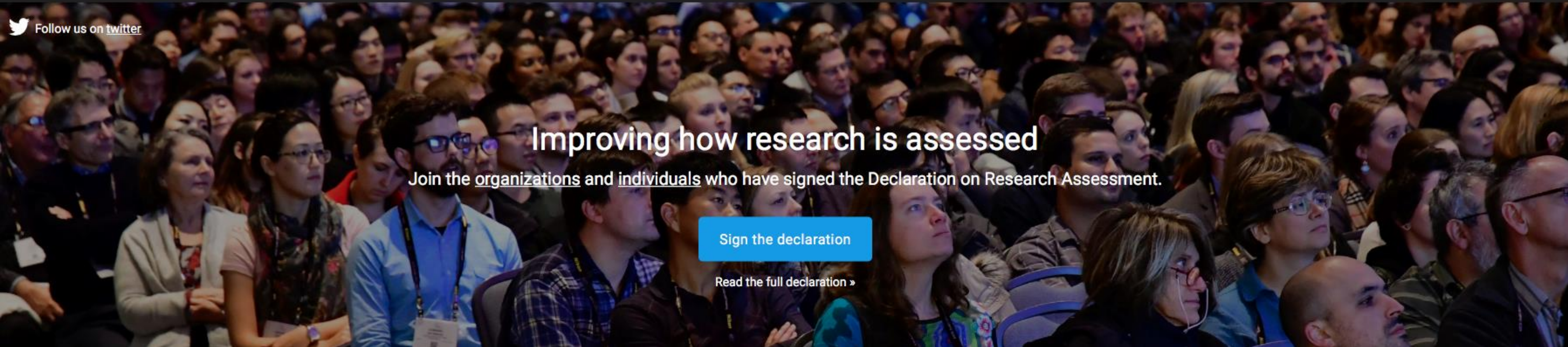


Revisiting research assessment approaches

Philosophy and Practice



[SIGN DORA](#) [READ THE DECLARATION](#) [SIGNERS](#) [BLOG](#) [GOOD PRACTICES](#) [CONTACT US](#)



[Follow us on twitter](#)

Improving how research is assessed

Join the organizations and individuals who have signed the Declaration on Research Assessment.

[Sign the declaration](#)

[Read the full declaration »](#)

Stephen Curry

Imperial College & DORA

EUA Workshop on Research Assessment in the Transition to Open Science | May 2019

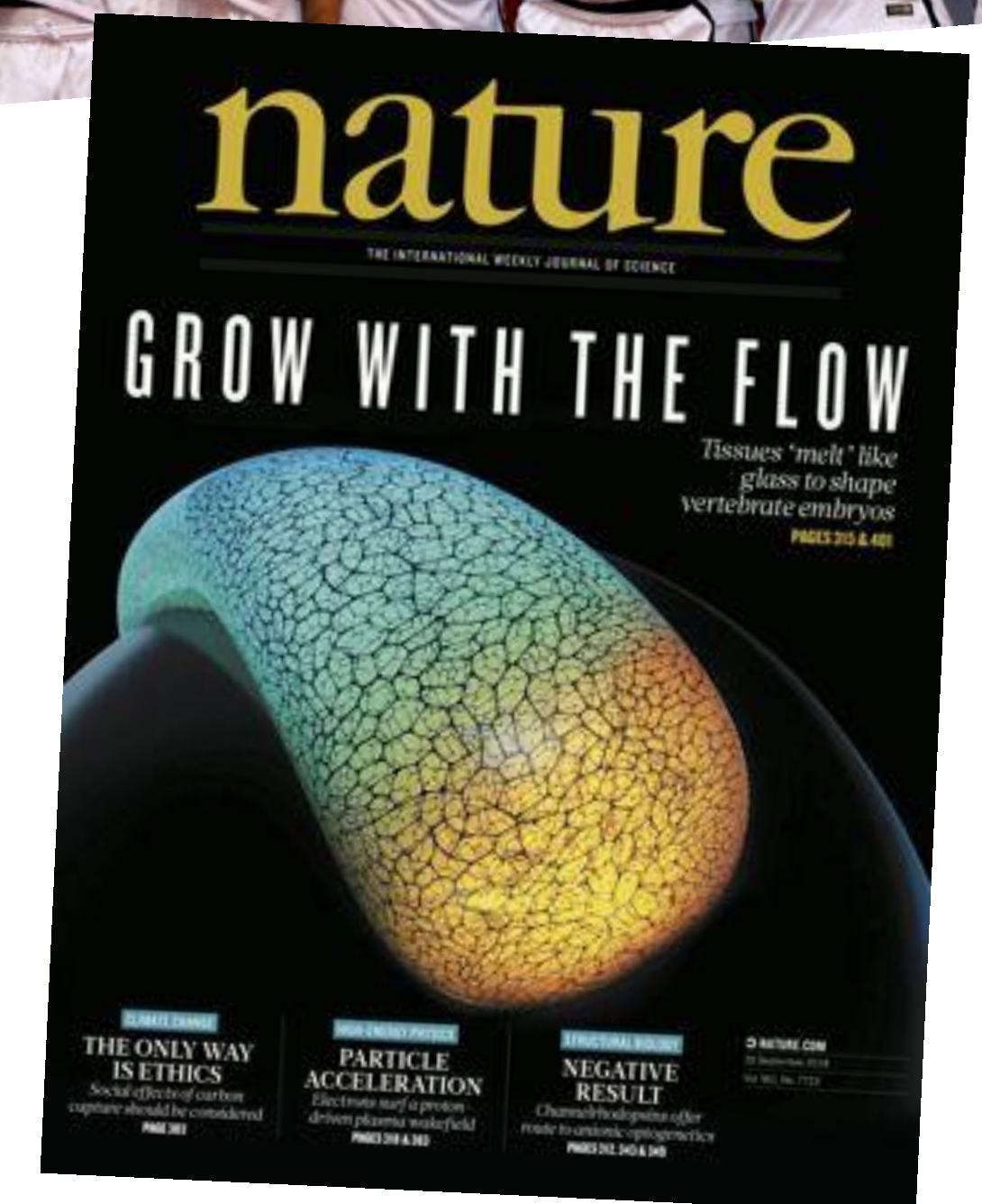
Why do we need research assessment?

