

DORA and reform of researcher evaluation



Stephen Curry & Anna Hatch

ACT Matching Event Careers | 29 Oct 2020



DORA: the declaration

One general recommendation:

Do not use journal-based metrics, such as Journal Impact Factors, as a **surrogate measure** of the quality of individual research articles, to assess an individual scientist's contributions, or in hiring, promotion, or funding decisions.



17 **positive** recommendations for different stakeholders:

- funders
- **institutions**
- publishers
- data providers
- researchers

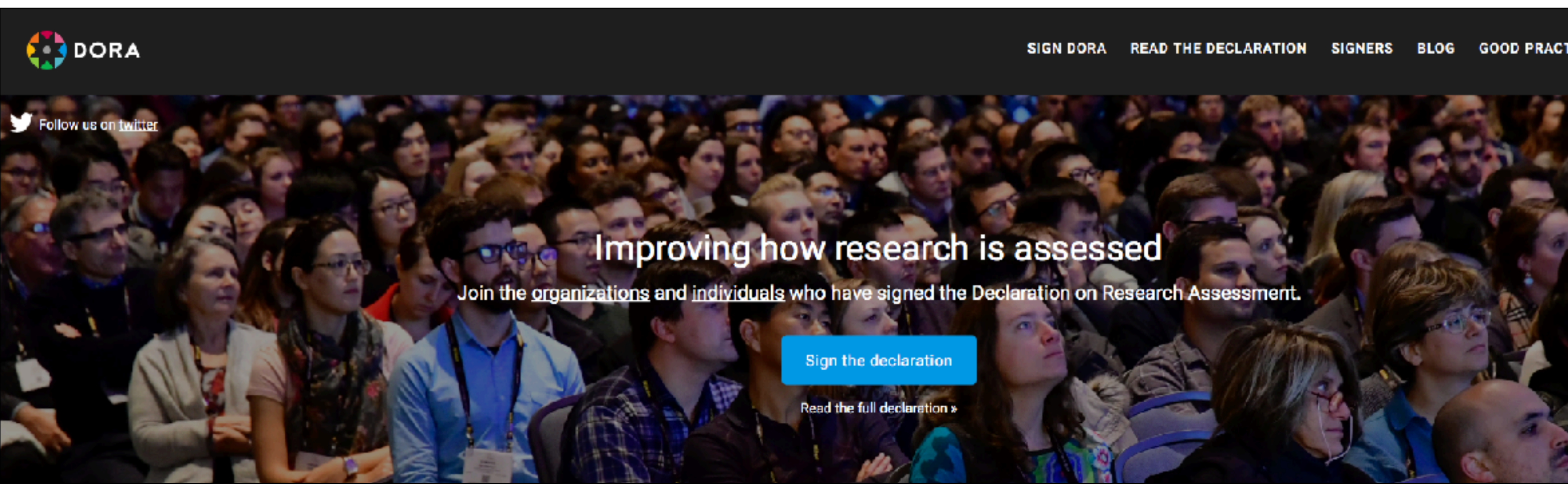
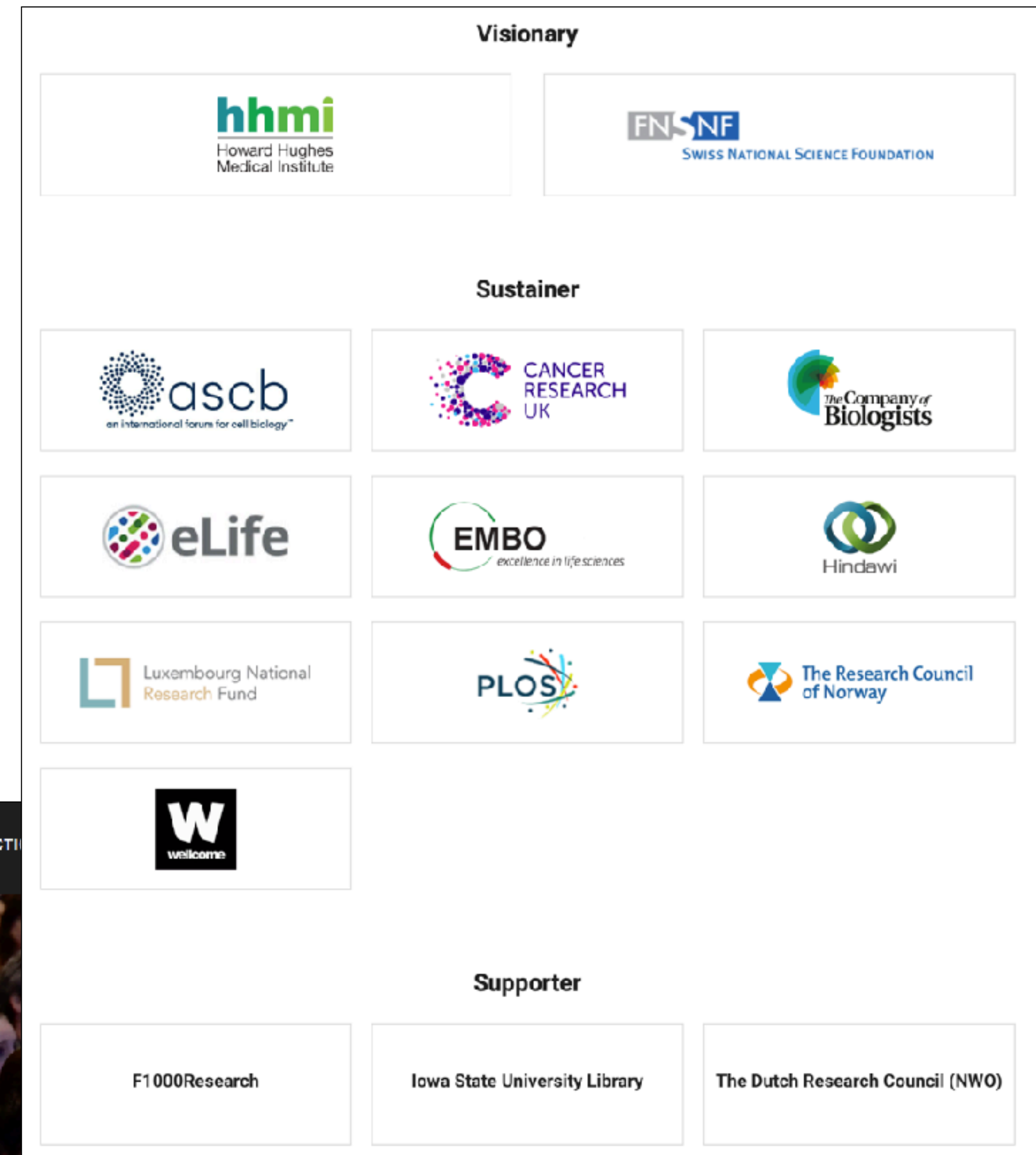
For institutions:

