINION 27 JUN 2023

REF 2028: Consultation has some heavy lifting to do

By Stephanie Smith

Share



Unanswered questions and uncertain trade-offs leave much still up for grabs, says Stephanie Smith

Fortune favours the bold. That seems to be what the team running the Research Excellence Framework

RoRI RESEARCH ON RESEARCH INSTITUTE

Research Professional

8am Playbook Trial Account

HE News Our Institution Help

News

UK Europe Australia Africa New Zealand USA World Views Politica

Bookmark Email Download

20 Oct 23, 07:00

By Stephen Curry, Elizabeth Gadd and James Wilsdon

Related Links

Stephen Curry on X

Elizabeth Gadd on X

James Wilsdon on X

Related Articles

Donelan's woke-bashing puts the REF in the crosshairs

REF 2028: How do you measure culture?

REF 2028: Rebalancing the definition of excellence

REF 2028: What could go wrong?

REF 2028: A clap for the Frap

Plans for REF 2028 should be debated, not throttled

Moves against emphasis on culture are mistaken, say Stephen Curry, Elizabeth Gadd and James Wilsdon

Comment on this article

When the proposals for the next Research Excellence Framework were published in June, reactions were broadly positive.

In the past five years, the acute problems in research cultures have received steadily more attention from university leaders, policymakers and funders. REF 2028's increased attention to people, culture and environment offered a route to tackling these concerns, rebalancing incentives towards the collective and collaborative aspects of a healthy, dynamic and fair research system.

Yet as summer edged into autumn, in some quarters that confidence seems to have collapsed. Critics are circling the REF 2028

REF 2028の計画は議論されるべきであり、抑制されるべきではない:　REF 2028において、出版物の相対的比重が減り、研究文化の比重が高まったことへの反発は誤り

UK Research and Innovation

Ref No: PROC/Form/04
V1.2

Specification for

UKRI 3341 -REF 2028 Environmental indicators commission

Annex A: Specification Document – Services / Supplies & Services

環境指標委員会の仕様書

RoRI RESEARCH ON RESEARCH INSTITUTE

People, Culture and Environment (PCE) 指標として考えられるもの

Possible PCE indicators?

PCE指標の候補となる項目(「良いデータ」)のバスケット

(The Metric TideおよびHarnessing the Metric Tide より抜粋):

研究文化/スタッフの幸福度調査の結果

研究リーダーシップコースへの参加(キャリアステージ/性別別

スタッフの民族的多様性

REF期間中の研究休暇取得スタッフの割合

ジャーナル編集者/編集委員会メンバーとしての役割

助成金/出版物のピアレビュー活動(FTE/年あたり

他機関(英国および海外)との共著論文の割合

他機関との共同プロジェクトによる助成金収入

REF期間内に昇進した教職員の割合

任期付き契約の研究者および技術者の割合

研究者における男女間および人種間の賃金格差

オープンアクセスによる成果

オープンデータセット

A basket of potential PCE indicators ('data for good')

(drawn from *The Metric Tide* and *Harnessing the Metric Tide*):

Research culture/staff well-being survey results

Participation in research leadership courses (by career stage/gender)

Ethnic diversity of staff

Proportion of staff with research leave over REF period

Roles as journal editors/editorial board members

Grant/publication peer review activity (per FTE/year)

Proportion of outputs co-authored with other institutions (UK and international)

Grant income from collaborative projects with other institutions

Proportion of staff promoted within the REF period

Proportion of researchers and technicians on fixed-term contracts

Gender and ethnicity pay gap amongst researchers

Open access outputs

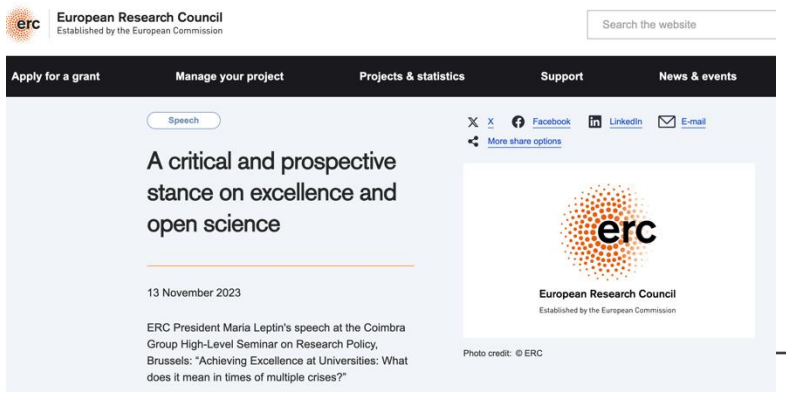
Open datasets

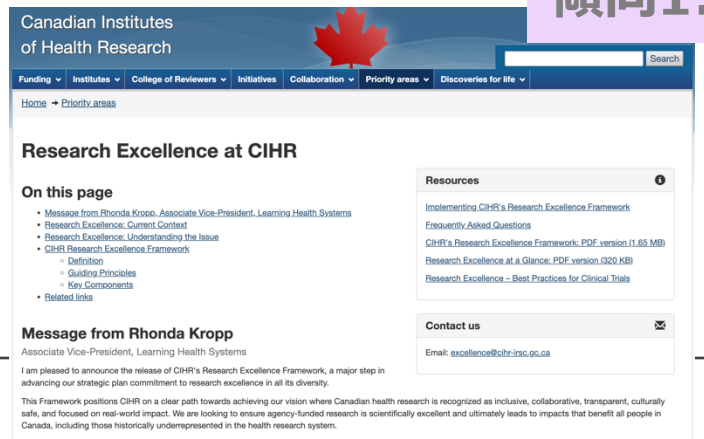
傾向1:　科学的な卓越性の概念の拡大

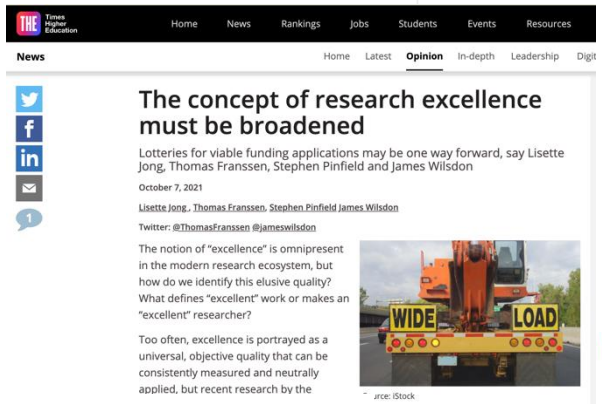
Trend 1: An expansion of notions of scientific excellence