To invest finite (and mostly public) resources wisely

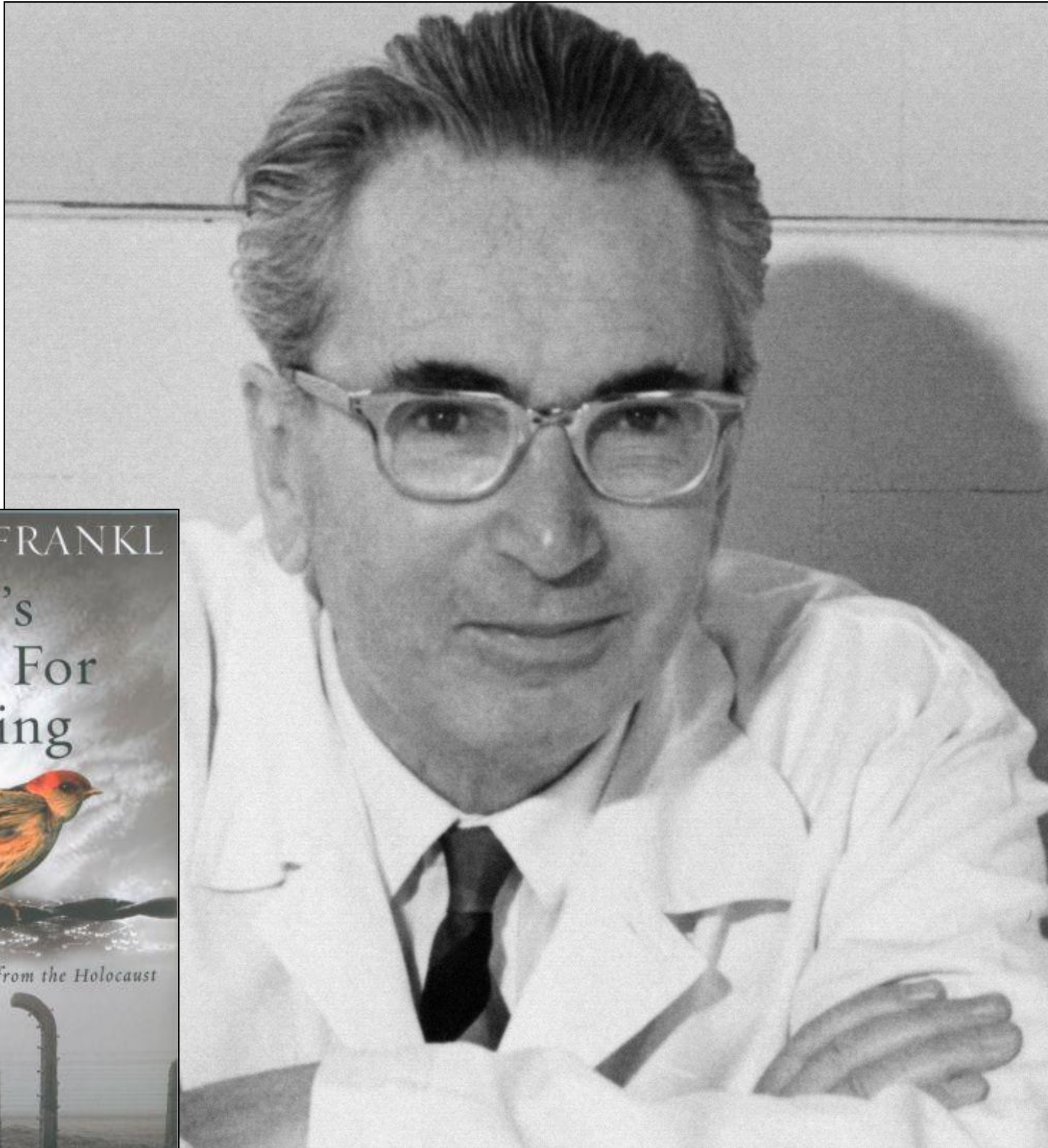
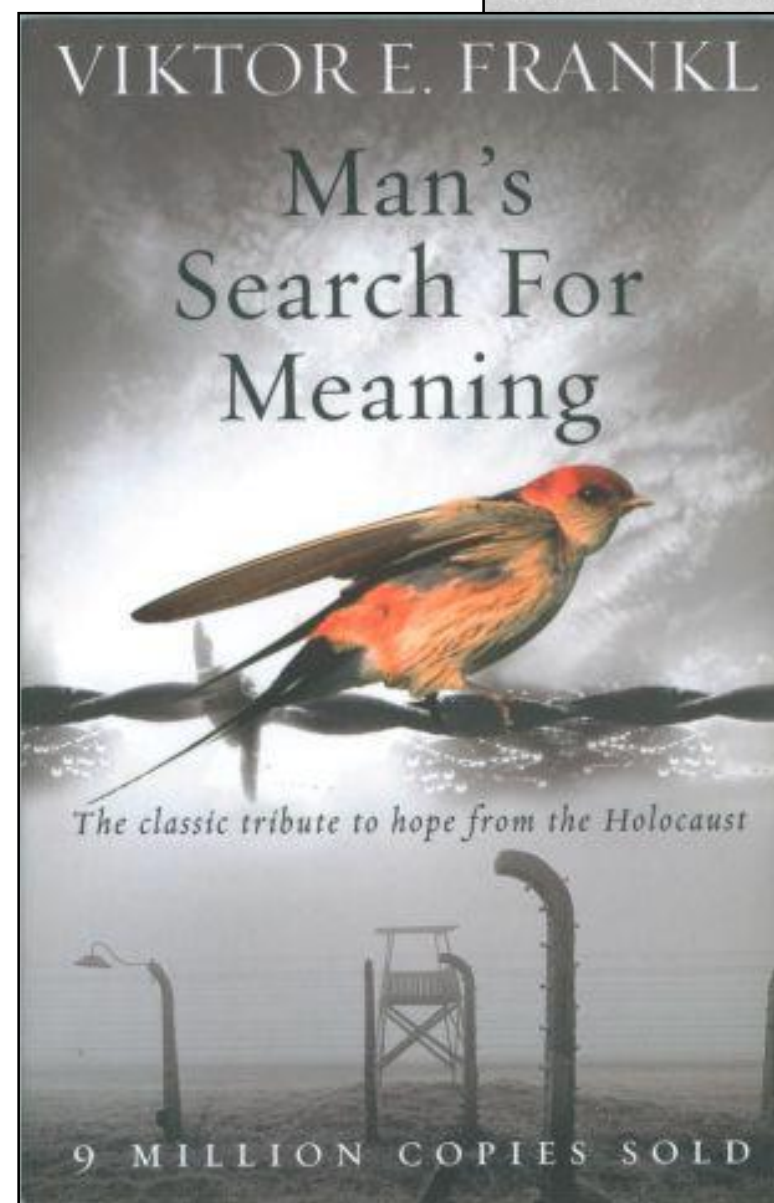
To evaluate returns on those investments

To support and encourage the best science and the best scientists

But what do we mean by ‘best’?



We need to assess research but how should we define success? (Ideal world...)



https://commons.wikimedia.org/wiki/File:Viktor_Frankl2.jpg

“Don’t aim at success [...] for success, like happiness, cannot be pursued; it must ensue, and it only does so as the unintended side-effect of one’s dedication to a cause greater than oneself...”

Viktor Frankl

We need to assess research but how should we define success? (Real world...)



Saving Science

Science isn't self-correcting, it's self-destructing. To save the enterprise, scientists must come out of the lab and into the real world.

Daniel Sarewitz

“much of the problem can be traced back to a **bald-faced but beautiful lie** upon which rests the political and cultural power of science. [...] It goes like this:

*Scientific progress on a broad front results from the **free play of free intellects**, working on **subjects of their own choice**, in the manner dictated by their **curiosity** for exploration of the unknown.”*

Measurement has its uses...



<http://www.aronline.co.uk/blogs/news/news-uk-car-manufacturing-enjoys-bumper-2013/>



<https://www.nuh.com.sg/patients-and-visitors/patients-and-visitors-guide/choice-of-accomodation/ward-types.html>

...but where are the limits?

Rank	Full Journal Title	Total Cites	Journal Impact
1	CA-A CANCER JOURNAL FOR CLINICIANS	28,839	244.585
2	NEW ENGLAND JOURNAL OF MEDICINE	332,830	79.258
3	LANCET	233,269	53.254
4	CHEMICAL REVIEWS	174,920	52.613
5	Nature Reviews Materials	3,218	51.941
6	NATURE REVIEWS DRUG DISCOVERY	31,312	50.167
7	JAMA-JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION	148,774	47.661
8	Nature Energy	5,072	46.859
9	NATURE REVIEWS CANCER	50,407	42.784
10	NATURE REVIEWS IMMUNOLOGY	39,215	41.982
11	NATURE	710,766	41.577
12	NATURE REVIEWS		

The Times Higher Education World University Rankings World University Rankings 2013-2014

1	California Institute of Technology (Caltech)	United States	94.9
2	Harvard University	United States	93.9
2	University of Oxford	United Kingdom	93.9
4	Stanford University	United States	93.8
5	Massachusetts Institute of Technology (MIT)	United States	93.0
6	Princeton University	United States	92.7
7	University of Cambridge	United Kingdom	92.3
8	University of California, Berkeley	United States	89.8
9	University of Chicago	United States	87.8
10	Imperial College London	United Kingdom	87.5
11	Yale University	United States	87.4
12	University of California, Los Angeles (UCLA)	United States	86.3



...and who gets to decide when to use metrics?