4. **Be explicit about the criteria** used to reach hiring, tenure, and promotion decisions, clearly highlighting, especially for early-stage investigators, that the **scientific content of a paper is much more important than publication metrics** or the identity of the journal in which it was published.

5. For the purposes of research assessment, **consider the value and impact of all research outputs** (including datasets and software) in addition to research publications, and **consider a broad range of impact measures** including qualitative indicators of research impact, such as influence on policy and practice.

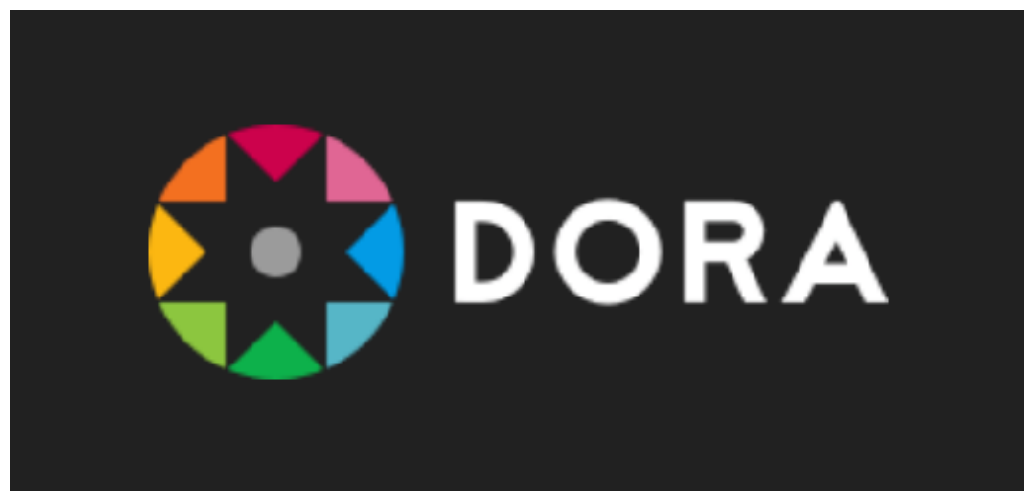
DORA: the organisation

- sfdora.org
- Launched in 2013
- >16,500 individuals and >2,000 organisations have signed
- From 2017: significant new financial support
 - 1.2 members of staff (plus an intern)
 - International steering group & a global advisory board (all volunteers)



DORA: our roadmap for action

- Recruit more signatories
- Extend the global and disciplinary impact of DORA
- Develop and promote best practice in research assessment



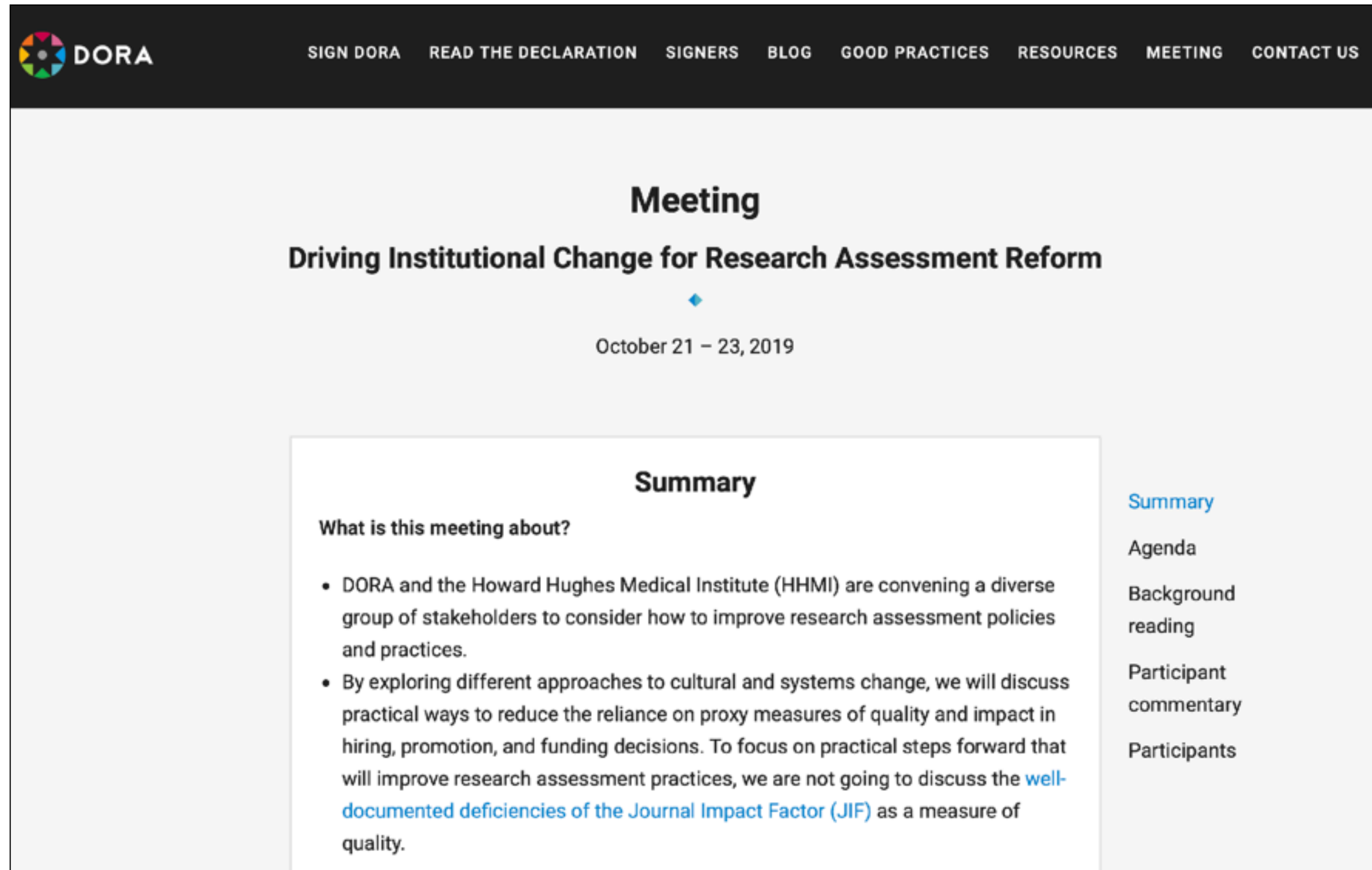
DORA session at ASCB|EMBO (Dec 2018)