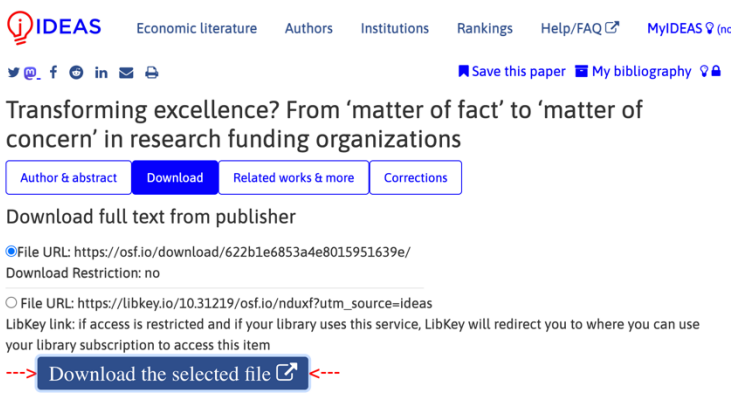
This is already well underway and is altering the demands and expectations placed on research, researchers and research funding – through a heightened focus on impact, TDR, team science, research culture, EDI etc.

これはすでに本格的に進行しており、インパクト、超学際研究、チームサイエンス、研究文化、公平性・多様性・包摂性(EDI)などへの注目が高まることで、研究、研究者、研究資金への需要と期待を変化させている









傾向2:学術コミュニケーションとオープンリサーチの急速に変化する状況

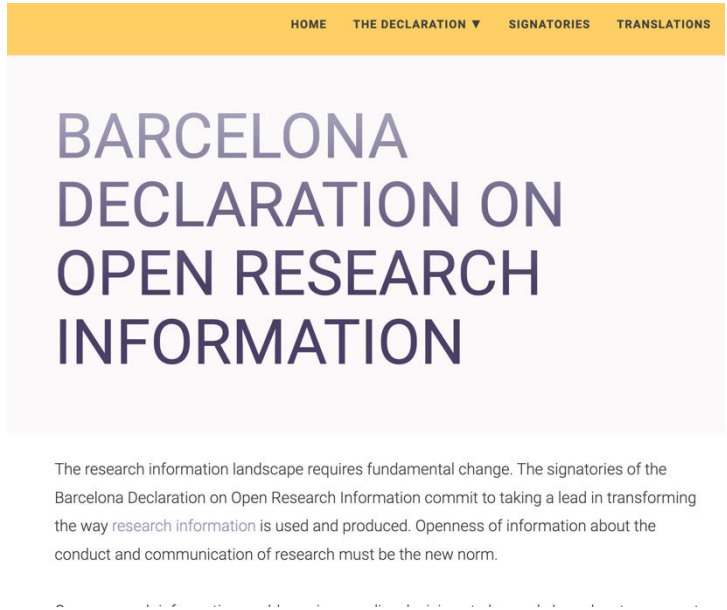
Trend 2: A rapidly-changing landscape for scholarly communication and open research

From Plan S to 'Plan U' and beyond, growth of preprints; the Publish-Review-Curate model; and the drive for open research data.

All of which means that traditional bibliometric measures (esp in more tightly-controlled commercial database) are increasingly inadequate to the tasks they purport to perform.

Plan Sから「Plan U」そしてそれ以降、プレプリントの増加、Publish-Review-Curateモデル、オープンな研究データの推進。

これらすべてが意味するのは、**従来の計量書誌学的手法**(特に、より厳重に管理された商業データベース)**は、本来の目的に対してますます不適切になっている**ということです。



RoRI RESEARCH ON RESEARCH INSTITUTE

傾向3:　測定・評価における新技術・手法の可能性

A personal take on science and society

World view

AI tools can help universities maximize research impacts

By Dashun Wang

Algorithms could identify scientists who need support with translating their work into real-world applications. Leaders must step up.

From the Internet to CRISPR–Cas9 gene editing, many seeds of progress were planted initially in the ivory tower of academia. Could research be doing even more for society? I argue that it could — if universities used artificial intelligence (AI) tools to maximize the impact of their scientists' outputs.

Each year, millions of grant proposals, preprints and research papers are produced, along with patents, clinical trials and drug approvals. Massive data sets storing details of these outputs can be scoured by AI algorithms to better understand how science and technology progress and to identify gaps and bottlenecks that hinder breakthroughs. Over the past few years, my colleague and close collaborator Ben Jones, my team and I have been working with large US universities to maximize their research impacts. We've already learnt a lot.

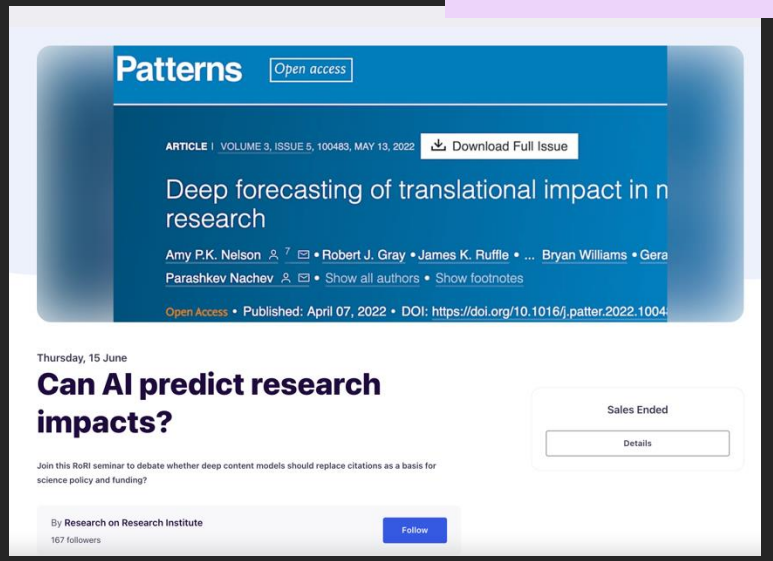
For example, during our pilot project at Northwestern University in Evanston, Illinois, we worked with one of its

> "The dichotomy of basic versus applied research is becoming inadequate."

Numerous factors can contribute to this gender gap, such as unequal access to education and mentorship, funding disparities, prevailing norms and stereotypes and structural barriers in patenting and commercialization processes. A better understanding of these challenges would help to broaden the pool of innovators.

Similarly, we see a large difference between tenure-track and tenured faculty members: tenured researchers patent their work at a higher rate. But one doesn't magically become more innovative the moment tenure is granted. The causes of this gap are probably distinct from those of the gender one, and might include promotion incentives and what counts towards tenure. But both discrepancies point to untapped opportunities for innovation.