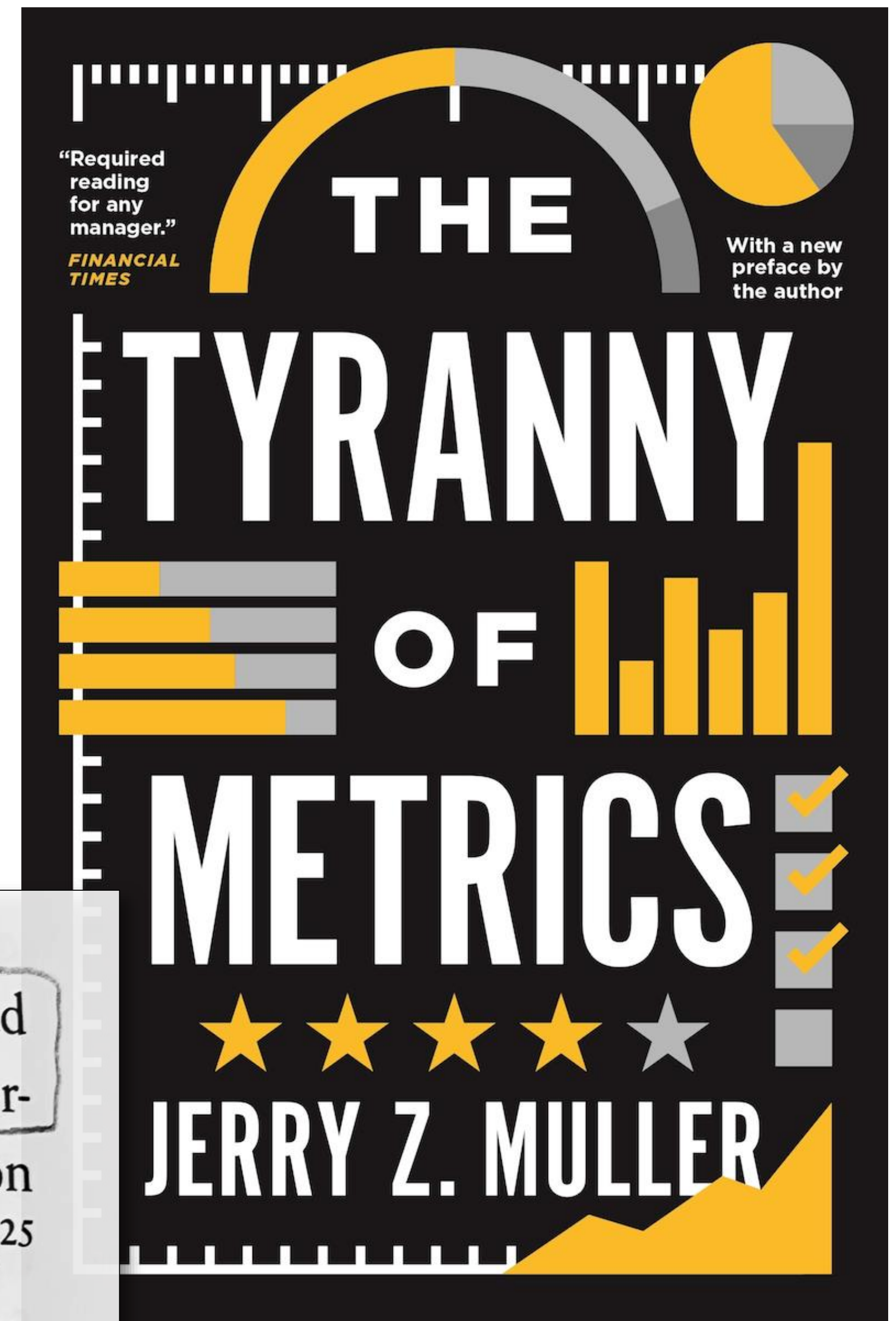
The new managerialism - generalists, not generals...

Metrics can be useful but:

- not if misapplied by people who do not understand their context
- not if tied too tightly to extrinsic rewards

onto them—known in the trade as “window dressing.”²⁴

A focus on measurable performance indicators can lead managers to neglect tasks for which no clear measures of performance are available, as the organizational scholars Nelson Repenning and Rebecca Henderson have recently noted.²⁵ Unable to count intangible assets such as reputation, em-



Negative effects of over-reliance on metrics based on academic outputs

Sick of Impact Factors

Posted on August 13, 2012 by Stephen

I am sick of impact factors and so is science.

The impact factor might have started out as a good idea, but its time has come and gone. [Conceived by Eugene Garfield](#) in the 1970s as a useful tool for research libraries to judge the relative merits of journals when allocating their subscription budgets, the impact factor is [calculated](#) annually as the mean number of citations to articles published in any given journal in the two preceding years.



<http://occamstypewriter.org/scurry/2012/08/13/sick-of-impact-factors/>

“I’m really excited. We just had a big paper in Cell... !”

Postdoc (University of Y)

*“Despite personal ideals and good intentions, in this incentive and reward system **researchers find themselves pursuing not the work that benefits public or preventive health or patient care the most, but work that gives most academic credit** and is better for career advancement.”*

Frank Miedema

<https://blogs.bmj.com/openscience/2018/01/24/setting-the-agenda-who-are-we-answering-to/>

- slows publication & reduces productivity
- positive bias in the literature
- JIF correlates with retraction rate
- impact on reliability & public trust?
- devaluation of other important activities
- stress on the individual

Accentuate the positive: how *open* science can be *better* science

Peer review and scientific publishing
Occam's corner

Stephen Curry
@Stephen_Curry


Monday 7 September 2015 11.00 BST

Shares 1 Comments 14

Save for later

Peer review, preprints and the speed of science

Peer review is often claimed to be the guarantor of the trustworthiness of scientific papers, but it is a troubled process. Preprints offer a way out



Subediting skills for writers Photograph: Victor Moriyama/Getty Images

A few weeks ago my collaborator and I submitted a paper to a journal. We have been investigating the machinery of infected cells and the peer review process. In review, our paper could be published or rejected.

Science
Occam's corner

Zika virus initiative reveals deeper malady in scientific publishing
Stephen Curry

Moves to speed up the release of Zika virus research in response to the public health crisis highlight a systemic failure in scientific publishing. Help could be at hand at the ASAPbio meeting today in the USA

Contact author
@Stephen_Curry

Tuesday 16 February 2016 11.54 GMT

Shares 539 Comments 4

Save for later

Preprints: faster communication; worldwide access
Focus on the content, not the container (journal)

- Valuable groundwork for journal-independent evaluation
Largest possible audience (sharing + scrutiny = reliability)

- Same applies to **OA papers**

Practice encourages **open peer review**

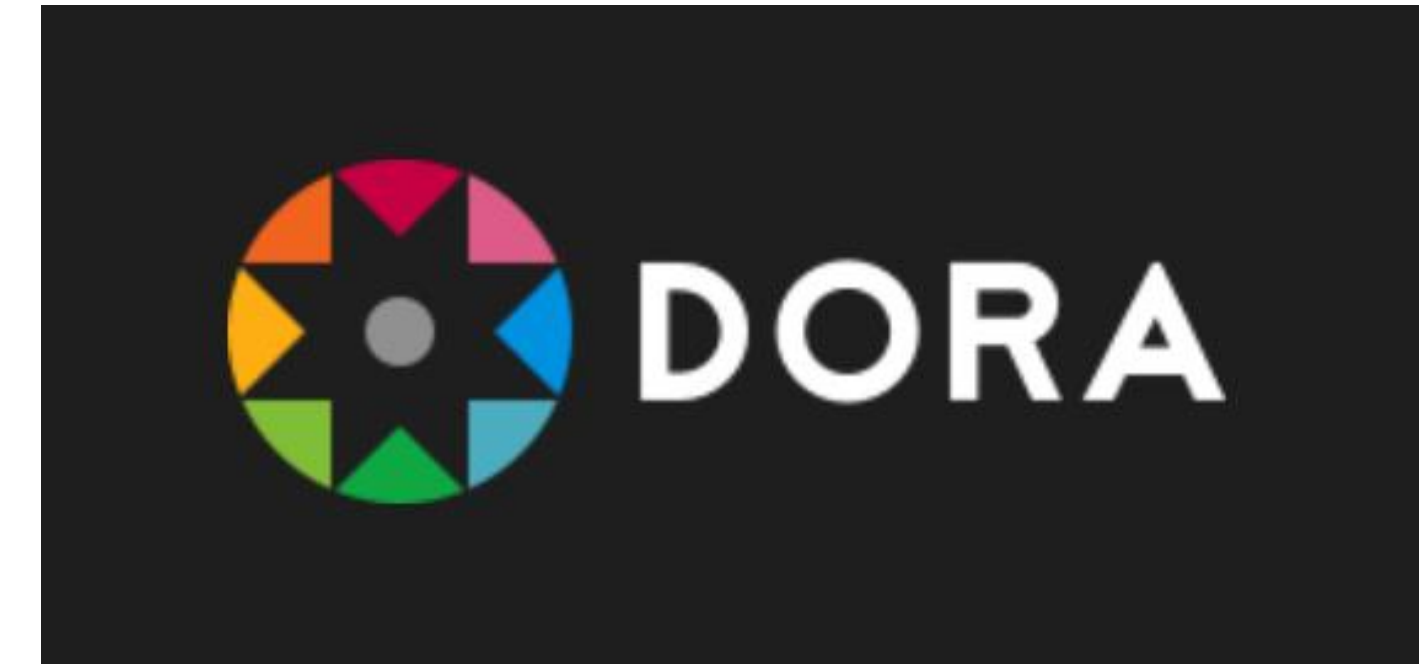
Data sharing: scrutiny benefits (reliability)