The screenshot shows the eLife website interface. At the top, there's a navigation bar with the eLife logo, a hamburger menu, and links for 'HOME', 'MAGAZINE', and 'INNOVATION'. On the right, there are links for 'ABOUT', 'COMMUNITY', 'SUBMIT MY RESEARCH', and 'LOG IN/REGISTER'. The main content area features the article title 'Research Assessment: Reducing bias in the evaluation of researchers' in a large, bold font. Below the title is a subtitle: 'A workshop run by DORA identified a number of ways to reduce bias in hiring and funding decisions.' There are social media sharing icons for Facebook, Twitter, Email, and Reddit. Below these, it says 'INSIDE ELIFE Apr 17, 2019'. Further down, it shows 'VIEWS 1,375' and 'ANNOTATIONS 0'. The authors are listed as 'By Anna Hatch (DORA), Veronique Kiermer (PLOS), Bernd Pulverer (EMBO), Erika Shugart (American Society for Cell Biology), and Stephen Curry (Imperial College London)'. The section 'Introduction' begins with the text: 'Hiring and funding decisions influence academic priorities directly by setting research agendas. They also shape priorities indirectly by affecting the diversity of the scientific workforce, which in turn influences the questions that'.

<https://elifesciences.org/inside-elif/1fd1018c/research-assessment-reducing-bias-in-the-evaluation-of-researchers>

DORA: developing and promoting best practice

DORA/HHMI joint meeting (Oct 2019)



The screenshot shows the DORA website with a dark header containing the DORA logo and navigation links: SIGN DORA, READ THE DECLARATION, SIGNERS, BLOG, GOOD PRACTICES, RESOURCES, MEETING, and CONTACT US. The main content area is titled "Meeting" with the subtitle "Driving Institutional Change for Research Assessment Reform" and the dates "October 21 – 23, 2019". Below this is a "Summary" section with the heading "What is this meeting about?" and a bulleted list of points. A sidebar on the right lists links for Summary, Agenda, Background reading, Participant commentary, and Participants.

Meeting
Driving Institutional Change for Research Assessment Reform
October 21 – 23, 2019

