The 6 W‘s of Open Science

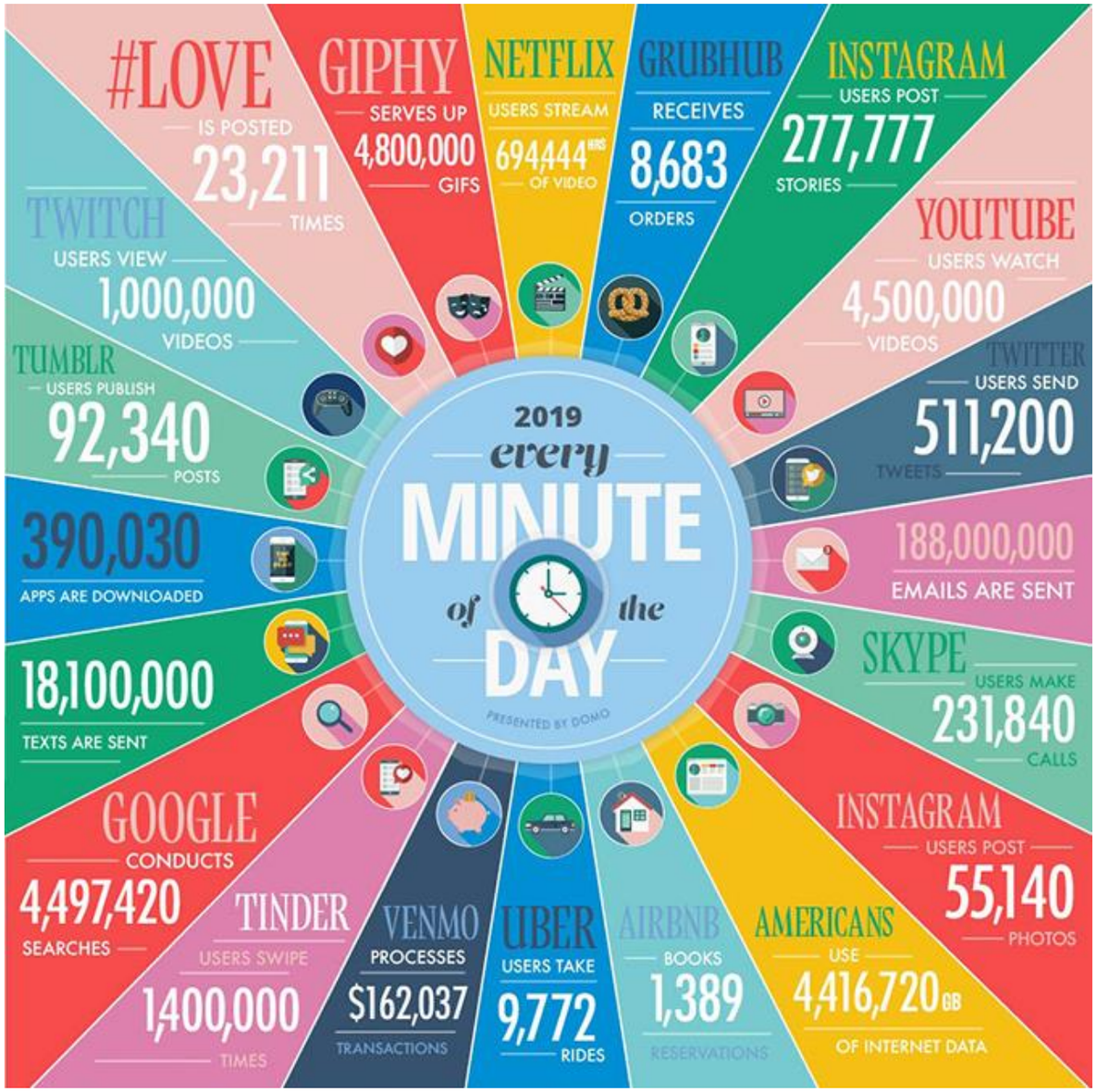
Prof. Dr. Isabella Peters, Web Science

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Leibniz-Gemeinschaft

CAU

What are the roots of open science?



What are the roots of open science?

Social media use of economists and other social scientists (N=3,400)

Most Popular Services

	Service	Use
1.	Google Scholar	75.2%
2.	JSTOR	60.7%
3.	Wikipedia	60.1%
4.	ResearchGate	55.6%
5.	EBSCO	36.7%

Most Frequently Used Services

	Service	Use daily
1.	Google Scholar	31.0%
2.	Wikipedia	21.0%
3.	Facebook	14.0%
4.	Twitter	10.0%
5.	JSTOR	10.0%

Long-Established Services

	Service	Use for 5+ years
1.	Wikipedia	81.9%
2.	Amazon	73.6%
3.	JSTOR	70.2%
4.	Facebook	66.6%
5.	ICPSR	65.6%

Up-and-Coming Services

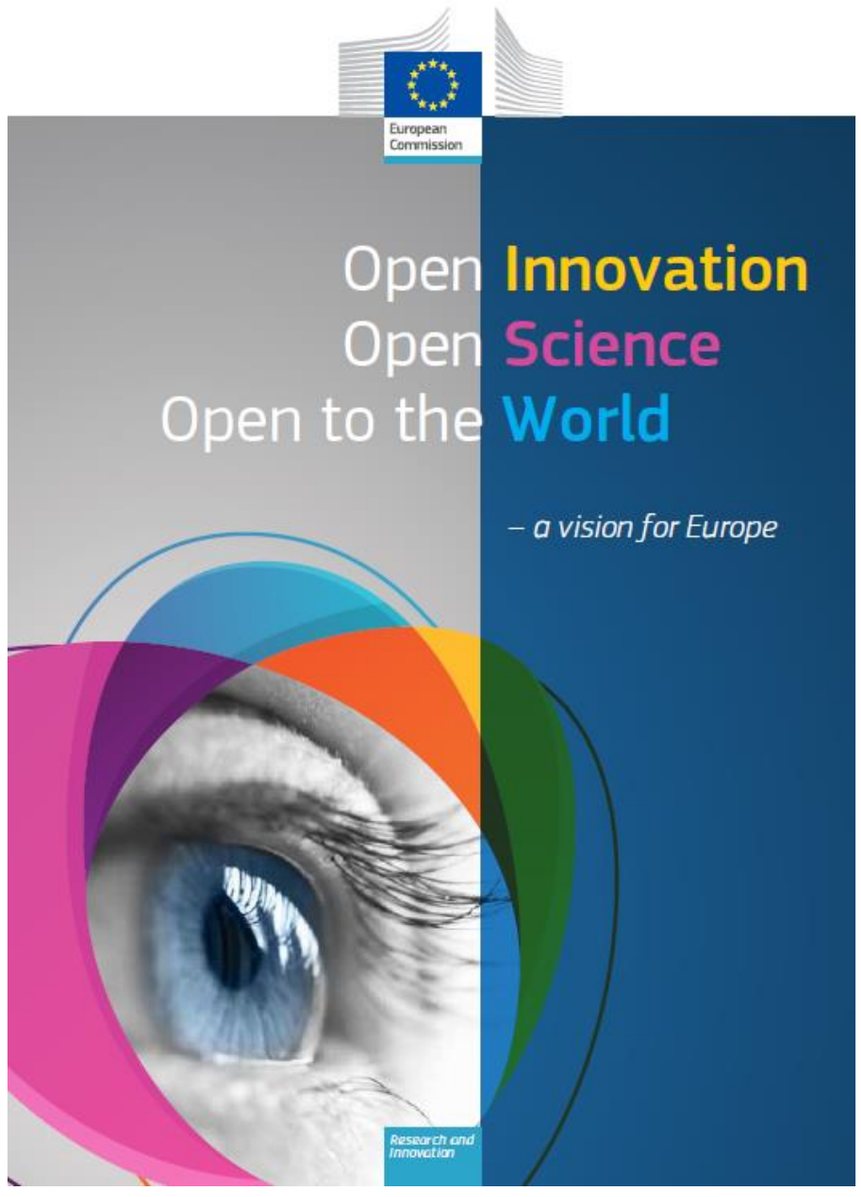
	Service	Use <1 year
1.	ResearchGate	5.1%
2.	Academia.edu	2.7%
3.	Mendeley	2.6%
4.	sci-hub.io	2.5%
5.	Twitter	2.5%

Based on: Lemke et al. (2017). Exploring the Meaning and Perception of Altmetrics. http://doi.org/10.5281/zenodo.1037146

What are the roots of open science?