Thus, data and AI tools can help institutions to identify people and ideas that are overlooked, both in a research institution and globally. But universities must take care. They have many roles and responsibilities — from educating future leaders to advancing fundamental knowledge — that must not be eclipsed by efforts to promote practical applications. Some people might argue that scientists don't need to commercialize their ideas themselves, because industry can pick up the ball. Or there might be unintended conse-



Patterns Open access

ARTICLE | VOLUME 3, ISSUE 5, 100483, MAY 13, 2022 Download Full Issue

Deep forecasting of translational impact in n research

Amy P.K. Nelson • Robert J. Gray • James K. Ruffle • ... Bryan Williams • Gera
Parashkev Nachev • Show all authors • Show footnotes

Open Access • Published: April 07, 2022 • DOI: https://doi.org/10.1016/j.patter.2022.1004

Thursday, 15 June

Can AI predict research impacts?

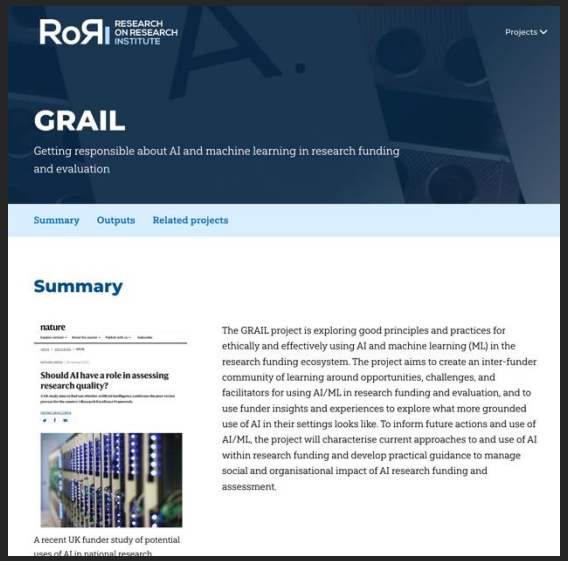
Join this RoRI seminar to debate whether deep content models should replace citations as a basis for science policy and funding?

By Research on Research Institute
167 followers

Follow

Sales Ended

Details



RoRI RESEARCH ON RESEARCH INSTITUTE

Projects

GRAIL

Getting responsible about AI and machine learning in research funding and evaluation

Summary Outputs Related projects

Summary

nature

Should AI have a role in assessing research quality?

A recent UK funder study of potential uses of AI in national research

The GRAIL project is exploring good principles and practices for ethically and effectively using AI and machine learning (ML) in the research funding ecosystem. The project aims to create an inter-funder community of learning around opportunities, challenges, and facilitators for using AI/ML in research funding and evaluation, and to use funder insights and experiences to explore what more grounded use of AI in their settings looks like. To inform future actions and use of AI/ML, the project will characterise current approaches to and use of AI within research funding and develop practical guidance to manage social and organisational impact of AI research funding and assessment.

Trend 3: Possibilities of new technologies and methods for measurement & evaluation

Possibilities of new technologies and methods – esp. LLMs.

These need to be deployed carefully and with thought but potentially they break down some quantitative/qualitative divides, and combined with other methods, point towards richer proxies, reduced burdens, and real-time systems for monitoring research performance.