Better for changing the world (utility & impact; *e.g.* Zika crisis)

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 *campaign*

San Francisco Declaration on Research Assessment

- 6 years old; >14,000 individuals and >1300 organisations signed
- 2018: New funding, new steering group, new URL - [sfdora.org](https://www.nature.com/articles/d41586-018-01642-w)
- New Roadmap for **action**:
 - Increase awareness of the need to develop alternatives to the JIF
 - **Research and promote best practice in research assessment.**
 - Extend the global and disciplinary impact of DORA
- New international advisory board – a truly global initiative

<https://www.nature.com/articles/d41586-018-01642-w>

WORLD VIEW

A personal take on events



Words were a good start — now it is time for action

Five years ago, the Declaration on Research Assessment was a rallying point. It must now become a tool for fair evaluation, urges **Stephen Curry**.

Declarations are bound to fall short. The 240-year-old United States Declaration of Independence holds it self-evident that “all men [sic] are created equal”, but equality remains a far-off dream for many Americans.

The San Francisco Declaration on Research Assessment (DORA; <https://sfdora.org>) is much younger, but similarly idealistic. Conceived by a group of journal editors and publishers at a meeting of the American Society for Cell Biology (ASCB) in December 2012, it proclaims a pressing need to improve how scientific research is evaluated, and asks scientists, funders, institutions and publishers to forswear using journal impact factors (JIFs) to judge individual researchers.

DORA’s aim is a world in which the content of a research paper matters more than the impact factor of the journal in which it appears. Thousands of individuals and hundreds of research organizations now agree and have signed up. Momentum is building, particularly in the United Kingdom, where the number of university signatories has trebled in the past two years. This week, all seven UK research councils announced their support.

Impact factors were never meant to be a metric for individual papers, let alone individual people. They’re an average of the skewed distribution of citations accumulated by papers in a given journal over two years. Not only do these averages hide huge variations between papers in the same journal, but citations are imperfect measures of quality and influence. High-impact-factor journals may publish a lot of top-notch science, but we should not outsource evaluation of individual researchers and their outputs to seductive journal metrics.

Most agree that yoking career rewards to JIFs is distorting science. Yet the practice seems impossible to root out. In China, for example, many universities pay impact-factor-related bonuses, inspired by unwritten norms of the West. Scientists in parts of Eastern Europe cling to impact factors as a crude bulwark against cronyism. More worryingly, processes for JIF-free assessment have yet to gain credibility even at some institutions that have signed DORA. Stories percolate of research managers demanding high impact factors. Job and grant applicants feel that they can’t compete unless they publish in prominent journals. All are fearful of shunning off the familiar harness.

So, DORA’s job now is to accelerate the change it called for. I feel the need for change whenever I meet postdocs. Their curiosity about the world and determination to improve it burns bright. But their desires to pursue the most fascinating and most impactful questions are subverted by our systems of evaluation. As they apply for their first permanent positions, they are already calculating how to manoeuvre within the JIF-dependent managerialism of modern science.

There have been many calls for something better, including the Leiden Manifesto and the UK report “The Metric Tide”, both released in

2015. Like DORA, these have changed the tenor of discussions around researcher assessment and paved the way for change.

It is time to shift from making declarations to finding solutions. With the support of the ASCB, Cancer Research UK, the European Molecular Biology Organization, the biomedical funder the Wellcome Trust and the publishers the Company of Biologists, *eLife*, F1000, Hindawi and PLOS, DORA has hired a full-time community manager and revamped its steering committee, which I head. We are committed to getting on with the job.

Our goal is to discover and disseminate examples of good practice, and to boost the profile of assessment reform. We will do that at conferences and in online discussions; we will also establish regional nodes across the world, run by volunteers who will work to identify and address local issues.

This week, for example, DORA is participating in a workshop at which the Forum for Responsible Metrics — an expert group established following the release of “The Metric Tide” — will present results of the first UK-wide survey of research assessment. This will bring broader exposure to what universities are thinking and doing, and put the spotlight on instances of good and bad practice.

We have to get beyond complaining, to find robust, efficient and bias-free assessment methods. Right now, there are few compelling options. I favour concise one- or two-page “bio-sketches”, similar to those rolled out in 2016 by the University Medical Centre Utrecht in the Netherlands. These let researchers summarize their most important research contributions, plus mentoring, societal engagement and other valuable activities. This approach could have flaws. Perhaps it gives too much leeway for ‘spin’. But, as scientists, surely we can agree that it’s worth doing the experiment to properly evaluate evaluation.

This is hard stuff: we need frank discussions that grind through details, with researchers themselves, to find out what works and to forestall problems. We need to be mindful of the damage wrought to the careers of women and minorities by bias in peer review and in subjective evaluations. And we need to join in with parallel moves towards open research, data and code sharing, and the proper recognition of scientific reproducibility.


Declarations such as DORA are important; credible alternatives to the status quo are more so. True success will mean every institution, everywhere in the world, bragging about the quality of their research-assessment procedures, rather than the size of their impact factors. ■

IT’S WORTH DOING THE EXPERIMENT TO PROPERLY EVALUATE EVALUATION.


Stephen Curry is a professor of structural biology and assistant provost for equality, diversity and inclusion at Imperial College London. He is also chair of the DORA steering group.
e-mail: s.curry@imperial.ac.uk

8 FEBRUARY 2018 | VOL 554 | NATURE | 147

© 2018 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.



SIGN DORA READ THE DECLARATION SIGNERS BLOG GO


 Follow us on twitter

Improving how research is assessed


Join the organizations and individuals who have signed the Declaration on Research Assessment.


Sign the declaration


Read the full declaration »




an international forum for cell biology™













excellence in life sciences



Open for Science







11

New tools and processes for assessment

<http://www.nature.com/news/fewer-numbers-better-science-1.20858>



Fewer numbers, better science

Scientific quality is hard to define, and numbers are easy to look at. But bibliometrics are warping science — encouraging quantity over quality. Leaders at two research institutions describe how they do things differently.

Researcher assessment at UMC Utrecht

1. Research, publications, grants
2. Managerial & academic duties
3. Mentoring & teaching
4. Clinical work (if applicable)
5. Entrepreneurship & community outreach

Charité University Hospital, Berlin

- Scientific contribution to your field
- Your 5 most important papers
- Contribution to open science
- Your most important collaborations

New tools and processes for assessment

DORA session at AAAS (Feb 2019)

SCIENCE TRANSCENDING BOUNDARIES

AAASANNUAL MEETING
Washington, DC | Feb 14–17, 2019

REGISTRATIONWHAT'S NEWPROGRAME-POSTERSATTENDSPONSORS & EXHIBITS

HOME

SCIENTIFIC SESSION

Academic Research Assessment: Reducing Biases in Evaluation

SEARCH

DORA

SIGN DORAREAD THE DECLARATIONSIGNERSBLOGGOOD PRACTICESRESOURCESCONTACT US

Blog

DORA – accentuating the positive

May 9, 2019

The San Francisco Declaration on Research Assessment (DORA) is well known for its strong position on the need to eliminate the use of journal-based metrics in decisions on hiring, promotion, or funding of academics. As such, it is sometimes taken to be an initiative merely focused on criticising the undue influence of one specific metric, the journal impact factor (JIF). But to see DORA just in those terms overlooks the many positive prescriptions that the declaration lays out for how to reform research assessment