Carlos Moedas
Commissioner for Research, Science & Innovation

„…the way that science works is fundamentally changing and an equally important transformation is taking place in how companies and societies innovate. Put simply, the advent of digital technologies is making science and innovation more open, collaborative, and global.“



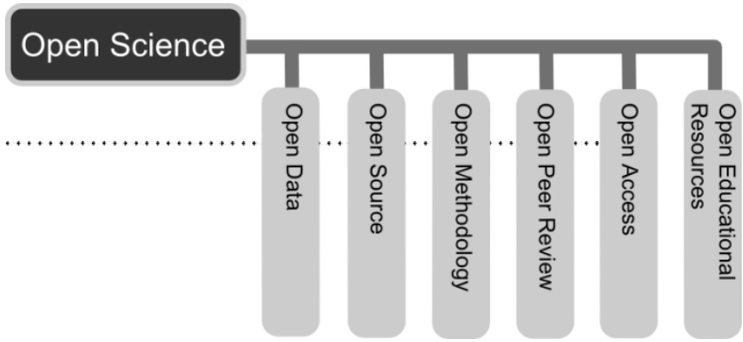
- Mai 2016
- Directorate-General for Research & Innovation
- DOI: 10.2777/061652

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Leibniz-Gemeinschaft

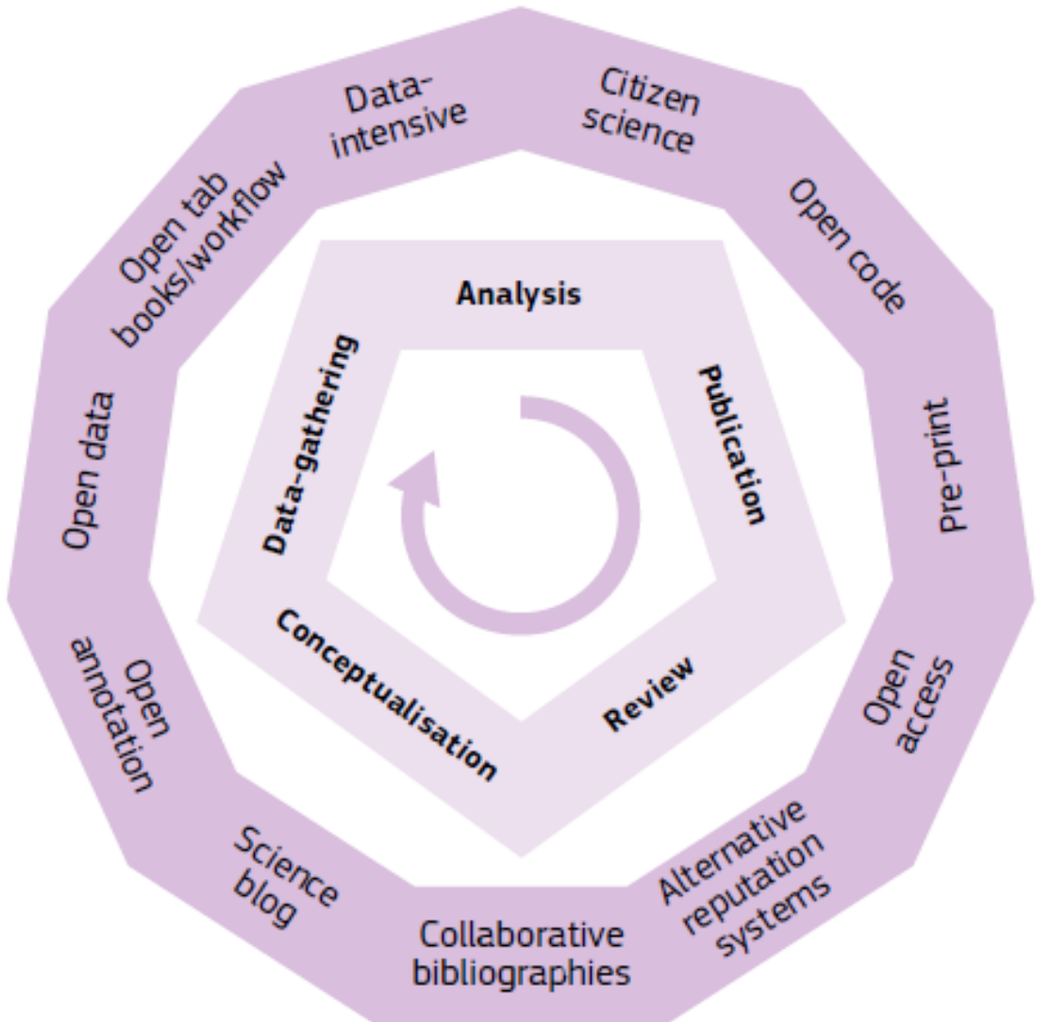
C A U

What is meant by „open science“



- Open Science opens up the entire research enterprise (inner circle) by using a variety of means and digital tools (outer circle)

- From publishing as fast as possible to sharing as fast as possible



EC, 2016, DOI: 10.2777/061652
https://de.wikipedia.org/wiki/Offene_Wissenschaft#/media/Datei:Open_Science_-_Prinzipien.png

1. Publication of gene sequence

download

readme
readme

Assembly
03/06/11 Ion Torrent mapped assembly: Escherichia coli TY-2482 20110606 upload2ncbi.fa.gz
06/06/11 Ion Torrent+Illumina hybrid assembly: Es…erichia coli TY-2482.contig.fa.gz
06/06/11 Ion Torrent+Illumina hybrid assembly (NC…sion): Escherichia coli TY-2482.contig.20110606.fa.gz
11/06/11 Illumina de novo assembly: Escherichia co…2.scaffold.20110610.fa.gz
16/06/11 Gapless Illumina de novo assembly (chrom…herichia coli TY-2482.chromosome.20110616.fa.gz
16/06/11 Gapless Illumina de novo assen…

Raw Data
Illumina reads
11/06/11 Illumina reads: 110601 I238 FCE…
Ion Torrent reads
02/06/11 Ion Torrent run 1: run1.fastq.gz
02/06/11 Ion Torrent run 2: run2.fastq.gz
02/06/11 Ion Torrent run 3: run3.fastq.gz
02/06/11 Ion Torrent run 4: run4.fastq.gz
02/06/11 Ion Torrent run 5: run5.fastq.gz
03/06/11 Ion Torrent run 6: run6.fastq.gz
03/06/11 Ion Torrent run 7: run7.fastq.gz

Additional information
2011vs2001 v2.xls
Specific primers for PCR detection.pdf

History
June 3, 2011: Data released.

In further accordance with our terms of use, please cite this dataset as:
Li, D; Xi, F; Zhao, M; Chen, W; Cao, S; Xu, R; Wang, G; Wang, J; Zhang, Z; Li, Y; Cui, C; Chang, C; Cui, C; Luo, Y; Qin, J; Li, S; Li, J; Peng, Y; Pu, F; Sun, Y; Chen, Y; Zong, Y; Ma, X; Yang, X; Cen, Z; Song, Y; Zhao, X; Chen, F; Yin, X; Rohde, H; Liang, Y; Li, Y and the *Escherichia coli* O104:H4 TY-2482 isolate genome sequencing consortium (2011): Genomic data from *Escherichia coli* O104:H4 isolate TY-2482. BGI Shenzhen. http://dx.doi.org/10.5524/100001

Related manuscript available at:
10.1056/NEJMoa1107643

Accession codes associated with this data:
NCBI Study SRP006916
NCBI BioProject PRJNA67657

EHEC
Escherichia coli bacterium

Pathogens: Genes and Genomes

A heady mix of bacterial pathogenomics, next-generation sequencing, type-III secretion, bioinformatics and evolution!

2. Analysis, discussion in blogs, wikis

You are here: Home / 2011 / June / *EHEC Genome Assembly*

EHEC Genome Assembly

By Nick Loman on June 2, 2011

Keep track of the genomic analysis of the EHEC strains on our Github Wiki.

BGI have released 5 runs of Ion Torrent data for the German EHEC/V…utbreak strain. I hope it is released with no specific restrictions on use for the benefit of the entire community, but th…n't make that entirely clear. Thanks to the BGI for putting it up!

Shall we crowd source some analysis? This come…t as I am currently help organise the Applied Bioinformatics & Public Health conference i… we are discussing the use of whole-genome sequencing in

ehec-outbreak-crowdsourced / BGI-data-analysis
Watch 43 Fork 5
Code | Network | Pull Requests 0 | Issues 2 | Wiki | Graphs
Home | Pages | Wiki History | Git Access

E. coli O104:H4 Genome Analysis Crowdsourcing

Page History

In this wiki we aim to gather all the results of the analysis of the *E. coli* O104:H4 strain responsible for the current outbreak in Germany and Europe.

TEN isolates from the outbreak have been sequenced so far:

- TY2482 (BGI in collaboration with University Medical Centre Hamburg-Eppendorf)
- LB226692 (Life Tech in-house in collaboration with University Hospital Muenster)

3. Scientific publication

We released these data into the public domain under a Creative … a burst of crowd-sourced, curiosity-driven analyses carried out by … continents.[1] … after its disse… release of the … sequences, a … that facilitates … timely informa… lineage.

[11] GitHub. E. coli O104:H4 genome analysis crowd sourcing, 2011. (https://github.com/ehec-outbreak-crowdsourced/BGI-data-analysis/wiki.)

The NEW ENGLAND JOURNAL of MEDICINE

BRIEF REPORT

Open-Source Genomic Analysis of Shiga-Toxin–Producing *E. coli* O104:H4

…D., Junjie Qin, Ph.D., Yujun Cui, Ph.D., Dongfang Li, M.E., …n, M.B., B.S., Moritz Hentschke, M.D., Wentong Chen, B.S., …, Yangqing Peng, B.S., Junhua Li, B.E., Feng Xi, B.E., Shenghui Li, B.S., Yin Li, B.S., Zhaoxi Zhang, B.S., Xianwei Yang, B.S., Meiru Zhao, M.S., Peng Wang, B.M., Yuanlin Guan, B.E., Zhong Cen, M.E., Xiangna Zhao, B.S., Martin Christner, M.D., Robin Kobbe, M.D., Sebastian Loos, M.D., Jun Oh, M.D., Liang Yang, Ph.D., Antoine Danchin, Ph.D., George F. Gao, Ph.D., Yajun Song, Ph.D., Yingrui Li, B.S., Huanming Yang, Ph.D., Jian Wang, Ph.D., Jianguo Xu, M.D., Ph.D., Mark J. Pallen, M.D., Ph.D., Jun Wang, Ph.D., Martin Aepfelbacher, M.D., Ruifu Yang, M.D., Ph.D., and the *E. coli* O104:H4 Genome Analysis Crowd-Sourcing Consortium*

SUMMARY

An outbreak caused by Shiga-toxin–producing *Escherichia coli* O104:H4 occurred in Germany in May and June of 2011, with more than 3000 persons infected. Here, we report a cluster of cases associated with a single family and describe an open-source

Ways of doing open science: success stories

GitHub This repository https://github.com/leereilly/swot Explore Features Enterprise Pricing Sign up Sign in

leereilly / swot Watch 22 Star 394 Fork 2,206

Code Issues 22 Pull requests 90 Pulse Graphs

Identify email addresses or domains names that belong to colleges or universities. Help automate the process of approving or rejecting academic discounts.