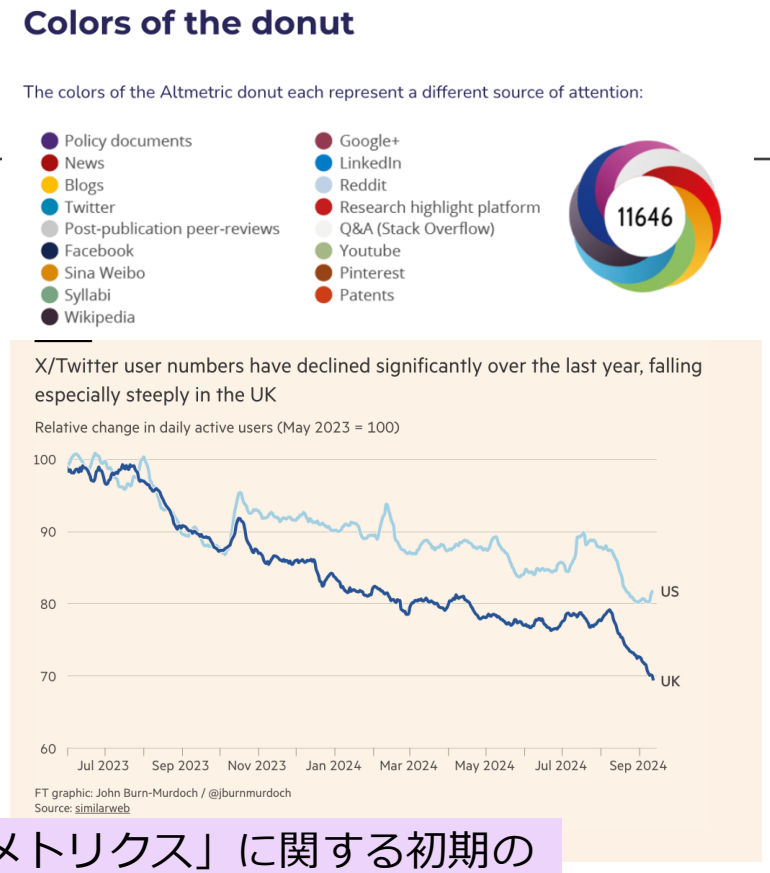
新しいテクノロジーや手法の可能性 – 特に大規模言語モデル（LLM）

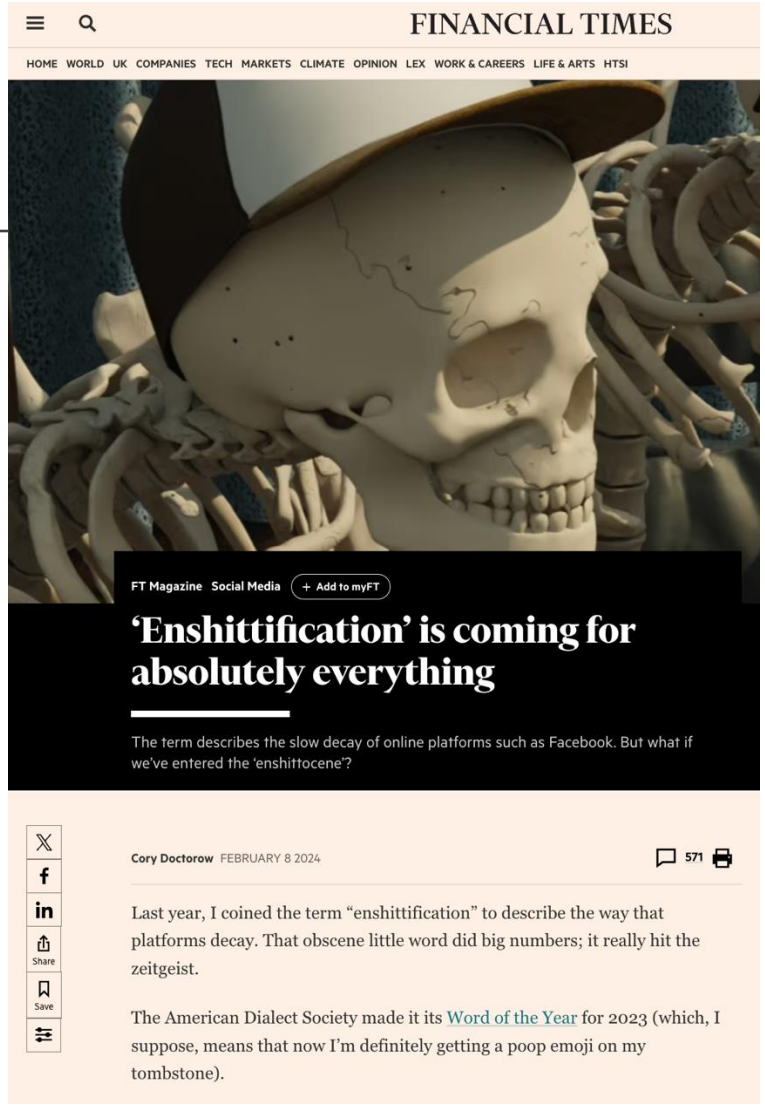
これらは慎重に、かつ熟考した上で展開する必要がありますが、定量的・定性的な格差を解消する可能性があり、他の手法と組み合わせることで、より豊かな代替指標、負担の軽減、研究業績のモニタリングのためのリアルタイムシステムにつながるでしょう。

傾向4:　第一世代の「オルトメトリクス」を独占していたプラットフォームの浸食

Trend 4: Erosion of platforms that dominated first generation ‘altmetrics’

There’s an accelerating erosion of social media platforms that dominated early debates over “altmetrics”. For example, how do we now treat X/Twitter as a measure worth focusing on (versus Mastodon, Bluesky, TikTok, Instagram etc). It’s messy and fluid – and shows how one needs to avoid “baking in” particular indicators or systems that may not last and are weak proxies at best for ‘impact’.



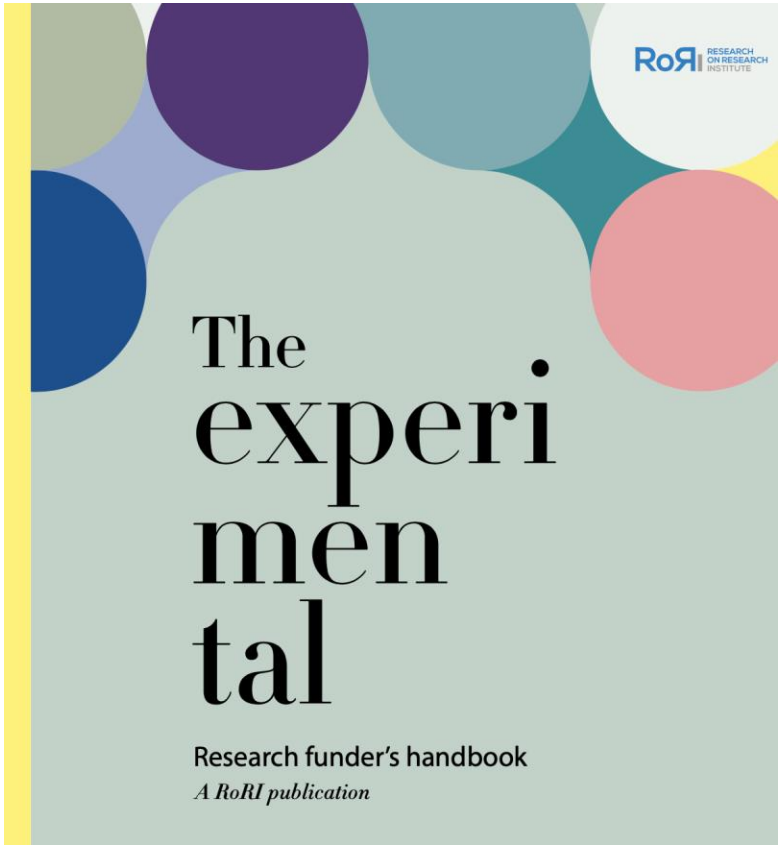


新しい指標として注目されていた「オルトメトリクス」に関する初期の議論を独占していたソーシャルメディアプラットフォームの浸食が加速。例えば、X/Twitterを重視すべき指標として扱うべきかどうか(Mastodon、Bluesky、TikTok、Instagramなどより新しい研究紹介ツールと比較して)。それは混沌として流動的であり、「影響」を測る指標やシステムとして長続きしない可能性があり、せいぜい弱い代替指標にすぎないものを「焼き付ける」ことは避けなければならない。

RoRI RESEARCH ON RESEARCH INSTITUTE

傾向5：　大学と資金提供者が評価へのアプローチをより系統的かつ実験的に

Trend 5. Universities and funders becoming more systematic & experimental in their approaches to assessment



AGORRA(責任ある研究評価のための世界的な監視機関)

RoRI RESEARCH ON RESEARCH INSTITUTE

Projects Resources Events About Get Involved

AGORRA

A global observatory of responsible research assessment

Summary Outputs Related projects

Summary

Less than a year after **it started rolling**, the **CoARA (Coalition for Advancing Research Assessment)** convoy is gathering speed. The tally of organisational signatories to its underpinning **Agreement on Reforming Research Assessment** is rising by dozens each week.