Read more

DORA news and announcements

Impacts, Outputs, and Approaches to Assessment in the Humanities

Categories

Community interviews

DORA news and announcements

Guest posts

Miscellaneous

The DORA Roadmap

Recent Posts

DORA – accentuating the positive
May 9, 2019

Impacts, Outputs, and Approaches to Assessment in the Humanities
April 2, 2019

#sfDORA community interview: exploring impact in the Humanities and Social Sciences
March 14, 2019

Breaking habits: reducing bias in hiring, promotion and

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

DORA session at ASCB | EMBO (Dec 2018)

eLIFE

ABOUTCOMMUNITYSUBMIT MY RESEARCHID LOG IN/REGISTER

HOME MAGAZINE INNOVATION

Research Assessment: Reducing bias in the evaluation of researchers

A workshop run by DORA identified a number of ways to reduce bias in hiring and funding decisions.

f

t

e

r

INSIDE ELIFE Apr 17, 2019


VIEWS 1,375ANNOTATIONS 0

By Anna Hatch (DORA), Veronique Kiermer (PLOS), Bernd Pulverer (EMBO), Erika Shugart (American Society for Cell Biology), and Stephen Curry (Imperial College London)

Introduction

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

Plan S and research evaluation



SearchE-alertSubmitLogin

News & CommentMore ▾




Research


NEWS • 04 SEPTEMBER 2018

Radical open-access plan could spell end to journal subscriptions

Eleven research funders in Europe announce 'Plan S' to make all scientific works free to read as soon as they are published.

Holly Else








Robert-Jan Smits, the European Commission's special envoy on open access, spearheaded the Plan S initiative. Credit: Nikolay Doychinov/EU2018BG

PDF version

LATEST NEWS ARTICLES

Ice-tracking satellite launches after 10 years in the works

Stand back, Aquaman: Harpoon-throwing satellite takes aim at space junk

AI helps unlock 'dark matter' of bizarre superconductors

Research funders from France, the United Kingdom, the Netherlands and eight other European nations have

Display a menu

dical open-access initiative that could

“We also understand that researchers may be driven to do so by a misdirected reward system which puts emphasis on the wrong indicators (e.g. journal impact factor). **We therefore commit to fundamentally revise the incentive and reward system of science, using the San Francisco Declaration on Research Assessment (DORA) as a starting point.**

<https://www.scienceeurope.org/coalition-s/>

News | 5 November 2018

Wellcome is updating its open access policy

Following a six-month review, we're updating our open access (OA) policy. The changes will apply from 1 January 2020. Robert Kiley, Head of Open Research, explains what will be different and why.

“5. Wellcome-funded organisations must sign or publicly commit to the San Francisco Declaration on Research Assessment (DORA), or an equivalent. We may ask organisations to show that they're complying with this as part of our organisation audits. This is a new requirement to encourage organisations to consider the intrinsic merit of the work when making promotion and tenure decisions, not just the title of the journal or publisher.”

<https://www.nature.com/articles/d41586-018-06178-7>

We need to assess research but how should we define success?



<http://www.newyorker.com/magazine/2013/07/29/slow-ideas>

“We yearn for frictionless, technological solutions. But people talking to people is still how the world’s standards change.”

Atul Gawande

What should success look like?

Reliable, rapidly communicated, accessible, high-quality **research** that transforms our understanding of the world and can change it for the better.

Researchers who collaborate, who feel a duty of care to group members & colleagues, and a responsibility to the societies of which they are an integral part.

A **research system** that values the people within it, that considers their quality of life, their mental health, and that provides the training and processes to seek out the creative vigour of diversity.

Thank you

s.curry@imperial.ac.uk

@Stephen_Curry

Wouldn't a focus on research teams be better...?

Molecular Cell
Forum

Cell
PRESS

How to Build a Motivated Research Group

Uri Alon^{1,*}

¹Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot 76100, Israel

*Correspondence: urialon@weizmann.ac.il

DOI 10.1016/j.molcel.2010.01.011

Motivated group members experience a full sense of choice: of doing what one wants. Such behavior shows high performance, is enjoyable, and enhances innovation. This essay describes principles of building a motivated research group.

[https://www.cell.com/molecular-cell/fulltext/S1097-2765\(10\)00040-7](https://www.cell.com/molecular-cell/fulltext/S1097-2765(10)00040-7)

A Bigger Prize

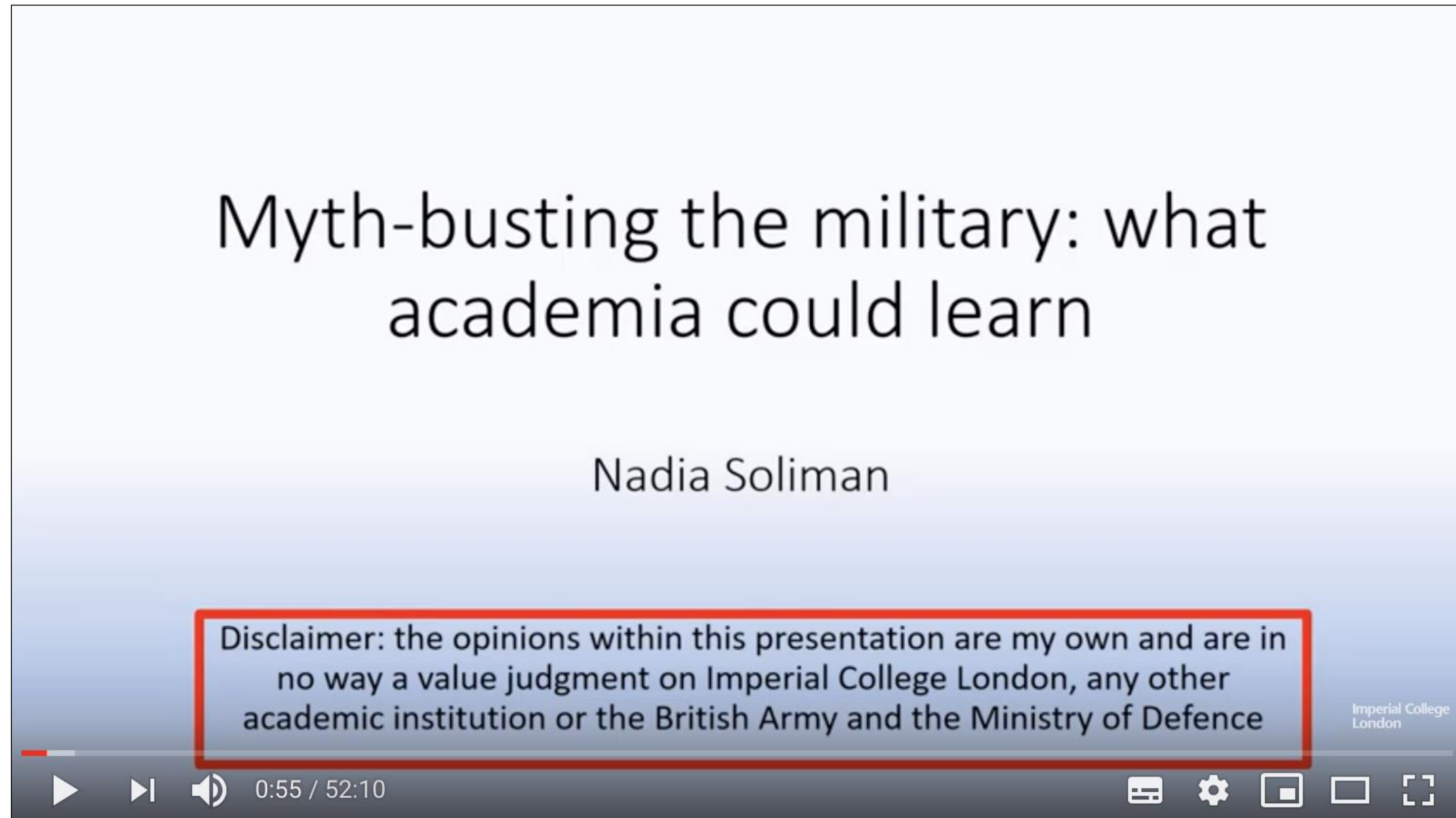
Why no one wins unless everyone wins



MARGARET HEFFERNAN

Author of shortlisted FT/Goldman Sachs Business Book of the Year *Wilful Blindness*

Or a focus on better training of research *leaders*...?