1,473 commits 2 branches 26 releases 378 contributors

Branch: master New pull request New file Find file HTTPS https://github.com/leerei Download ZIP

ReidWeb Merge pull request #988 from janyksteenbeek/master ... Latest commit dcad913 6 days ago

lib	Merge pull request #988 from janyksteenbeek/master	6 days ago
test	Add more blacklisted domains	3 months ago
.document	Initial commit to email.	4 years ago

Ways of doing open science: success stories

- Early feedback: pregistration of studies

Call for Papers

Technology and Human Behavior

A Preregistered Special Issue of the Journal of Media Psychology

Guest Editors: Malte Elson (Ruhr U Bochum) and Andrew K. Przybylski (U of Oxford)
Editor-in-Chief: Nicole Krämer (U of Duisburg-Essen)

Psychological research practice has undergone some remarkable changes in recent years. The discipline's attention towards issues of theory robustness, replicability/reproducibility of empirical findings, and open science practices has substantially increased.Media psychologyis part of this improvement as 2015 marks the beginning of a new phase. For the first time, media psychologists are provided with the opportunity to submit and publish pre-registered research reports at the Journal of Media Psychology. Following the preregistered report format, peer review happens in two steps: First, the theoretical background, hypotheses, methods, and analysis plans of a study are peer-reviewed before the data are collected. If the theoretical derivation of hypotheses as well as methods are evaluated as sound, the study receives an "in-principle" acceptance, and researchers can proceed to conduct it (taking potential changes or additions suggested by the reviewers into consideration). By doing so, the collected data can be used as a true hypothesis test. In a second step, the soundness of the analyses and discussion section are reviewed, but the publication decision is not contingent on the outcome of the study. This way of conducting peer review is special because it draws a clear line between work that is "exploratory" in natureand that which is "confirmatory" empirical work that employs these practices. In order to allow many media psychologists to participate, the Special Issue presents a broad theme: We are inviting original research proposals aiming to investigate how the selection, use, and experience of technology affect human behavior. Proposals can include additional hypotheses pertaining to effects on cognitions, emotions, and motivations, but primary interest should be to explain behavioral characteristics. In keeping with the journal's domains of inquiry, we are particularly interested in research on entertainment media and social media, but also welcome studies on human-computer interaction, e-learning, virtual environments, advertising, or other areas of media psychology. We also encourage submissions describing replications of existing research, as those naturally allow a clear, motivated rationale for preregistration. In addition, we will encourage submissions employing alternative types of analysis plans, such as Bayesian statistics (which require quantification of beliefs prior to data collection). The Special Issue is also dedicated to promote open science practices, requiring all authors to share all data and materials with the reviewers during the two evaluation stages, and with the wider community upon publication (as publicly available supplementary materials on the journal's website).

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

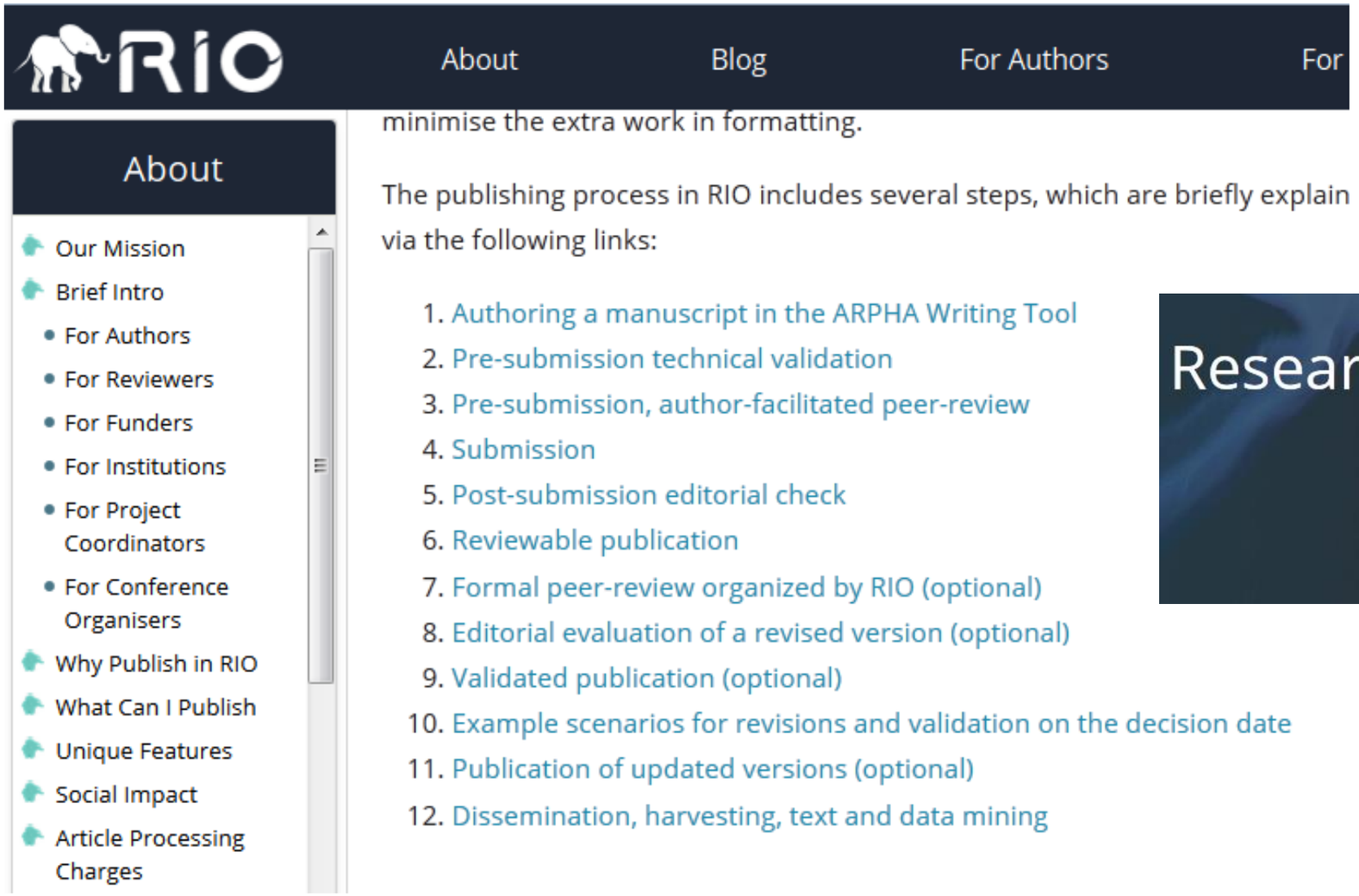
https://www.icahdq.org/pubs/calls/JournalOfMedia.asp

Leibniz-Gemeinschaft

CAU

Ways of doing open science: success stories

- Early feedback: non-traditional publications



http://riojournal.com/

Our Mission

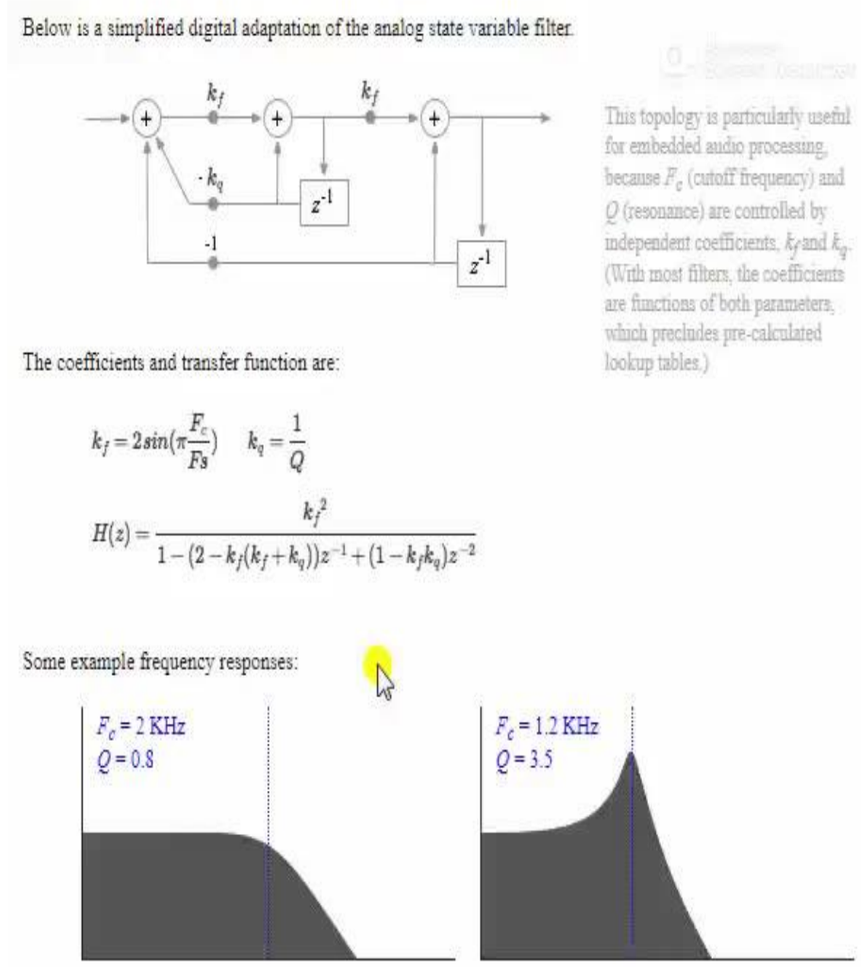
Research Ideas and Outcomes (RIO) aims to catalyse change in research communication by publishing ideas, proposals and outcomes in a comprehensive way. By doing so, we hope to increase transparency, trust and efficiency of the whole research ecosystem.