In May 2023, **DORA (The Declaration on Research Assessment)**—on whose foundations CoARA builds—marked its 10th anniversary with a series of workshops around the world. And at a national level, we've seen a sharper focus on these agendas in light of ongoing or proposed

Research Team

Claire Fraser
Senior Policy Adviser, Research England

Alex Rushforth
Senior Research Fellow, RoRI and CWTS-Leiden

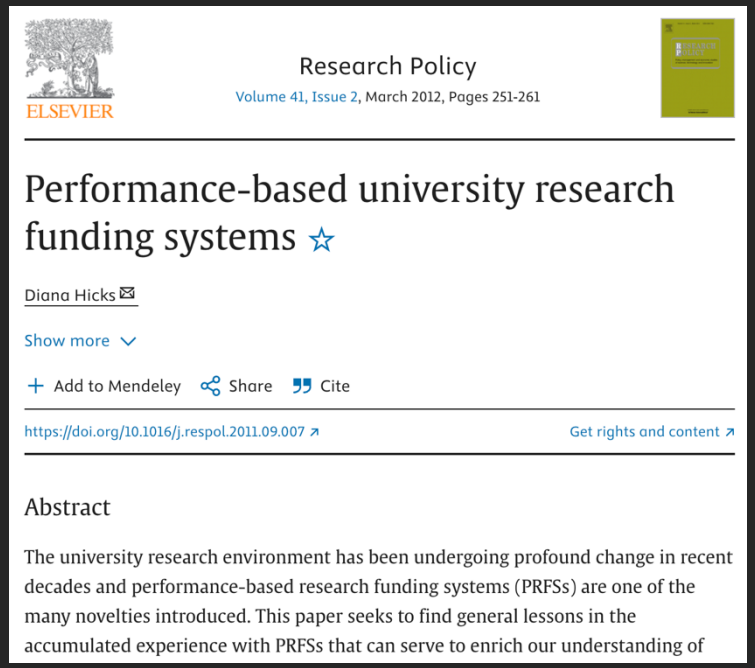
James Wilsdon
Executive Director, RoRI & Director, RoRI CIC

全国的な研究評価システム: 新しい類型論

なぜ類型論なのか?

National research assessment systems: a new typology

Why a typology?



ELSEVIER

Research Policy

Volume 41, Issue 2, March 2012, Pages 251-261

Performance-based university research funding systems ☆

Diana Hicks ✉

Show more

+ Add to Mendeley Share Cite

https://doi.org/10.1016/j.respol.2011.09.007 ↗ Get rights and content ↗

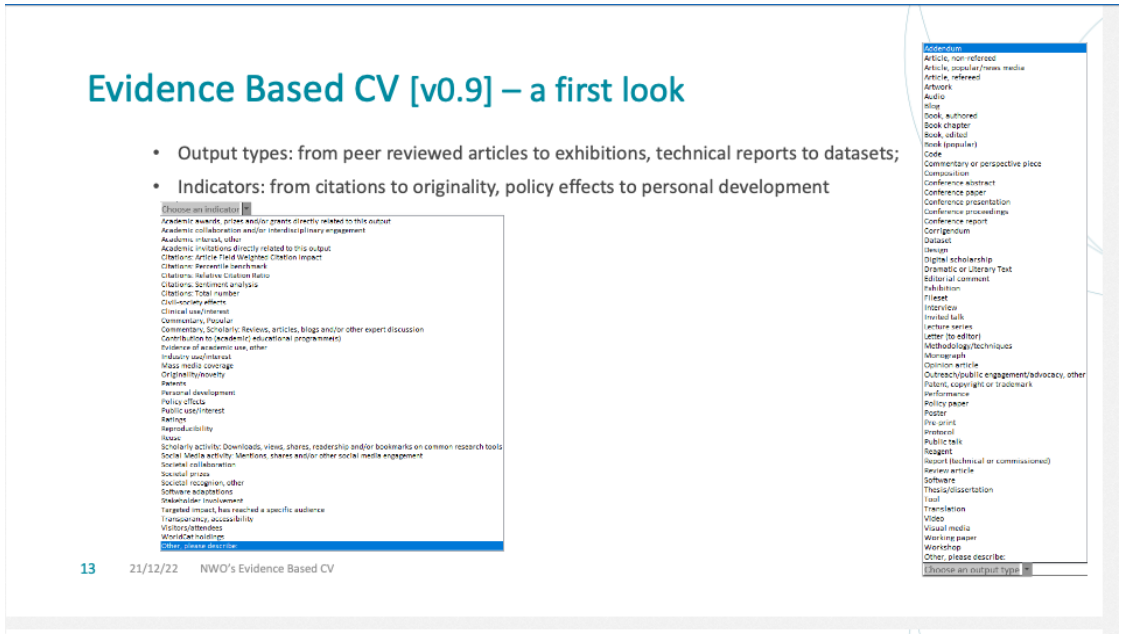
Abstract

The university research environment has been undergoing profound change in recent decades and performance-based research funding systems (PRFSs) are one of the many novelties introduced. This paper seeks to find general lessons in the accumulated experience with PRFSs that can serve to enrich our understanding of

To help identify key characteristic properties of a given national assessment system	特定の国家評価システムの主な特性を特定するのに役立つ
To support comparison between major types of system	主要なシステムタイプ間の比較をサポートする
To support mutual learning and exchange within the AGORRA projects among partners	パートナー間のAGORRAプロジェクトにおける相互学習と交流をサポートする
Data collected in the workstream on 13 countries also very useful:	13カ国を対象としたワークストリームで収集されたデータも非常に有用
To explore major developments and patterns in national systems over time (2010-2024)	国家システムにおける主な進展とパターンを長期的に(2010年から2024年まで)調査する
To reflect on how current designs, purposes and rationales might shape 'responsible research assessment' transitions differently.	現在の設計、目的、根拠がどのように「責任ある研究評価」への移行を形づくるかを考察する

日本のシステムに対する質問: (1) 研究評価の改革や他国で新たに登場した好事例から、どのような洞察が得られるだろうか?

その他の新しいアイデアとしては、CWTSが開発した「つながりに基づく地域性」という手法があり、つながりによって潜在的な社会的重要性を示すことができる。



This is the NWO's model for evidence based CVs

Other newer ideas include the *'area based connectedness' method* developed by CWTS to indicate potential societal relevance by connectedness

DE GRUYTER OPEN

Research Paper

Measuring Societal Impact Is as Complex as ABC

Ed Noyons†

Centre for Science and Technology Studies, Leiden University, Leiden, The Netherlands

Citation: Ed Noyons (2019). Measuring societal impact is as complex as ABC. Journal of Data and Information Science, 4(3), 6–21

DOI: 10.2478/jdis-2019-0012

Received: June 28, 2019
Revised: July 30, 2019
Accepted: Aug. 6, 2019

Abstract

Purpose: This paper describes an alternative way of assessing journals considering a broader perspective of its impact. The Area-based connectedness (ABC) to society of journals applied here contributes to the assessment of the dissemination task of journals but with more data it may also contribute to the assessment of other missions.

Design/methodology/approach: The ABC approach assesses the performance of research actors, in this case journals, considering the characteristics of the research areas in which they are active. Each paper in a journal inherits the characteristics of its area. These areas are defined by a publication-based classification. The characteristics of areas relate to 5 dimensions of connectedness to society (news, policy, industrial R&D, technology and local interest) and are calculated by bibliometric indicators and social media metrics.