<https://www.youtube.com/watch?v=b2sgp7Kjy8>



“I wish I’d had the courage to live a life true to myself, not the life others expected of me.”

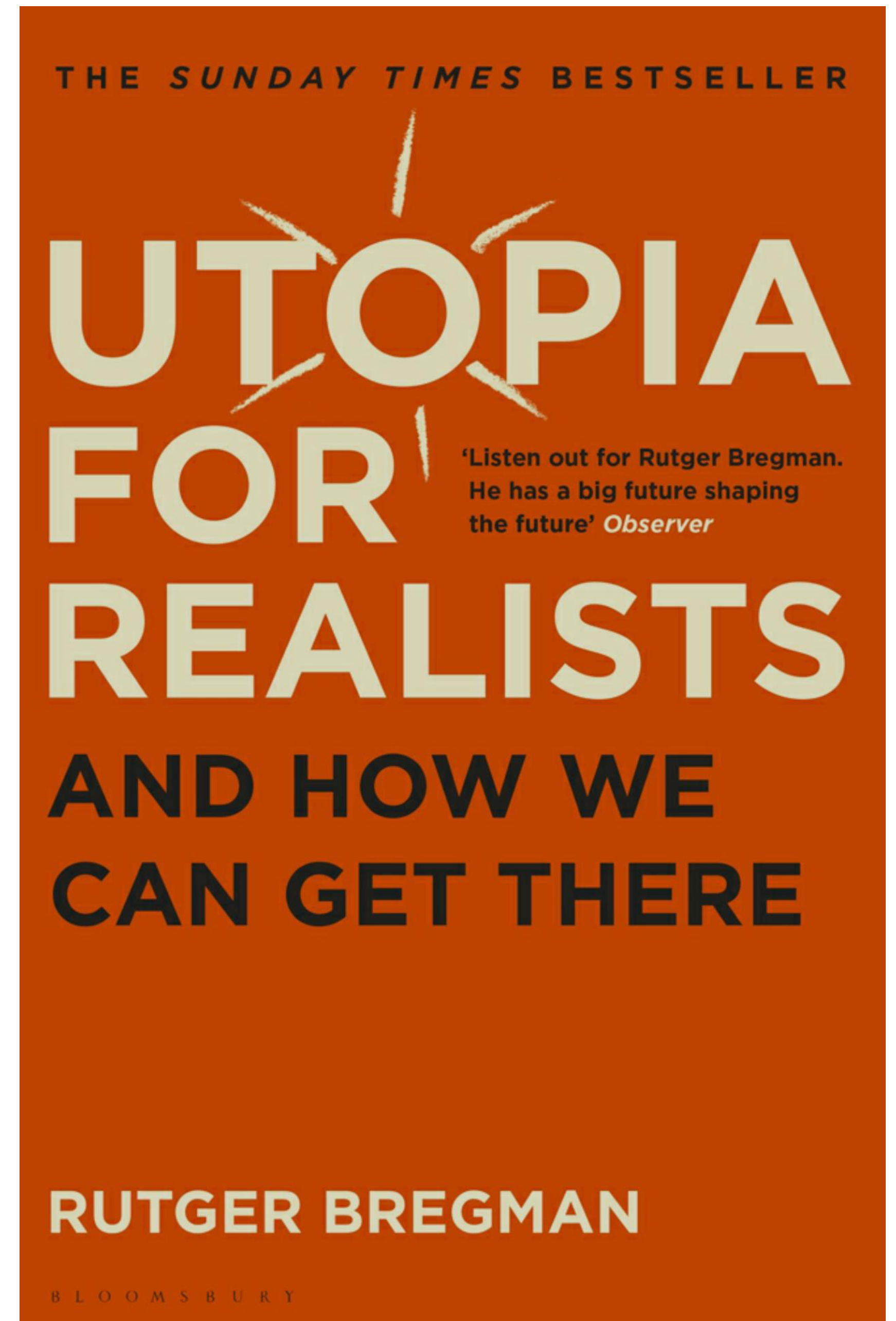
From ‘The top five regrets of the dying’ by Bronnie Ware (No. 1)

Utopia for Realists, Rutger Bregman

“Governing by numbers is the last resort of a **country that no longer knows what it wants**, a country with no vision of utopia.”

“What we need is a ‘**dashboard**’ complete with an **array of indicators** to track the things that make life worthwhile – money and growth, obviously, but also community service, jobs, knowledge, social cohesion. And, of course, the scarcest good of all: time. ‘But such a dashboard couldn’t possibly be objective,’ you might counter. True. But **there’s no such thing as a neutral metric**. Behind every statistic is a certain set of assumptions and prejudices.”

“If we want to change the world, we need to be **unrealistic, unreasonable, and impossible**.”



The future...



<https://publications.europa.eu/en/publication-detail/-/publication/47a3a330-c9cb-11e7-8e69-01aa75ed71a1/language-en>