Ways of doing open science: success stories

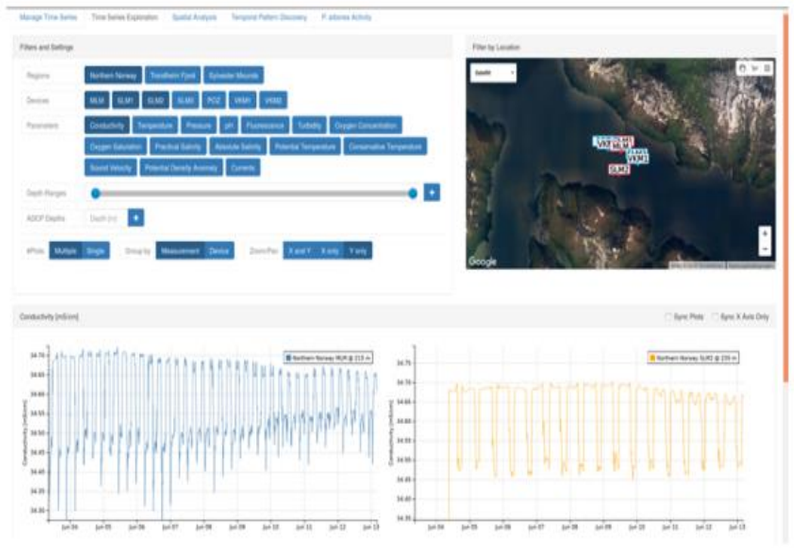
- Early feedback: non-traditional publications + processes

Ways of doing open science: success stories

- ‚computational essays‘ (Stephen Wolfram)
- allow for participation/ engagement
- comprehensible
- interactive



OceanTEA - interactive exploration oceanographic datasets



Why open science?

- Science system vs. individual researcher
- Quality
- Efficiency
- Reproducibility
- Credibility
- Visibility

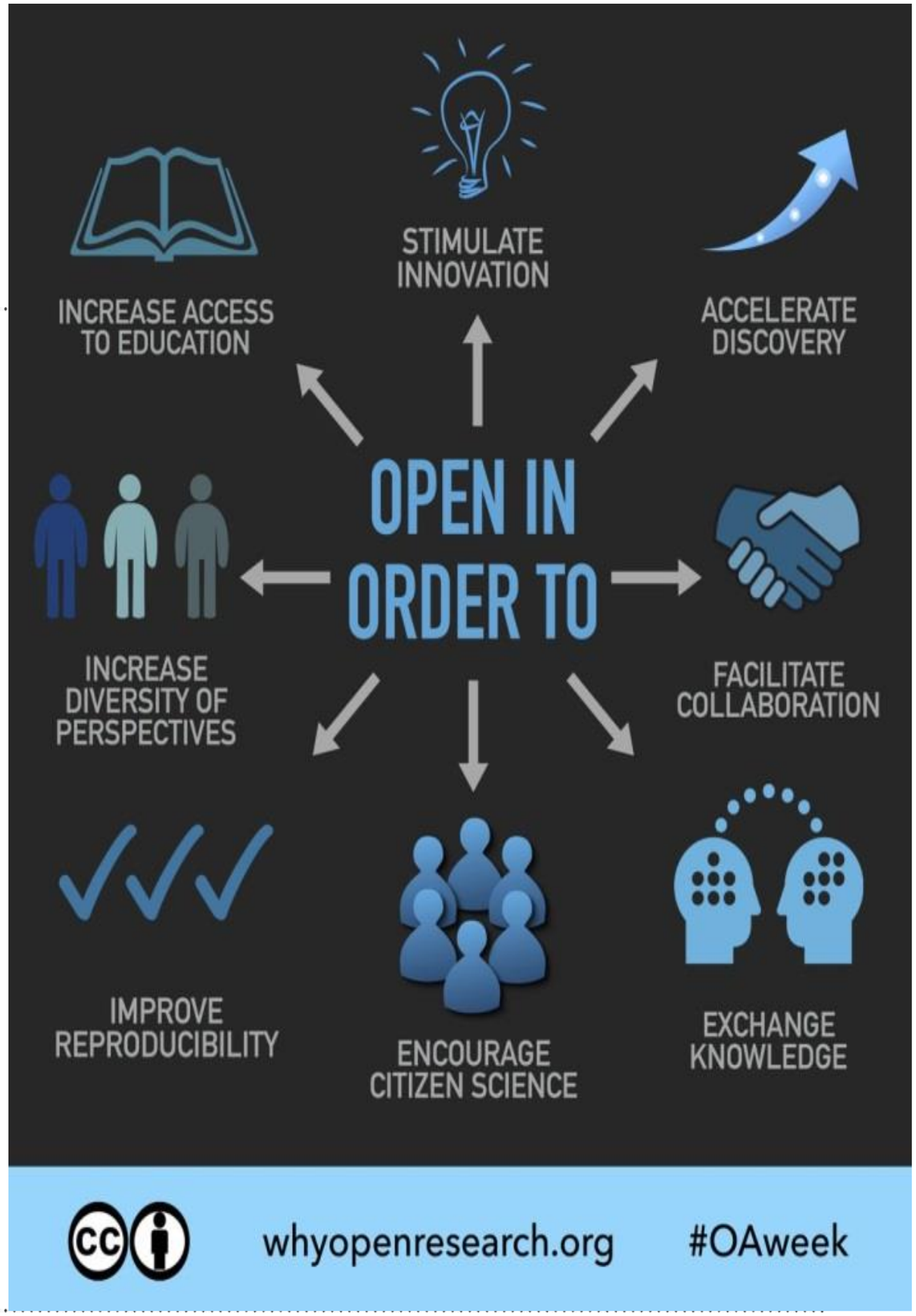
ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

EC, 2016, DOI: 10.2777/061652

http://whyopenresearch.org

Why open science?

- “Open science is about improving the quality, accountability and social contribution of research…” (p. 96)



McKiernan, Erin (2017): OA Week graphics, "Open in order to...". figshare. Fileset. https://doi.org/10.6084/m9.figshare.5527900.v3
MLE on Open Science: Final Report -Altmetrics and Rewards (2018). https://rio.jrc.ec.europa.eu/en/library/mle-open-science-final-report-altmetrics-and-rewards

Why open science?

Lizzie Gadd:

- Openness leads to QUALITY by enhancing rigour and reproducibility as evidenced by registered reports, open methods, FAIR data and diversity of both those undertaking the research, and research participants.
- Openness leads to VISIBILITY by making the whole research lifecycle more transparent and accessible
- Openness leads to IMPACT through improved engagement with openly accessible and understandable outputs, and with the communities on which our research impacts.

Why open science?

- Visibility and citations

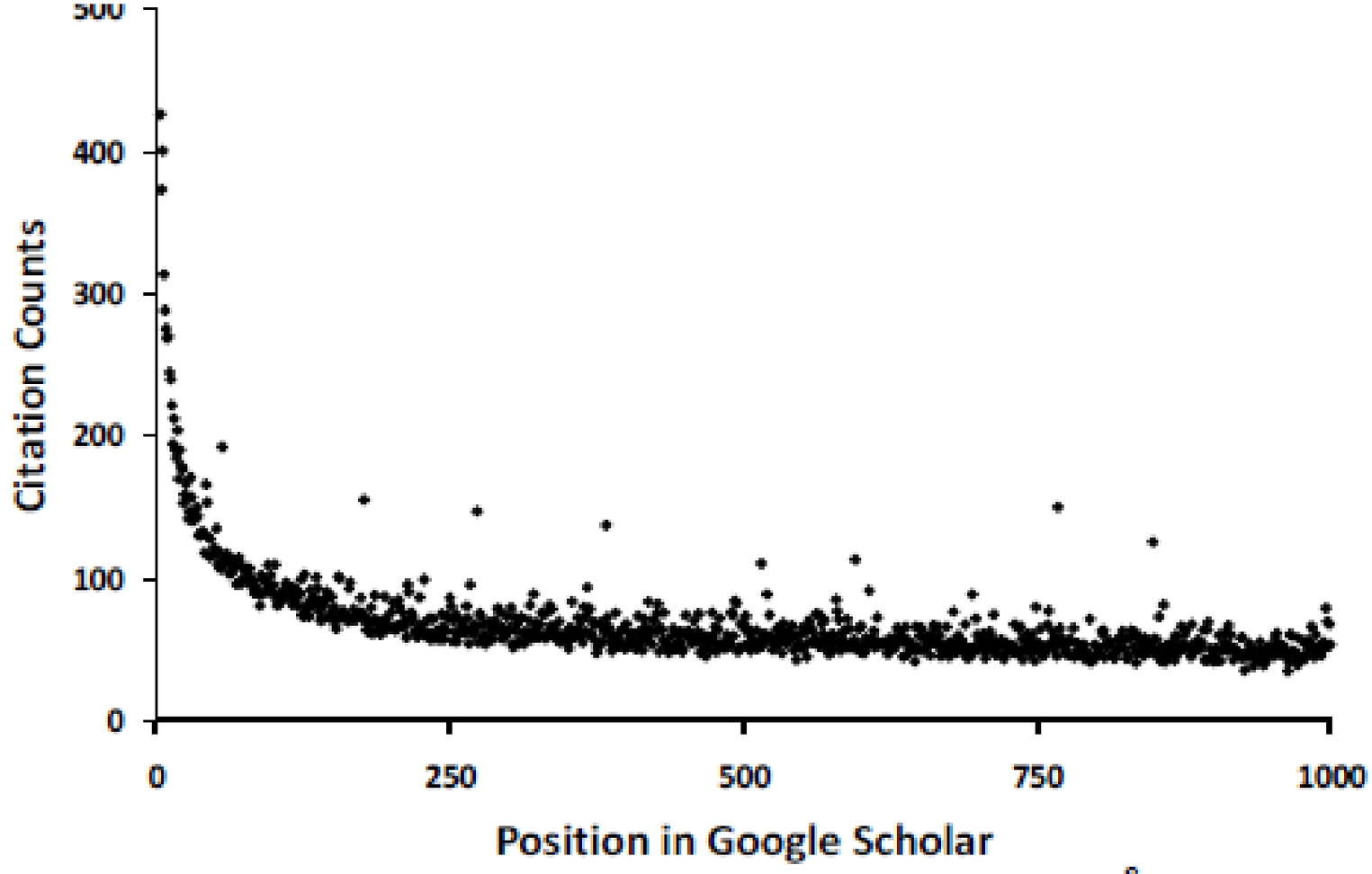


Figure 3: Mean Citation Count per Position[8]

Beel, J., Gipp, B., & Wilde, E. (2010). Academic Search Engine Optimization (ASEO): Optimizing Scholarly Literature for Google Scholar and Co. Journal of Scholarly Publishing, 41(2), 176–190. doi: 10.3138/jsp.41.2.176

Why open science?