Findings: In the paper, I illustrate the approach by showing the results for a few journals. They illustrate the diverse profiles that journals may have. We are able to provide a profile for each journal in the Web of Science database. The profiles we present show an appropriate view on the journals' societal connectedness.

Research limitations: The classification I apply to perform the analyses is a CWTS in house classification based on Web of Science data. As such the application depends on the (updates of) that system. The classification is available at www.leidenranking.com.

Practical implications: The dimensions of connectedness discussed in this paper relate to the dissemination task of journals but further development of this method may provide more options to monitor the tasks/mission of journals.

Questions for the Japanese system: (1) what insights might be drawn from research assessment reforms and emerging good practices elsewhere?

New Zealand: rethinking the PBRF

THE Times Higher Education — Home News Rankings Jobs Students Events Resources

News — Home Latest Opinion In-depth Leadership Digital e

New Zealand cancels research quality evaluation

New government's decision spells the end of 'back-breaking' exercise, commentators say

April 9, 2024

John Ross

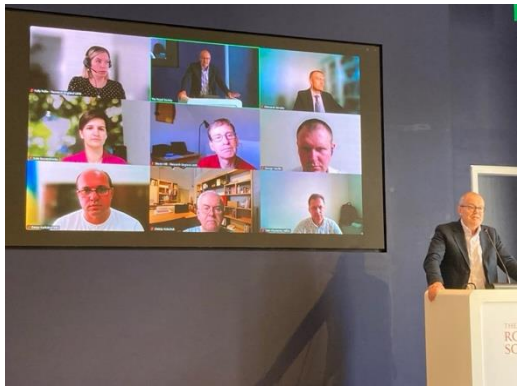
Twitter: @JohnRoss49

New Zealand has cancelled the upcoming round of its national research assessment, in what is widely regarded as a death blow for the decades-old exercise.

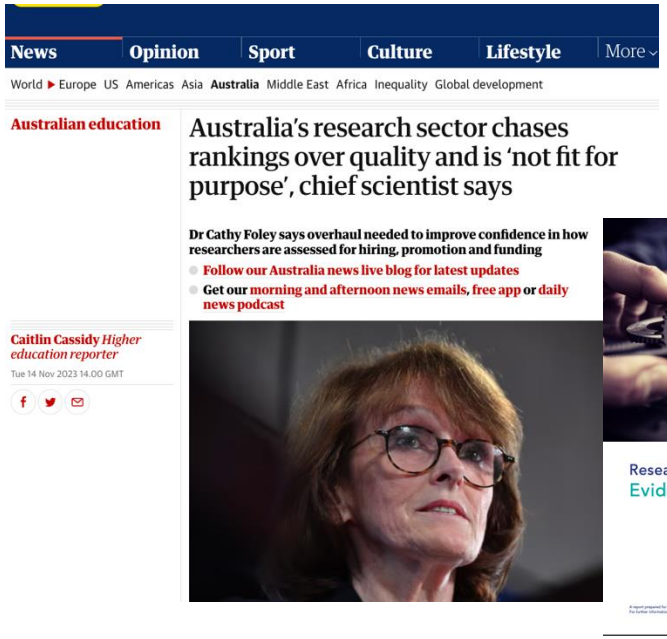
Tertiary education minister Penny Simmonds has decided not to proceed with data collection for the Performance-Based Research Fund (PBRF), which guides the allocation of NZ$315 million (£150 million) of block grants each year.

Source: THE

Ukraine: newcomers to national assessment



Australia: end of an ERA & what next?



News Opinion Sport Culture Lifestyle More

World Europe US Americas Asia Australia Middle East Africa Inequality Global development

Australian education

Australia's research sector chases rankings over quality and is 'not fit for purpose', chief scientist says

Dr Cathy Foley says overhaul needed to improve confidence in how researchers are assessed for hiring, promotion and funding

Follow our Australia news live blog for latest updates

Get our morning and afternoon news emails, free app or daily news podcast

Caitlin Cassidy *Higher education reporter*

Tue 14 Nov 2023 14.00 GMT

China: curbing the "siwei" ("four-onlys")

Citation: LI Xiaoxuan, XU Fang. How to Break the "Siwei"?—Practice and Enlightenment Based on Research Institute Evaluation of Chinese Academy of Sciences [J]. Bulletin of Chinese Academy of Sciences, 2020 (12): 1431–1438.

How to Break the "Siwei"?—Practice and Enlightenment Based on Research Institute Evaluation of Chinese Academy of Sciences

LI Xiaoxuan[1,2], XU Fang[1,2]

1. Institutes of Science and Development, Chinese Academy of Sciences, Beijing 100190, China;
2. School of Public Policy and Management, University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: In October 2018, five ministries and institutions, i.e., Ministry of Science and Technology, Ministry of Education, Ministry of Human Resources and Social Security, Chinese Academy of Sciences (CAS), and Chinese Academy of Engineering, collaboratively started the special action of breaking "Siwei," which means "Four-Only" problems, i.e., only papers, only titles, only education background, and only awards. Most researchers in universities and research institutions have both expectations and concerns. There are different opinions on how to break the "Siwei." On the basis of the analysis of the development of evaluation conducted by CAS for more than 20 years, this study holds the view that CAS has explored a way of breaking the "Siwei" and formed the CAS mode in research institute evaluation, which is expected to provide a case for reference on how to break the "Siwei."

DOI: 10.16418/j.issn.1000-3045.20201116002-en

Keywords: break the "Siwei"; Chinese Academy of Sciences; research institute evaluation; CAS mode; science evaluation

Over the years, science evaluation, particularly in the basic research, has been a hot topic in the scientific and techno- indicators sh the breaking

Spain: changing the balance of assessment?

AGENCIA ESTATAL DE INVESTIGACIÓN

ABOUT AEI · ANNOUNCEMENTS · THEMATIC AREAS · EVALUATION · AWARDED GRANTS · INTERNATIONAL COLLABORATION

La Agencia Estatal de Investigación y la Fundación "la Caixa" presentan claves para nuevos métodos de evaluación científica

March 18, 2024

- En un evento celebrado en la AEI, expertos de organismos como el Consejo Europeo de Investigación (ERC), RoRI (Research on Research Institute), la Fundación "la Caixa" y la AEI compartieron y debatieron acerca de cómo el uso de inteligencia artificial puede contribuir a mejorar los sistemas de evaluación científica.