Summary

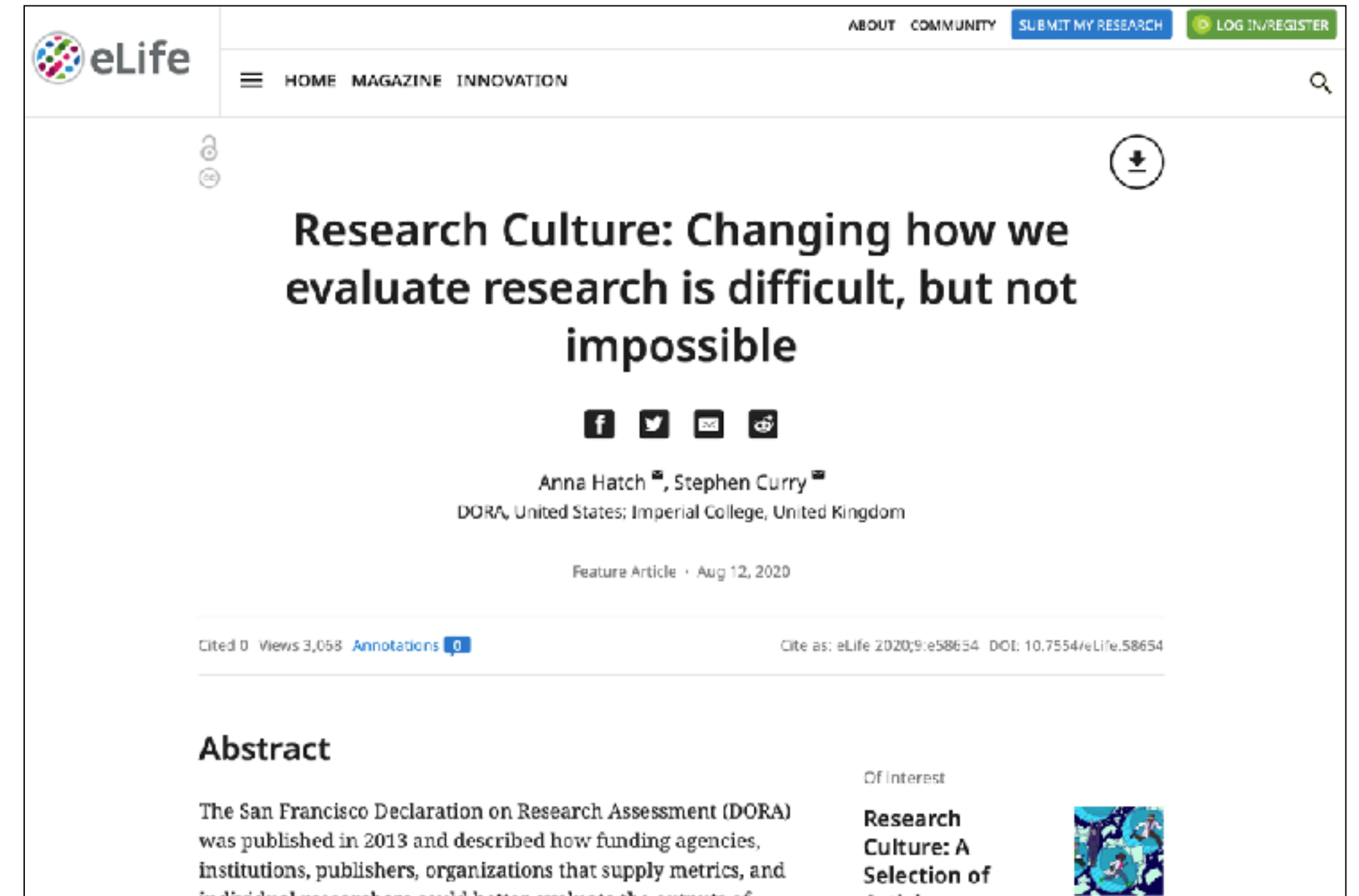
What is this meeting about?

- DORA and the Howard Hughes Medical Institute (HHMI) are convening a diverse group of stakeholders to consider how to improve research assessment policies and practices.
- By exploring different approaches to cultural and systems change, we will discuss practical ways to reduce the reliance on proxy measures of quality and impact in hiring, promotion, and funding decisions. To focus on practical steps forward that will improve research assessment practices, we are not going to discuss the well-documented deficiencies of the Journal Impact Factor (JIF) as a measure of quality.