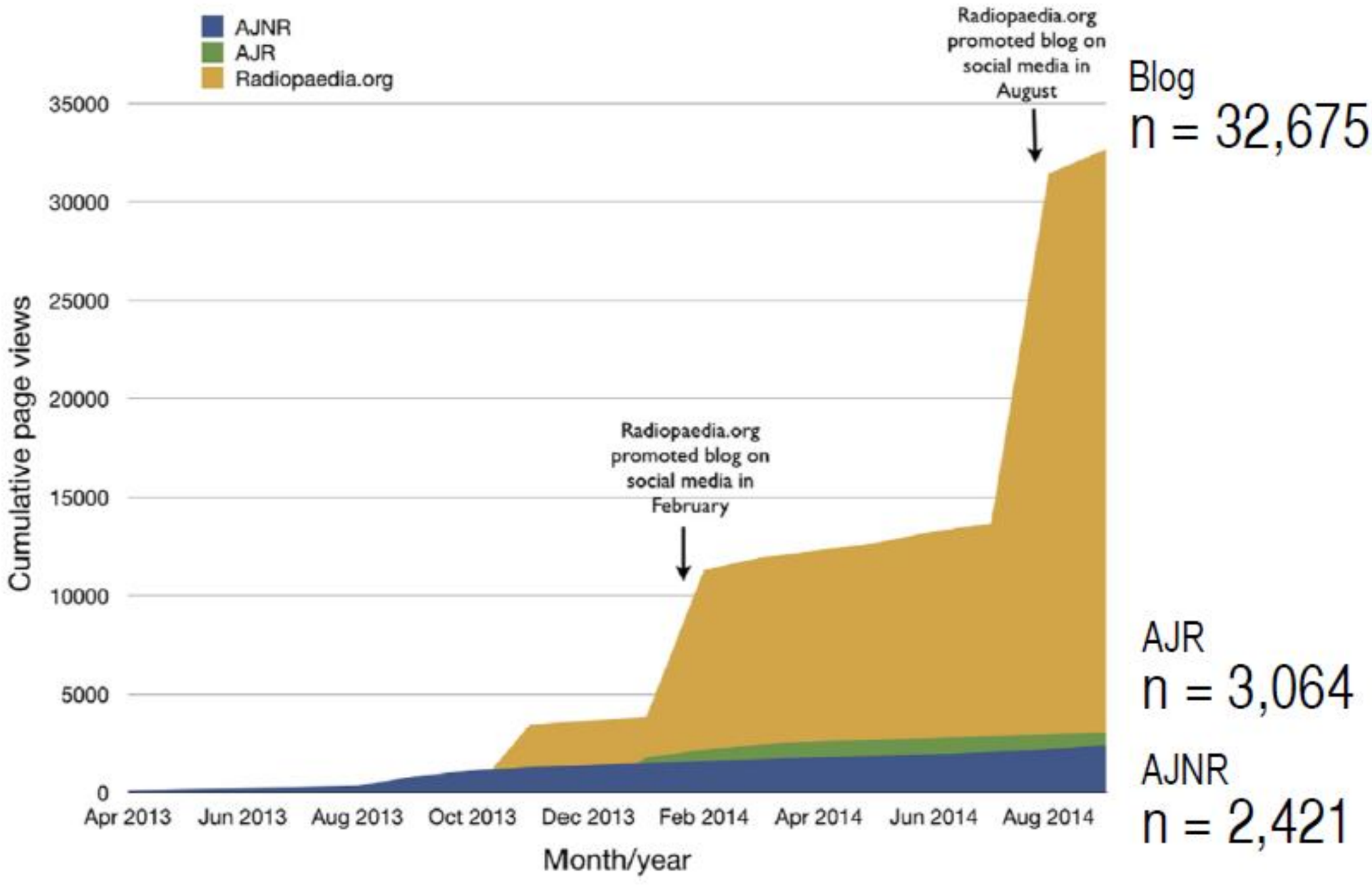
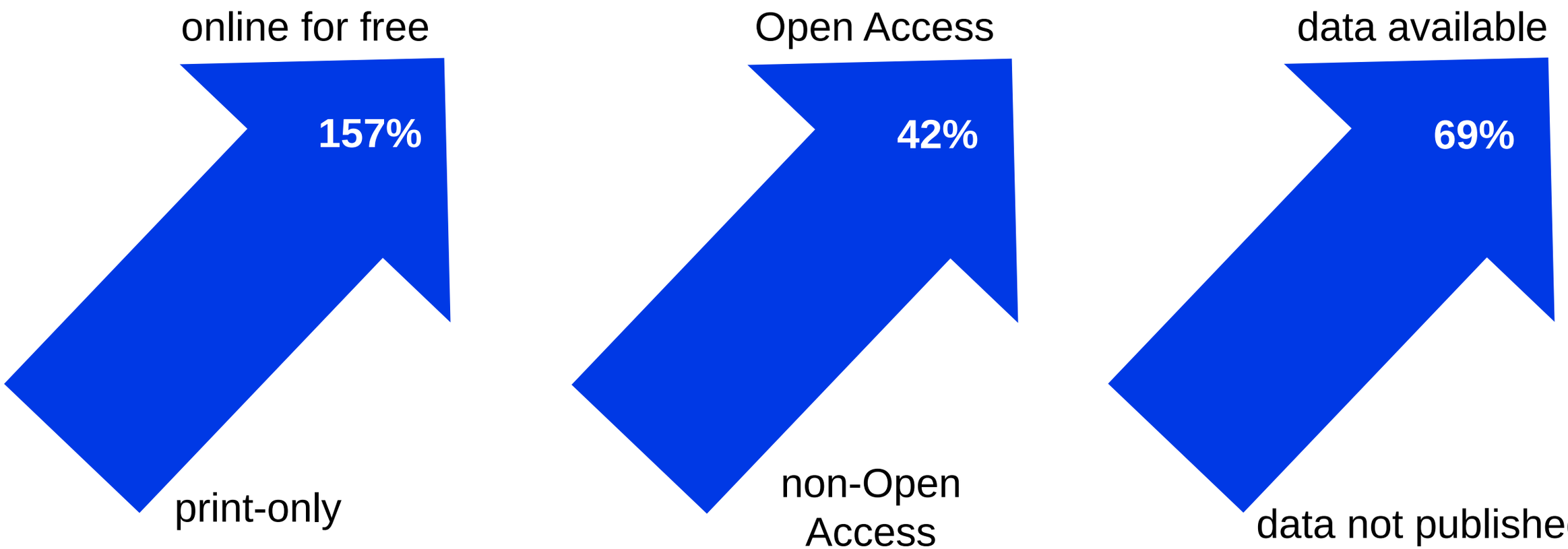


Fig 1. Cumulative monthly page views for the research articles in the *American Journal of Neuroradiology* (*AJNR*) and the *American Journal of Roentgenology* (*AJR*) and the blog article on Radiopaedia.org.

Hoang, J. K., McCall, J., Dixon, A. F., Fitzgerald, R. T., & Gaillard, F. (2015). Using Social Media to Share Your Radiology Research: How Effective Is a Blog Post? Journal of the American College of Radiology, 12(7), 760-765. doi: 10.1016/j.jacr.2015.03.048

Why open science?

- Visibility and citations



Lawrence, S. (2001). Free online availability substantially increases a paper's impact. Nature, 411(6837), 521
Piwowar, H. A., Day, R. S., & Fridsma, D. B. (2007). Sharing detailed research data is associated with increased citation rate. PLoS ONE, 2(3): e308.
Eysenbach, G. (2006). Citation advantage of open access articles. PLoS Biol 4:e157.

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics
Leibniz-Gemeinschaft

Why open science?

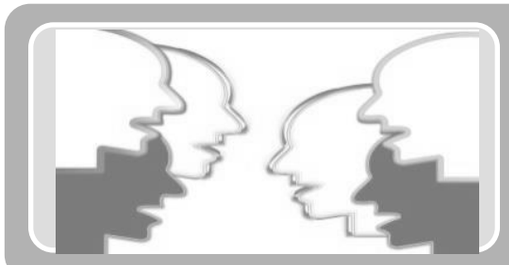
Open Access Advantage

Links:

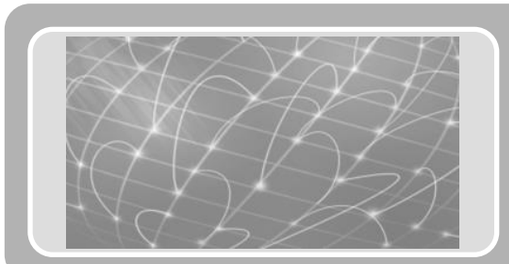
- sparceurope.org/oaca
- sparceurope.org/oaca_table (summary)

Ways of doing open science

3 bundles of functions in online environments



Interact and Exchange



Publish and Distribute

Present Oneself

Increases visibility

See also: https://osl.tib.eu/w/Handbuch_CoScience/Online-Profil-_und_Netzwerk-Dienste

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Leibniz-Gemeinschaft

C A U

Why (not) open science?