Evaluation of Research Careers fully acknowledging Open Science Practices

Rewards, incentives and/or recognition for researchers practicing Open Science

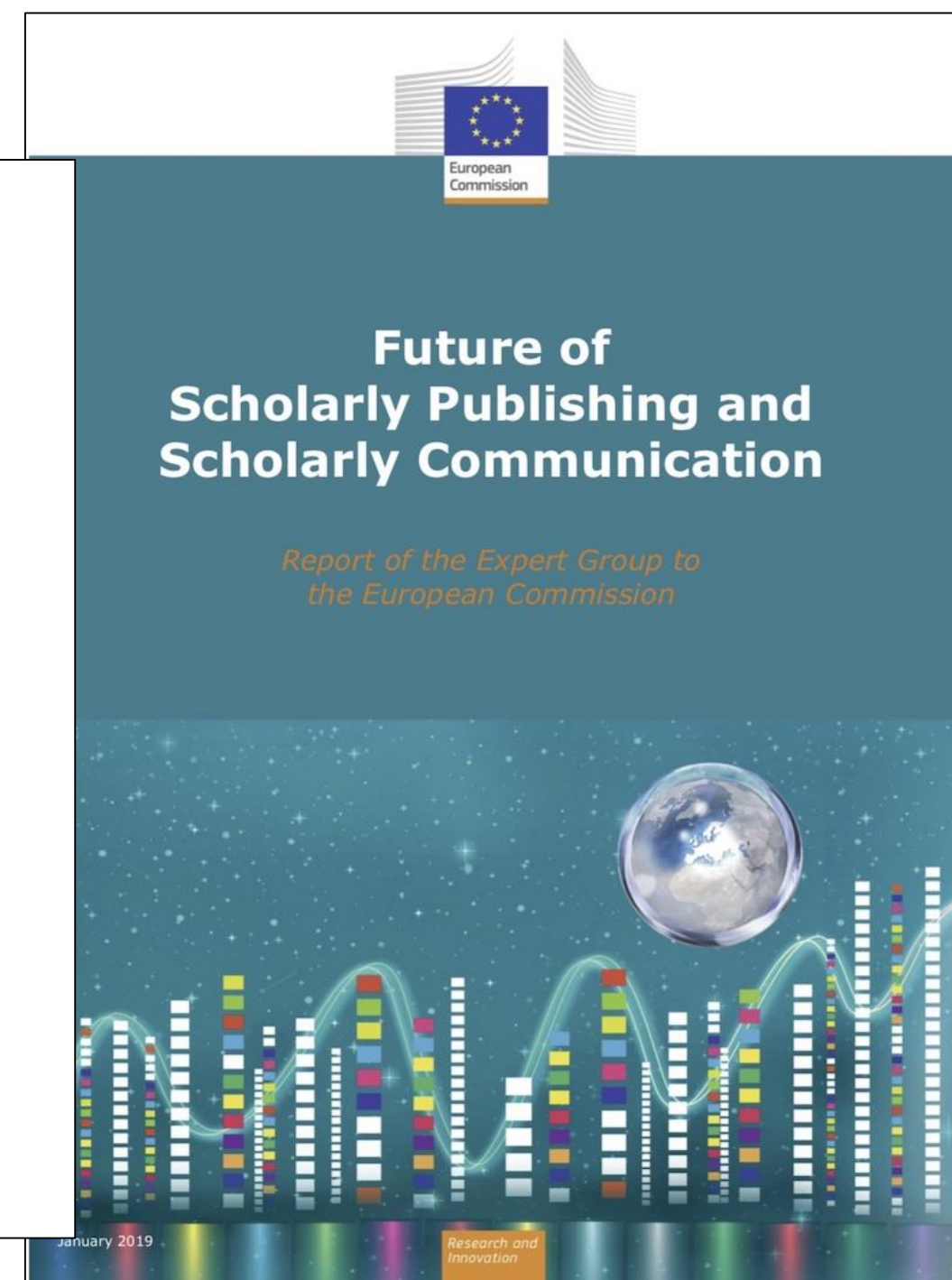
EXECUTIVE SUMMARY

Open Science represents an approach to research that is collaborative, transparent and accessible¹. There are a wide range of activities that come under the umbrella of Open Science that include open access publishing, open data, open peer review and open research. It also includes citizen science, or more broadly, stakeholder engagement, where non specialists engage directly in research. Open Science goes hand in hand with research integrity and requires legal and ethical awareness on the part of researchers. A driver for Open Science is improving the transparency and validity of research as well as in regards to public ownership of science, particularly that which is publicly funded.

The conclusion is actually simple: the evaluation of research is the keystone, and it has already been identified by scholars around the world, and by various expert groups within the European Commission, as structuring a global research architecture characterised by an unlimited quest for rankings. The ranking imperative affects all levels of the research structure, and it tends to constrain change for nearly all actors. This is true of individual researchers, of research groups, of whole research institutions, and even of whole countries. Symmetrically, publishers design their marketing strategies around journal rankings. But they too have become prisoners of this strategy, even though they benefit from it, and they have difficulties seeing beyond it.

Funding agencies also use rankings, sometimes abundantly. However, unlike the other actors, private funding charities are not ranked, and public, national, funders are ranked only indirectly, through their own country. As a result, funders in general enjoy more latitude than the other actors in scholarly communication and publishing. The European

<https://publications.europa.eu/en/publication-detail/-/publication/464477b3-2559-11e9-8d04-01aa75ed71a1>



A brief history of research assessment reform...