世界各国における研究評価の変化

RoRI RESEARCH ON RESEARCH INSTITUTE

日本のシステムへの質問: (2) 科学的卓越性、画期的な成果、質、影響力をどのように定義するか?

NOR-CAM –
A toolbox for recognition and rewards

I In recent years, initiatives to further develop ways of assessing research quality and academic careers have increased in scope and strength. This guide proposes a more flexible and holistic framework for recognition and rewards that adopts core principles for assessment: more transparency, greater breadth, and comprehensive assessments as opposed to one-sided use of indicators.*

Open Science
There are strong indications that, in the time ahead, openness will be an integral part of all knowledge production and dissemination. Open research is about to become the new norm, and it will therefore be natural for all results, activities and competencies to be assessed in the light of the aims of open research. Open research will therefore affect how careers are assessed.

Assessing and recognising a greater breadth of competencies in research and teaching and interaction with society. Many of the activities that academic staff perform in line with the institution's goals and work are not systematically assessed or valued. Individual research achievements in the form of published research results are more strongly incentivised than other work, and individual achievements are given greater weight than collaborative efforts. A better balance in the assessment of the various key activities has been called for.

The need to reduce and modify the reliance on quantitative publication metrics in academic career assessment. The privileging of quantitative research results and the traditional, quantifiable indicators with which they

academic activities, and ultimately to improve academic culture and the quality of research.

The knowledge sector is global. Changing the framework for recognition and rewards in one country at a time would be difficult. Researcher mobility and international funding makes it challenging to implement practices that are at odds with international norms. NOR-CAM is therefore developed in close contact with partners in several other countries, as well as in the European University Association (EUA) and the EU.

The guide is flexible but offers a systematic and structured framework for assessment. The assessment can be adapted to emphasise different competencies for different tasks/ positions/career stages depending on both the individual's career and the institutional needs.

An important goal of the guidance and framework is to make the assessment processes more transparent and predictable, both for the individual and for the institutions. What skills are needed for the position to be filled? How well does your own competence fit the position advertised? What are the requirements for promotion?

This will mean greater transparency about which assessment criteria are emphasised in the specific context, and will improve predictability, not least for applicants. It will also provide a better basis for career follow-up throughout the academic career path.

and supplemented with other information

2. **Recognise several competencies as merits but not in all areas at the same time or by each employee**
The individual academic is not expected to excel in all areas. It is the universities that must achieve the expected objectives given by the government regarding research, education and interaction with society, not the individual academic.
3. **Assess all results, activities and competencies in the light of Open Science principles**
Openness should be seen as an integrated part of the academic activity
4. **Practice transparency in the assessment and visibility of what should be recognised as merit**
Individuals must know what criteria are used to assess them and must be given insight into how the criteria are applied
5. **Promote gender balance and diversity**
Changes in the assessment criteria must be sensitive to impact on gender balance and diversity
6. **Assist in the concrete practice of announcements and assessment processes locally**
The framework should be a helpful tool in recruitment and appraisal processes in the institutions and within the academic communities

Four recommendations

1. **To establish a comprehensive framework for the assessment of academic careers that:**
balances quantitative and qualitative goals and forms of documentation for academic standards and competencies

enables diverse career paths and promotes high standards in the three key areas: education, research and interaction with society

recognises the independent and individual compe-

3. **To use NOR-CAM as a practical and flexible tool** for assessing academic results, competence and experience for academic personnel. NOR-CAM will highlight six areas of expertise through systematic documentation and reflection: *See the matrix on the next page.*

and project participants' competencies when assessing research projects.
Authorities
Ministry of Education and Research: incorporate the principles of NOR-CAM into the new national framework for the evaluation of Norwegian research and higher education.

Questions for the Japanese system: (2) How to define scientific excellence, breakthroughs, qualities & impacts?

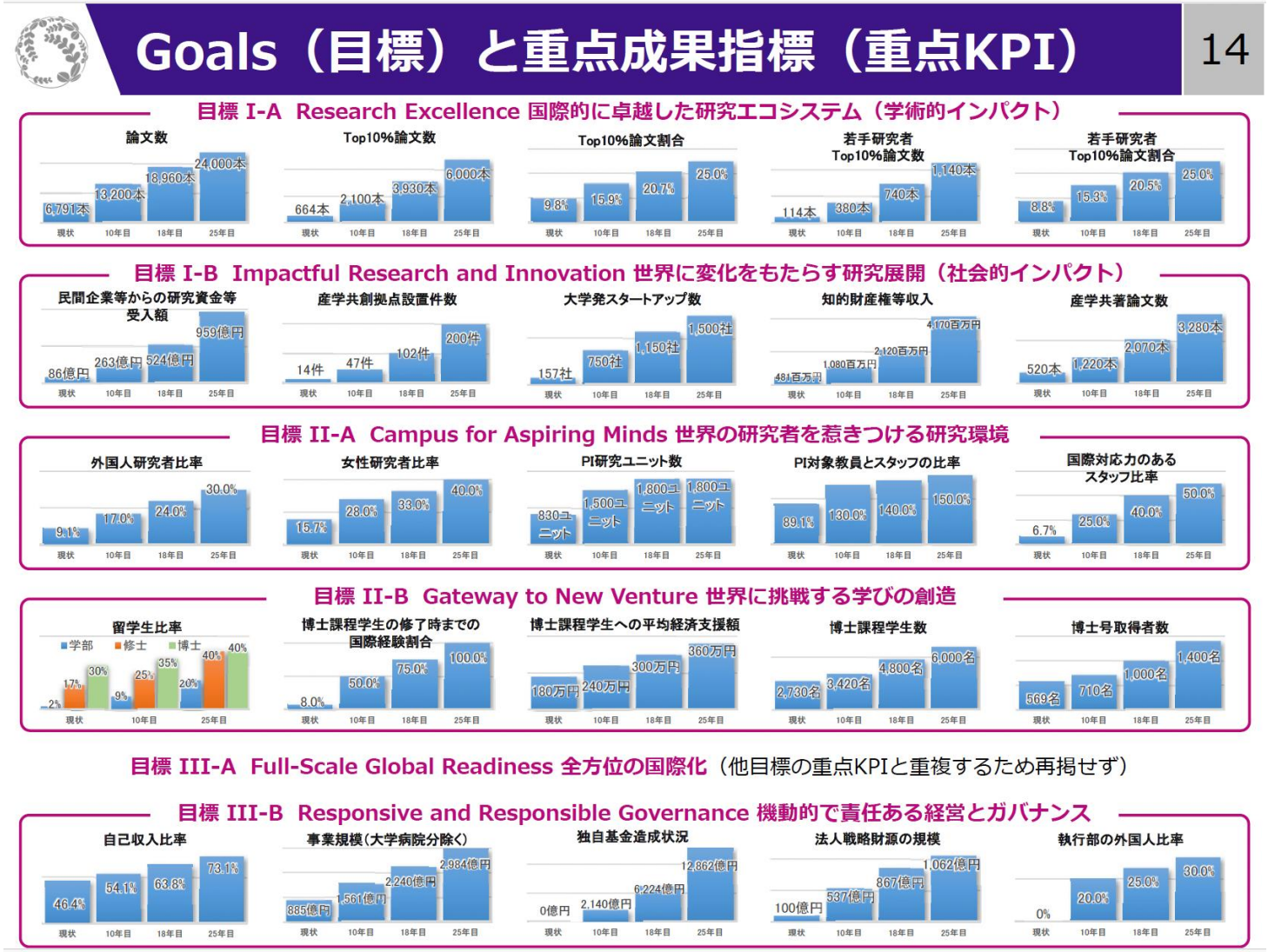
日本のシステムに対する質問: (3)