- Blogging is a **waste of precious time** that could be spent on “legitimate” publishing

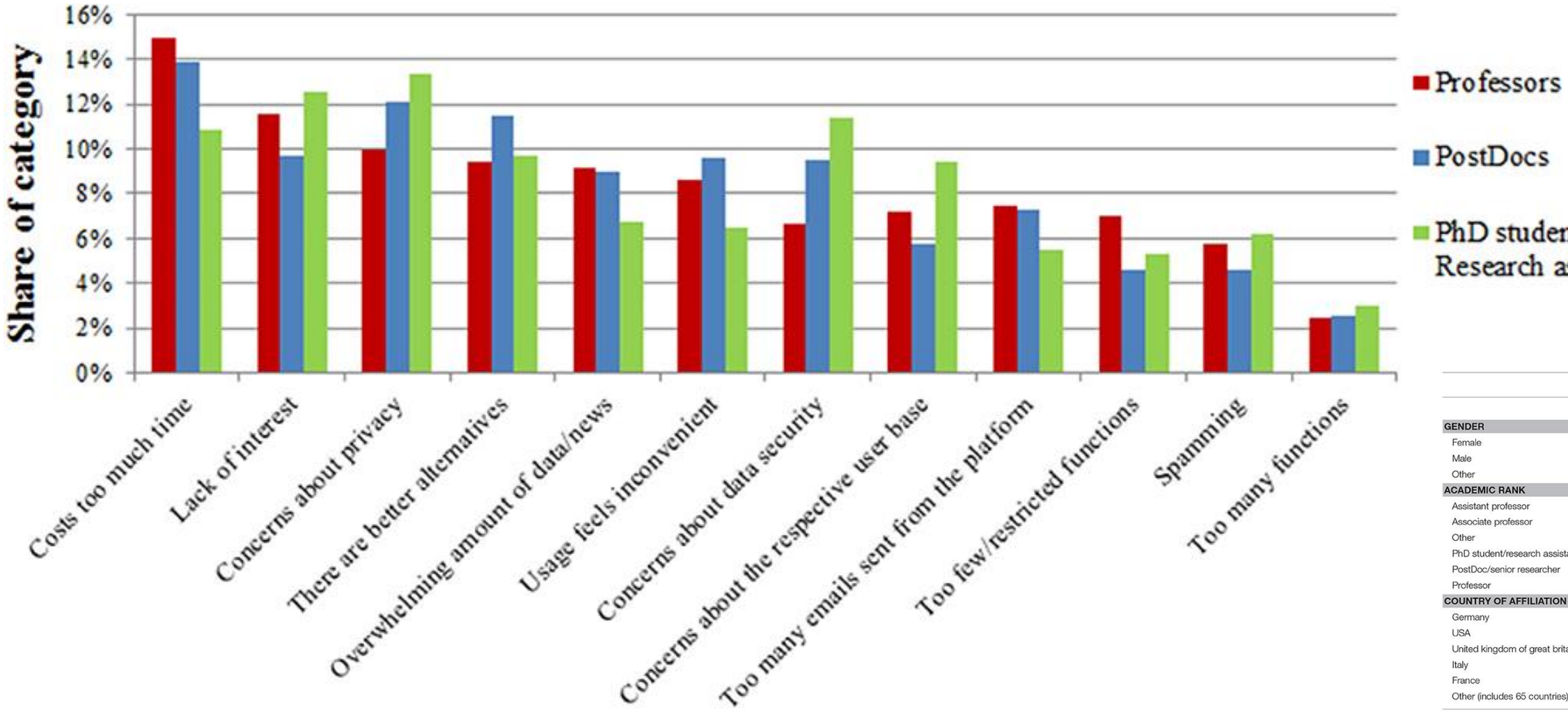
- Because it’s a form of self-publishing that lacks peer review, blogging **isn’t** usually viewed as a **legitimate form of scholarship**

- Dismissal of my work because it’s online [and] criticisms that my work **isn’t good enough** to be published anywhere else.

- Sometimes blogging is even seen as disseminating one’s ideas too freely. In a competitive academic field, research ideas could be “**scooped**” from a blog, while established journals may not want to publish work that’s available in some form online.

Why (not) open science?

- Survey and interviews
- R1: “When you use social media you’re not working”



	Survey A	Survey B
	N = 872 (%)	*N* = 948 (%)
GENDER		
Female	31.1	31.6
Male	68.7	68.3
Other	0.1	0.1
ACADEMIC RANK		
Assistant professor	12.3	12.6
Associate professor	16.8	16.6
Other	11.3	7.1
PhD student/research assistant	16.6	14.7
PostDoc/senior researcher	15.2	19.8
Professor	27.9	29.1
COUNTRY OF AFFILIATION (TOP 5 + OTHER)		
Germany	27.7	33.4
USA	14.0	14.7
United kingdom of great britain	5.8	5.9
Italy	5.8	5.3
France	3.7	3.5
Other (includes 65 countries)	43.0	37.3

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

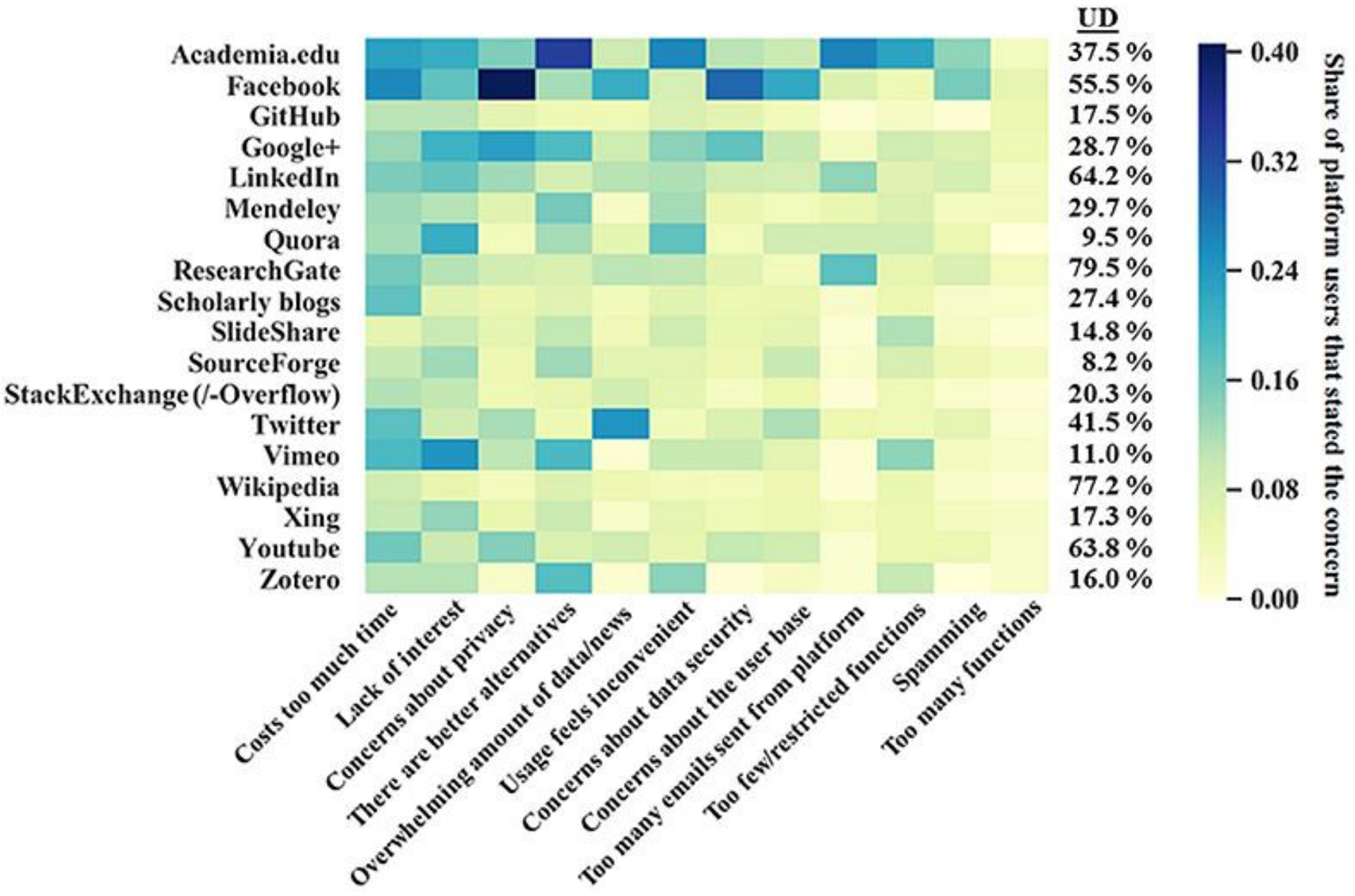
Lemke, S., Mehrazar, M., Mazarakis, A., & Peters, I. (2019). “When You Use Social Media You Are Not Working”: Barriers for the Use of Metrics in Social Sciences. Frontiers in Research Metrics and Analytics, 3, 39. https://www.frontiersin.org/articles/10.3389/frma.2018.00039/full

Leibniz-Gemeinschaft

C A U

Why (not) open science?