Summary
Agenda
Background reading
Participant commentary
Participants

<https://sfdora.org/assessingresearch/>

More info, ideas & resources at: <https://sfdora.org/>



The screenshot shows the eLife article page. The header includes the eLife logo and navigation links: ABOUT, COMMUNITY, SUBMIT MY RESEARCH, and LOG IN/REGISTER. The article title is "Research Culture: Changing how we evaluate research is difficult, but not impossible". Below the title are social media icons and the authors' names: Anna Hatch and Stephen Curry. The article is dated August 12, 2020. The abstract is visible, starting with "The San Francisco Declaration on Research Assessment (DORA) was published in 2013 and described how funding agencies, institutions, publishers, organizations that supply metrics, and individual researchers could better evaluate the outputs of".

eLife
ABOUT COMMUNITY SUBMIT MY RESEARCH LOG IN/REGISTER

HOME MAGAZINE INNOVATION

Research Culture: Changing how we evaluate research is difficult, but not impossible

Anna Hatch, Stephen Curry
DORA, United States; Imperial College, United Kingdom

Feature Article • Aug 12, 2020

Cited 0 Views 3,058 Annotations 0 Cite as: eLife 2020;9:e58654 DOI: 10.7554/eLife.58654

Abstract

The San Francisco Declaration on Research Assessment (DORA) was published in 2013 and described how funding agencies, institutions, publishers, organizations that supply metrics, and individual researchers could better evaluate the outputs of

Of interest
Research Culture: A Selection of

<https://elifesciences.org/articles/58654>

Framework for action:

- **understand** the obstacles to changes in the way research is assessed
- **experiment** with different approaches
- **create a shared vision** when revising policies and practices
- **communicate that vision** on campus and beyond

DORA: we collaborate

Royal Society - Résumé for Researchers



THE ROYAL SOCIETY Venu

Home Fellows Events Grants, Schemes & Awards Topics & policy Journals

Résumé for Researchers

Opening up conversations about researcher evaluation

Résumé for Researchers has been created to support the evaluation of individuals' varied contributions to research. Find out more about the background to the tool [in our blog](#).

Module 1 – How have you contributed to the generation of knowledge?

This module can be used to explain how you have contributed to the generation of new ideas and hypotheses and which key skills you have used to develop ideas and test hypotheses. It can be used to highlight how you have communicated on your ideas and research results, both written and verbally, the funding you have won and any awards that you have received. It can include a small selection of outputs, with a description of why they are of particular relevance and why they are considered in the context of knowledge generation. Outputs can include open data sets, software, publications, commercial, entrepreneurial or industrial products, clinical practice developments, educational products, policy publications, evidence synthesis pieces and conference publications that you have generated. Where outputs have a DOI please only include this.

Module 2 - How have you contributed to the development of individuals?

Module 3 - How have you contributed to the wider research community?

Module 4 - How have you contributed to broader society?

<https://royalsociety.org/topics-policy/projects/research-culture/tools-for-support/resume-for-researchers/>

DORA: we are helping to create new tools and processes for evaluation

RETHINKING RESEARCH ASSESSMENT

IDEAS FOR ACTION



5 COMMON MYTHS ABOUT EVALUATION

Hiring, promotion, and tenure decisions are largely made on “merit.”

Quality research is easy to recognize and rises to the top

JIF and other similar journal-based indicators measure research quality

Researchers mostly care about journal reputation

Assessment practices will naturally improve over time

Assessing research and researchers, especially in research-intensive institutions, frequently relies on indicators like Journal Impact Factor (JIF) and similar measures as proxies for quality in research, promotion, and tenure (RPT) decisions. But a closer examination indicates that the perceived value of JIF is often grounded in **five common myths**:

Large volumes of applications for faculty searches make it difficult for evaluators to distinguish between top-tier candidates, and unintended biases—like the halo effect, availability, and confirmation bias—influence decision making.

Novel research, including breakthrough Nobel-prize winning work², often becomes influential (and cited) outside of the JIF measurement window³, and findings with significant societal impact are not always published in journals with a high JIF.

JIFs are intended to reflect overall journal measures, and do not provide reliable or scientifically sound information about individual articles or researchers⁴.