May 2013



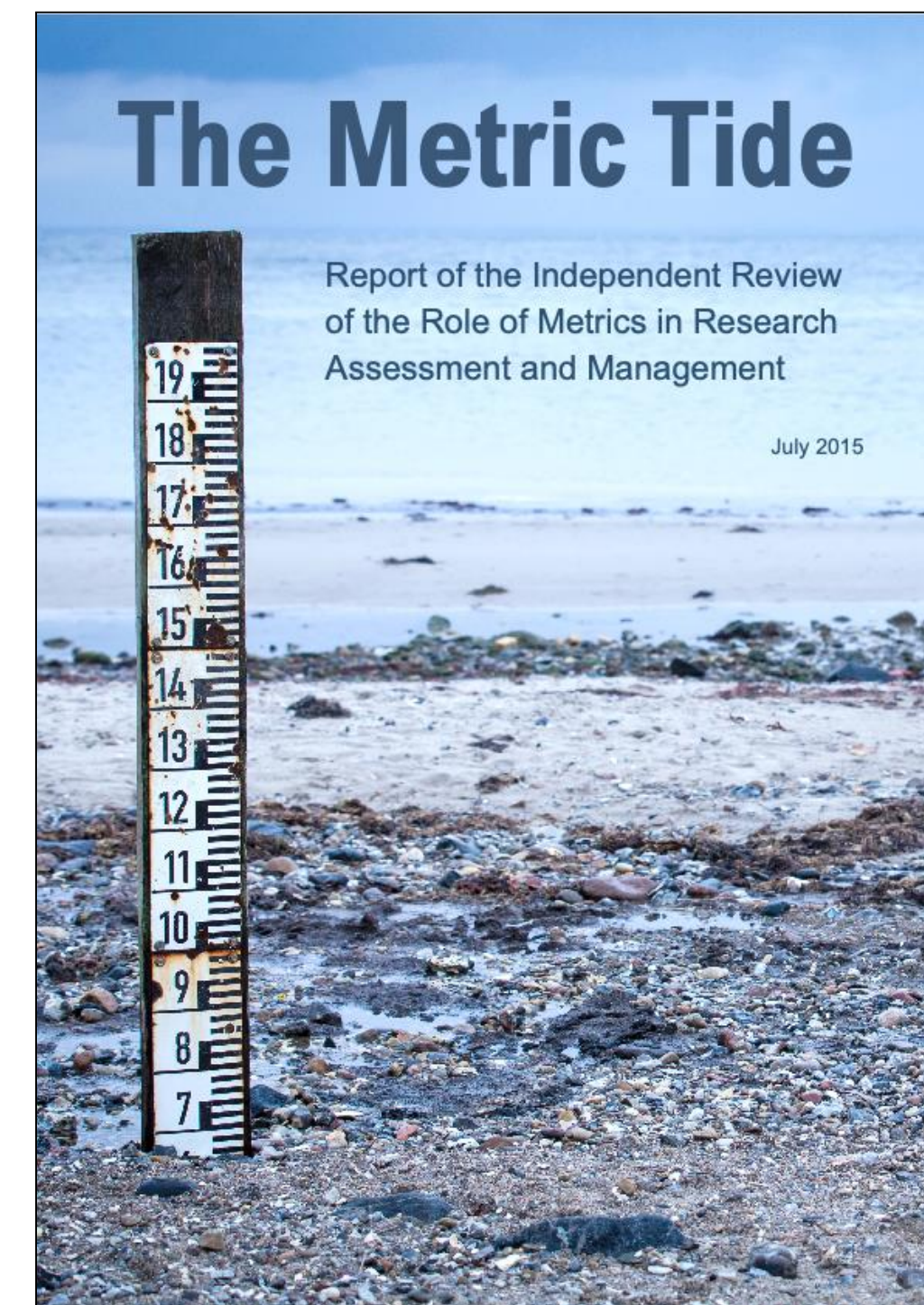
<https://sfdora.org>

Mar 2015



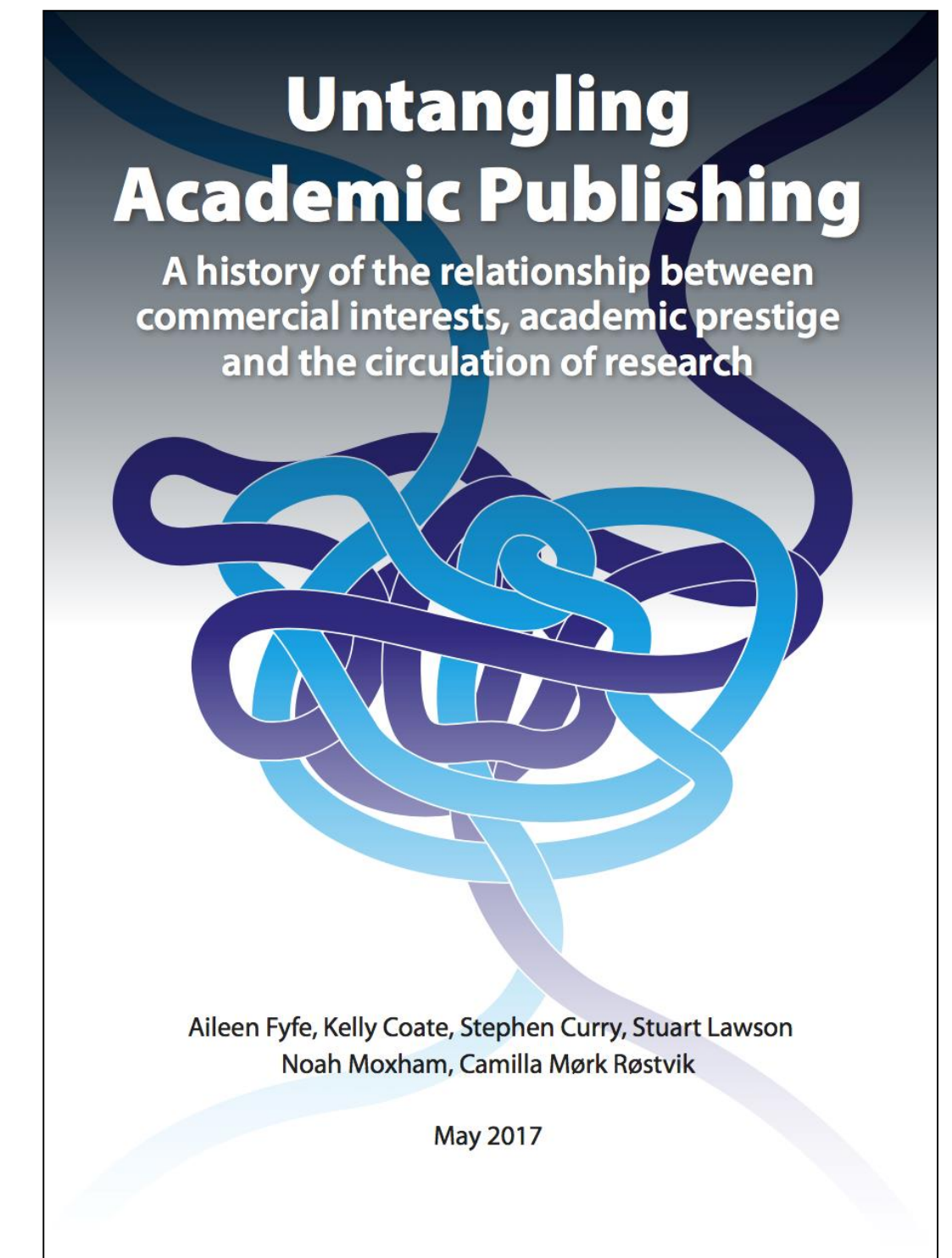
<http://www.leidenmanifesto.org>

Jul 2015



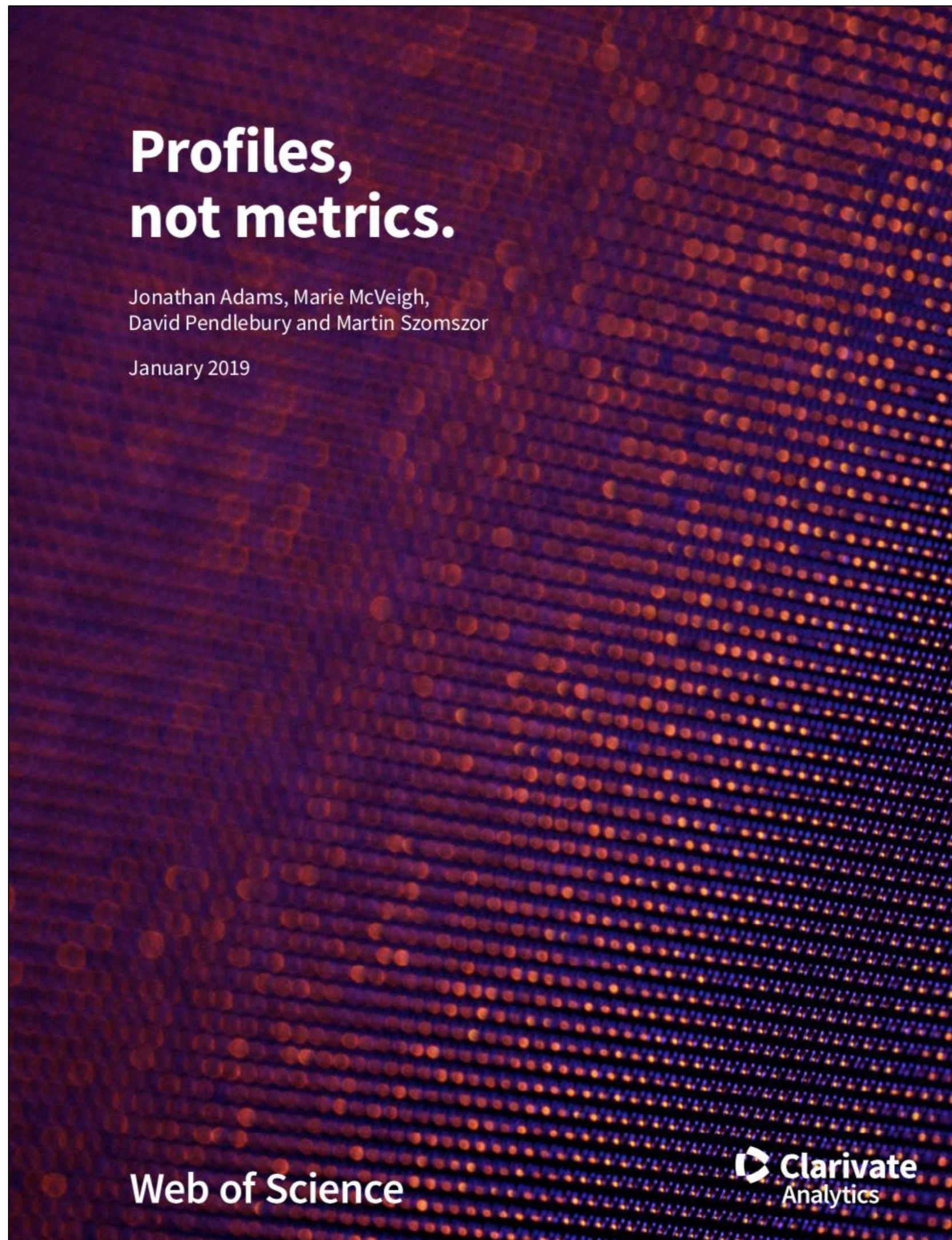
[UK Forum for Responsible Research Metrics](https://www.ukforumforresponsiblemetrics.org/)

May 2017



<https://zenodo.org/record/546100#.XKESdy2ZPOQ>

Profiles, not metrics



The point metrics (h-index, *Journal Impact Factor*, average citation impact) and the university ranking discussed in this report are all potentially informative but all suffer from widespread misinterpretation and irresponsible and often gross misuse. The alternative visual analyses are ‘picture profiles’ of research activity. They are graphical illustrations that: are relatively simple to produce; unpack a spread of much more valuable information; and support proper and responsible research management.

A beam-plot, not an h-index

The beam-plot is a single ‘picture’ of a researcher’s output and impact, showing how it varies within a year and evolves over time. The use of percentiles means that citation impact, which is highly skewed, can be seen in a context appropriate to both discipline and time since publication. Reducing this to the single value of an h-index may be an intriguing summary but it tells us nothing we can properly use in evaluation.

A Journal Profile Page, not just the JIF

The *Journal Impact Factor* (JIF) suffers from misapplication. It isn’t about research evaluation but about journal management. Putting JIF into a context that sets that single point value into a profile or spread of activity enables researchers and managers to see that JIF draws in a very wide diversity of performance at article level. JIF may be a guide but the full context is needed for real information outside the library and publishing house.

An Impact Profile™, not an isolated CNCI

A summary index of the average Category Normalized Citation Impact (CNCI) can also be misleading, because it submerges a diverse data spread which, as at individual and journal level, is highly skewed and subject to outlier values. The Impact Profile™ shifts that skew into a more digestible form and reveals the underlying distribution. It shows that the spread around a world average and an institutional average means that many outputs are inevitably cited more and others less often. Whereas the summary value told us nothing more than X had a higher average than Y, the Impact Profile™ points up a whole series of questions, but also provides routes to answers for research management: where are the collaborative papers; do the same people produce both high and low cited material; did we shift across time?

A Research Footprint, not a university ranking

The ranking table of universities suppresses far more information than most analyses. The Research Footprint can unpack performance by discipline or by data type. It can compare two institutions or countries, or it can compare a series of target organizations to a suitable benchmark. Critically, it shows that there is no sensible way to compare two complex research systems with a single number: it’s a bit more complicated than that!

The old proverb says that a picture is worth a thousand words. Visualizing a data distribution is worth a thousand single-point metrics.