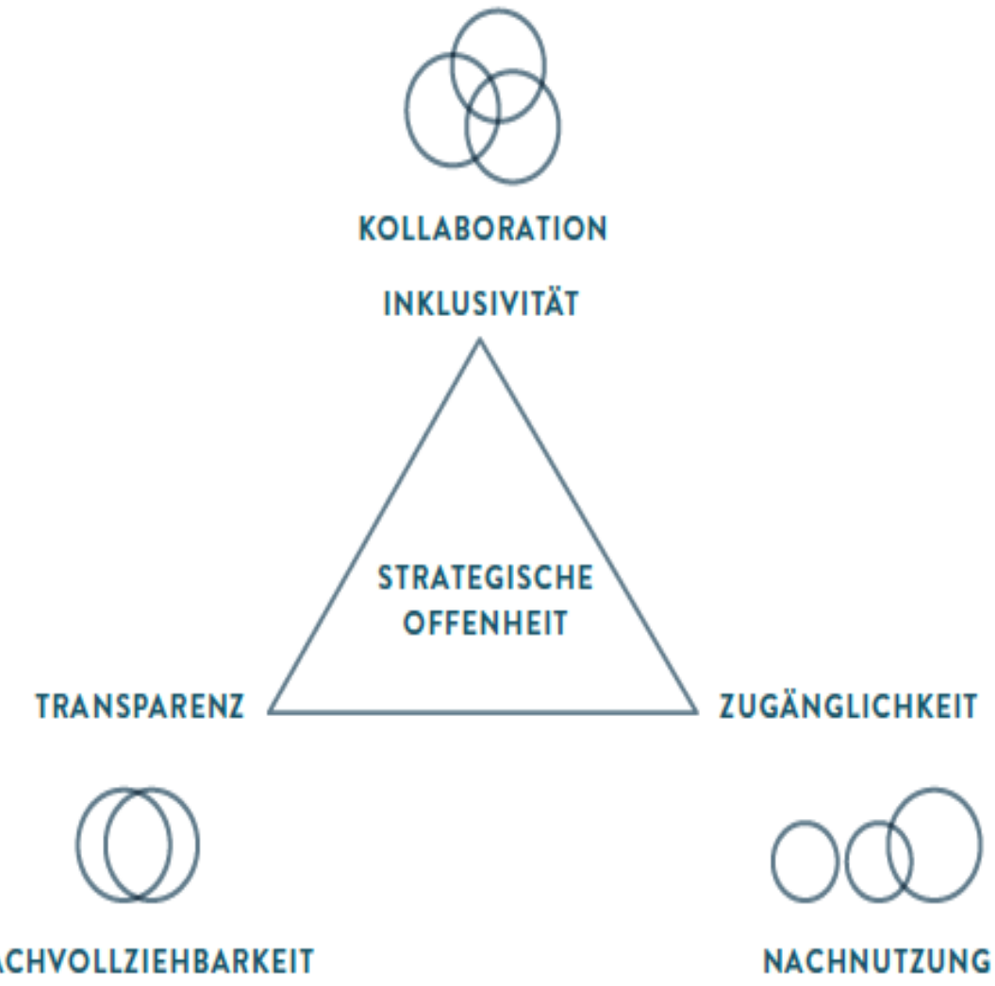
- Survey



Lemke, S., Mehrazar, M., Mazarakis, A., & Peters, I. (2019). “When You Use Social Media You Are Not Working”: Barriers for the Use of Metrics in Social Sciences. Frontiers in Research Metrics and Analytics, 3, 39. https://www.frontiersin.org/articles/10.3389/frma.2018.00039/full

Which environment for open science?

- Digital + open = two sides of the same coin



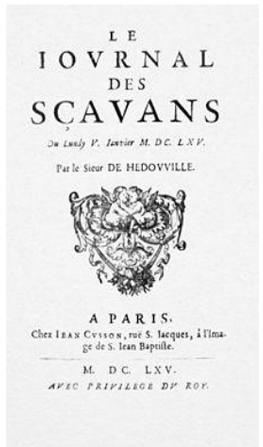
Fecher, B., Leimüller, G., & Blümel, C. (2018). Das Potenzial strategischer Öffnung. Stifterverband. DOI: 10.5281/zenodo.1246652

Which environment for open science?

Status Quo

- static content

- paper and PDF-publications



- discipline-specific data silos

- only limited repeatability, reusability, reproducibility and interactivity

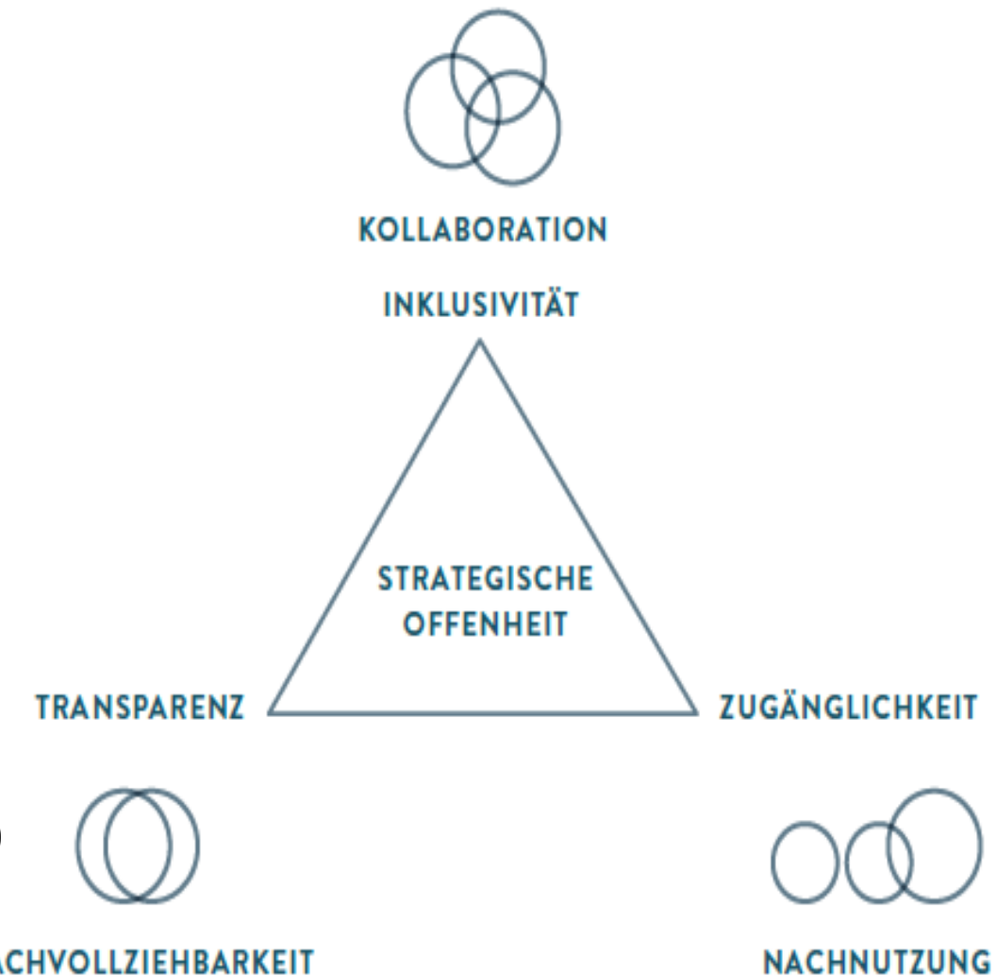
Which environment for open science?

- Digital + open = two sides of the same coin

- Share as early as is practicable in the discovery process (Nielsen, 2011)

- Stakeholder engagement

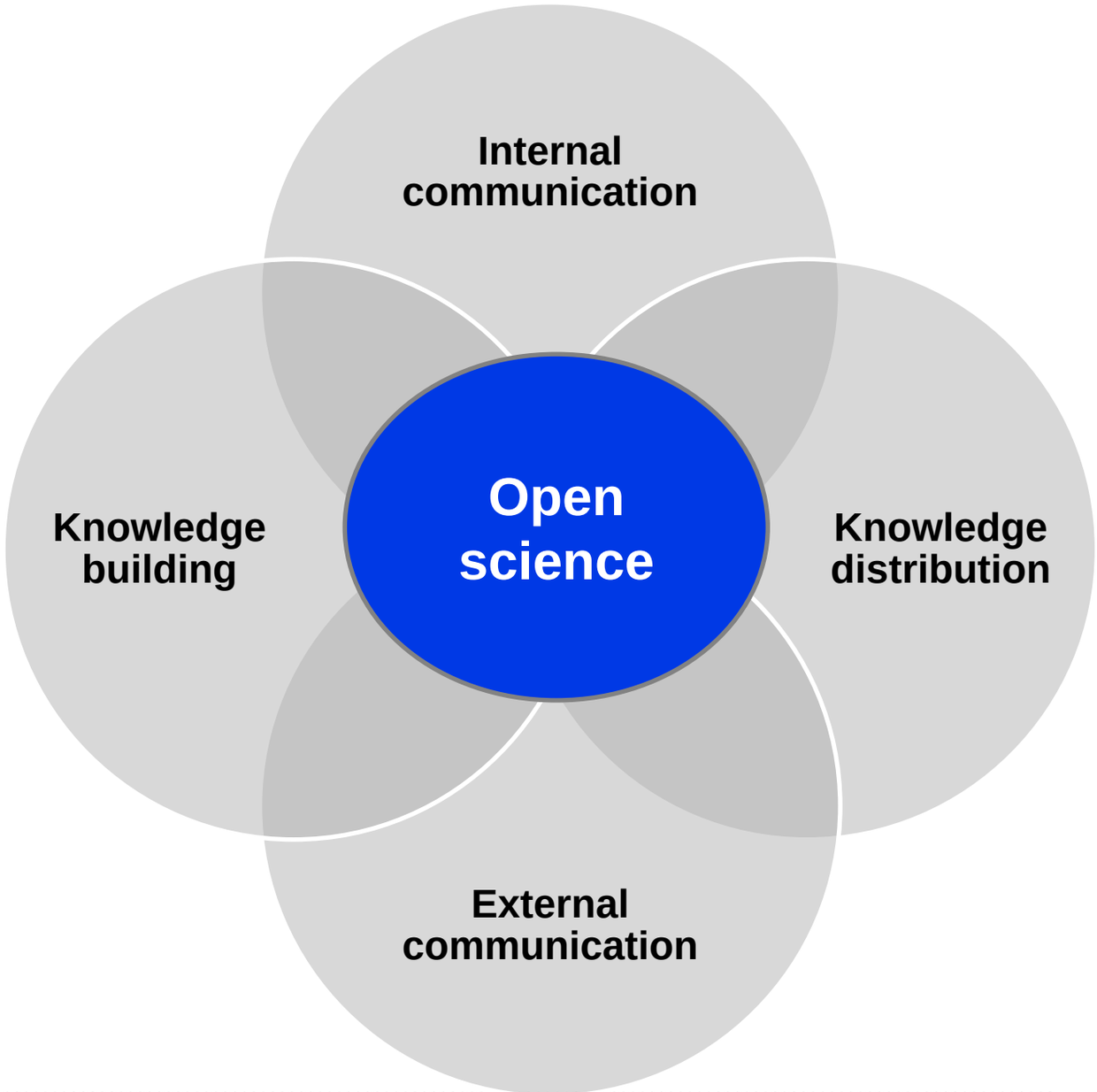
- Needs FAIR (findable, accessible, interoperable, reusable) infrastructures



Fecher, B., Leimüller, G., & Blümel, C. (2018). Das Potenzial strategischer Öffnung. Stifterverband. DOI: 10.5281/zenodo.1246652
Nielsen (2011). http://openscience.org/an-informal-definition-of-openscience/

Why open science?