Forty percent of research-intensive institutions in North America mention JIF in RPT documents, but interpret it inconsistently to mean quality, importance, or prestige⁶.

Faculty members claim to prioritize peer readership when publishing, yet the perception that their peers value prestige and a reliance on university rankings puts pressure on researchers to publish their work in high impact factor journals⁷.

“Invisible work” like service is typically not valued in RPT, yet disproportionately falls on women and other scholars historically excluded from research^{8,10}.

Based on a model of current post-doc to faculty transitions, faculty diversity will not significantly increase until 2080 without active intervention¹¹.

Analogous examples of these myths exist, both inside and outside of science:

There are more male CEOs named John than the total number of female CEOs¹.

John > All women

Low profits, high impact research on extending the life of mangoes transformed the industry, where transportation damage had historically reduced yield by 40% and incurred \$1 billion in losses⁵.

Many some medications are often tailored to genetics, even if they are the same formulation.

Generic Rx vs. Best Brand Rx


A 2019 US poll found that 74 percent of Democrats and Independents were comfortable with the idea of a woman president, but only 33 percent believed their neighbors were⁹.

Only forty-three percent of doctorates in the biomedical sciences are awarded to historically well-represented populations (i.e. white and Asian males), but this same group accounts for 82% of full professorships¹².

<https://sfdora.org/resources/>

RETHINKING RESEARCH ASSESSMENT

UNINTENDED COGNITIVE & SYSTEM BIASES



Judgment and decision-making biases that impact how we weigh options and make choices have been shown to result in inequitable review, promotion, and hiring practices. While recognizing these biases at a personal level is important, creating new structural and institutional conditions to reduce bias can be even more valuable.

People tend to dismiss evidence that doesn't fit their initial judgments or preconceptions

Example: Cherry picking information from a CV to confirm the view one already has, or dismissing potential warning signs because a candidate has already been accepted as a good fit.

Why it's problematic: Our initial conceptions are often based on subjective experiences and limited data. Failing to gather and consider counter-evidence makes us more likely to fall into old ways of thinking.

We often take the path of least resistance unless there are strong reasons not to

Example: Continuing to use citations from academic journals as a primary indicator of impact or quality, rather than considering alternate quantitative indicators of real world value.

Why it's problematic: People often stick with recognizably flawed processes because the effort to fix them or adopt new ones is perceived as too much effort.

Resources often flow to those who already have them

Examples: Highly cited references may be mentioned in part because researchers see that they're highly cited. Researchers with a long track record of grants receive a disproportionate amount of new funding.

Why it's problematic: When people lack the time or motivation to vet results, this can make access to resources even less equitable.

Once metrics are accepted as a way to gauge value, they start to lose meaning as objective measures

Example: Reward systems that rely on easily measurable qualities—like citations and publishing in high-JIF publications—can lead people to “game” the system.

Why it's problematic: When quantitative measures have an outside impact on how people are rewarded, it can increase the temptation to focus on a narrow set of activities and reduce investment in other meaningful, but less rewarded, achievements.

Tackling these infrastructural and institutional implications of common biases can help promote and support more equitable practices:

Incumbent processes and perceptions have the advantage

Many institutions have deep legacy traditions that become normalized over time, but these organizational habits can also keep new ideas and people out.

What can institutions do?

- Make the benefits of new behaviors concrete, salient, and easy to grasp
- Recognize where old assumptions may overly reward those who are more traditionally successful, at the expense of new or more diverse talent
- Set, publicize, and adhere to measurable goals that look beyond traditional norms of success when reviewing potential candidates to broaden the pool of

“Objective” comparisons are not necessarily equitable

Qualities that can be measured or ranked are tempting because they feel less subjective, but can feed a false sense of precision.

What can institutions do?