責任ある効果的な研究指標、管理体制、文化をどのように設計、選択、実施するか?

Questions for the Japanese system: (3)

how to design, select and implement responsible and effective research indicators, management systems and cultures?

責任ある効果的な研究指標、管理システム、文化をどのように設計、選択、実施するか？



Canadian Institutes of Health Research (CIHR)の研究の卓越性を目指す新たな枠組み

CIHR's new framework for research excellence

Guiding Principles

The following foundational principles will guide CIHR's work in integrating a more inclusive approach to research excellence across its programs, policies and practices:

Holistic: Research excellence is broad and spans how research is conceptualized, prioritized, taught, carried out, assessed, funded and used.

Adaptable: Research excellence is not one-size-fits-all, but rather context- and content-specific. Flexible, catered approaches are required to recognize and incent the breadth of research within CIHR's mandate.

Evergreen: Research excellence is a concept designed to be adjusted as new evidence emerges, and as science and society evolve.

Key Components

The following key components should be considered and addressed within all CIHR-funded research whenever relevant and appropriate:

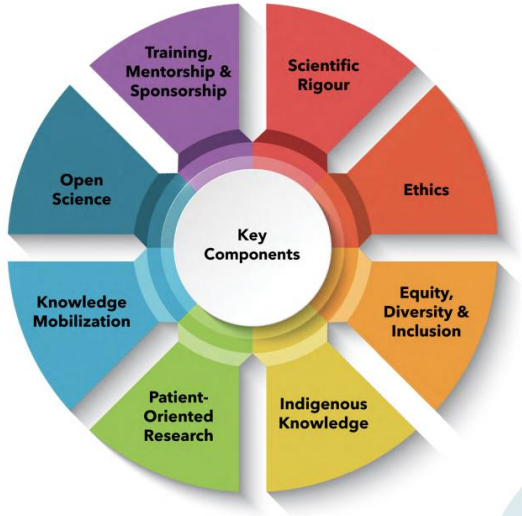


Figure 2: CIHR's key components of research excellence

GRASPOSプロジェクト (欧州のOpen Research Assessment Dataspace)は新しい指標フレームワークを開発しており、「研究評価改革の推進のための有志連合（Coalition for Advancing Research Assessment, COARA）」を通じて多くの作業が進行中

The GRASPOS project has developed new indicator frameworks and there is much work ongoing through COaRA

RoRI RESEARCH ON RESEARCH INSTITUTE

Japan

index

Will Japan's new ¥10-trillion university fund lift research performance?

Endowment inspired by US Ivy League model must be accompanied by other reforms, critics say. **By David McNeill**

Large endowment funds underpin the stability and independence of some of the world's most successful research institutions. Take Harvard University, the leading academic institution globally by Share in the Nature Index. Harvard has one of the world's largest endowments for an academic institution at more than $50 billion, a sum that allowed the fund to contribute $2.1 billion to the institution's operating budget from 2021 to 2022.

These endowments are usually made up of donations from private individuals; in the case of Harvard, the overall endowment includes thousands of separate charitable funds. Proponents say such endowments can give institutions a degree of freedom to pursue the scientific research they wish without the political interference that may come from relying entirely on government funding.

Inspired by this model, Japan unveiled plans in 2021 for a ¥10-trillion (US$75-billion) national endowment fund. The aim, say its architects, is to spur innovation and improve

and Germany spend twice that). Government funding for national universities, where most of Japan's leading science and technology research is carried out, has been cut by about 1% per year since they were freed from direct government control and partially privatized as part of sweeping reforms in 2004. This amounts to a total fall from ¥1.24 trillion in 2004 to ¥1.08 trillion in 2022.

Modest returns

MEXT now wants to nudge Japanese universities towards the US endowment funding model, forcing them to be more self-sustaining. Harvard's fund delivered a 33.6% return in the year to June 2021 (although such returns can easily fluctuate: it had a negative return of 1.8% last year). By contrast, the endowment fund of one of Japan's top higher-education institutions, the University of Tokyo, was around ¥19 billion in the year to March 2021, less than 0.5% of Harvard's total fund. Many other prestigious universities in Japan earn little or nothing from investments

between four and six, according to Ueyama, would share the profits from the estimated ¥300 billion a year earned from the fund. MEXT criteria for institutions wanting to take part include having "research capabilities for international excellence" and an "autonomous and responsible governance structure".

Ueyama says the estimated annual return is modest, but critics argue that the government forecasts are too optimistic. Japan's stock market doesn't function based on normal market conditions because the government has been artificially inflating its pension investment fund by buying stocks from the Bank of Japan, explains Takeshi Komagome, an education policy expert at the University of Kyoto. Without similar cash injections in the future, it's difficult to imagine big investment returns, he says. Komagome says the new initiative is effectively a confidence trick. "It looks good on the surface, but it won't deliver any extra money."

Then there is the issue of who will get funding. MEXT estimates that only a dozen or so universities will make the criteria, which

Finally, a few reflections from an RRA perspective about the role of indicators in the new University Fund:

how to strike the right balance between summative & formative modes of assessment?

How to balance universal applicability vs disciplinary & institutional diversity?

How to create long-term targets and incentives against a backdrop of continual change?

最後に、新しい大学基金における指標の役割について、RRAの観点からいくつかの考察を提示したい

総括的評価と形成的評価の適切なバランスをどのように取るか？

普遍的な適用可能性と学問分野や機関の多様性とのバランスをどのように取るか？

絶え間ない変化を背景に、長期的な目標とインセンティブをどのように設定するか？

RoRI RESEARCH ON RESEARCH INSTITUTE

j.wilsdon@ucl.ac.uk

researchonresearch.org