- Balance the use of quantitative metrics with qualitative inputs, like narrative CVs, that capture more intangible qualities
- Select standards based on a wide set of inputs rather than a narrow or anecdotal set
- Recognize where setting specific, quantifiable goals may be reinforcing some behaviors at the expense of others

Matthew effect

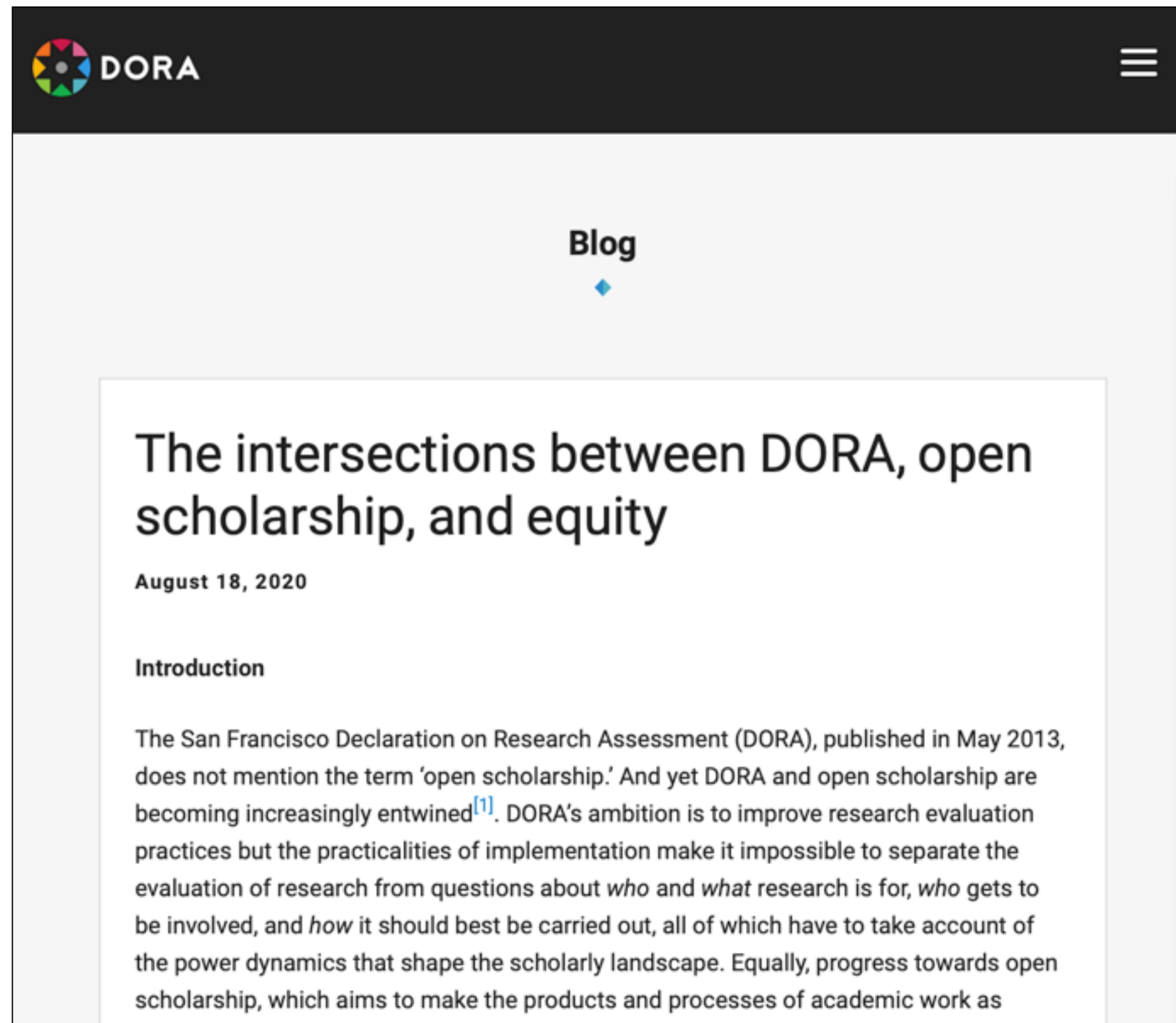
Status quo bias

Confirmation bias

Campbell's law

Anchoring

DORA: we are an important part of a bigger picture



<https://sfdora.org/2020/08/18/the-intersections-between-dora-open-scholarship-and-equity/>