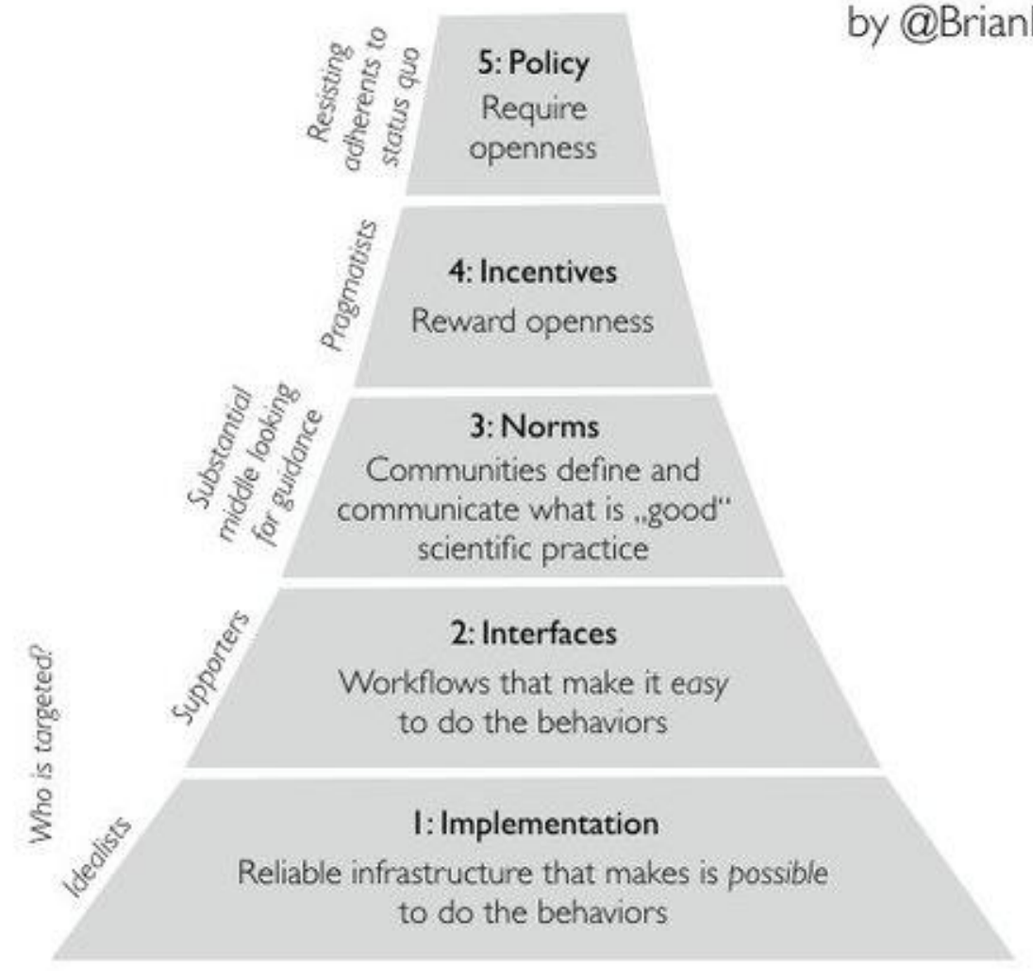
- Interact and exchange
- Publish and distribute
- Present oneself
- …



Which environment for open science?

How to achieve a cultural change towards open science

Based on a tweet storm by @BrianNosek



Which environment for open science?

- University Medical Center Utrecht, Netherlands

- Goals of policies
  - Individual researchers should be judged on their actual contributions, not number of publications
  - Research geared towards creating societal impact, not just scientific excellence
  - Goodhart‘s Law: the goal is the measure
    - „We shape the structures that shape science“.

- Portfolio approach
  - Plans of the researcher
  - Managerial responsibilities and academic duties
  - Teaching and supervision
  - Clinical work and research
  - Entrepreneurship and community outreach

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.

REDEFINE EXCELLENCE

Fix incentives to fix science

Rinze Benedictus and Frank Miedema

An obsession with metrics pervades science. Our institution, the University Medical Center Utrecht in the Netherlands, is not exempt. On our website, we proudly declare that we publish about 2,500 peer-reviewed scientific publications per year, with higher than average citation rates.

A few years ago, an evaluation committee spent hours discussing which of several faculty members to promote, only to settle on the two who had already been awarded particularly prestigious grants. Meanwhile, faculty members who spent time crafting policy advice had a hard time explaining how this added to their scientific output, even when it affected clinical decisions across the country.

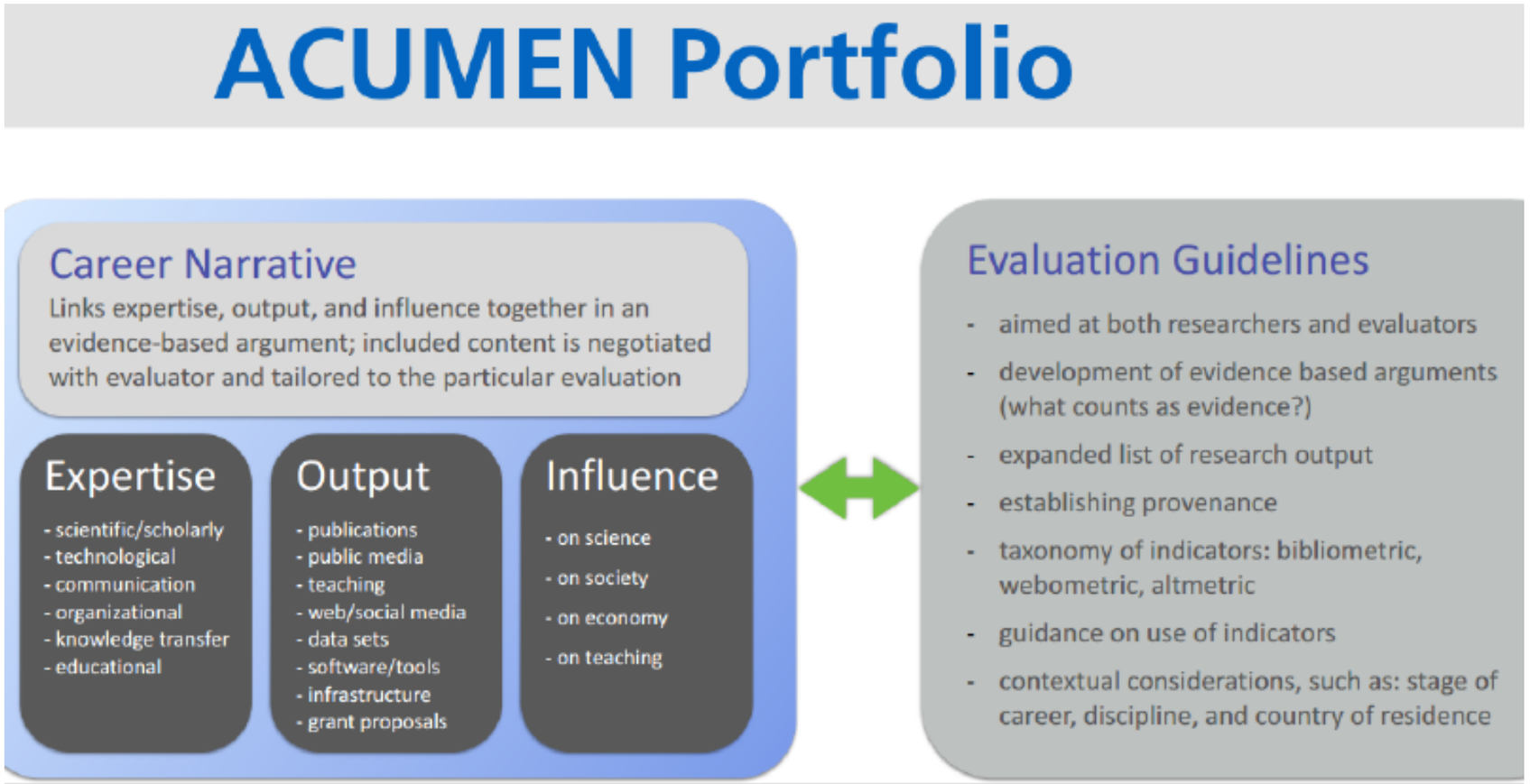
Publications that directly influenced patient care were weighted no higher in evaluations than any other paper, and ▸

27 OCTOBER 2016 | VOL 538 | NATURE | 453

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

doi:10.1038/538453a

Which environment for open science?



- Multidimensionality
- Data transparency
- Informed peer review
- Indicators as 'proxies' for career narratives

CWTS

14

Which environment for open science?

Open Science Career Evaluation Matrix (OS-CAM)

- Areas to be considered
  - Research output
  - Research process
  - Service and leadership
  - Teaching and supervision
  - Professional experience

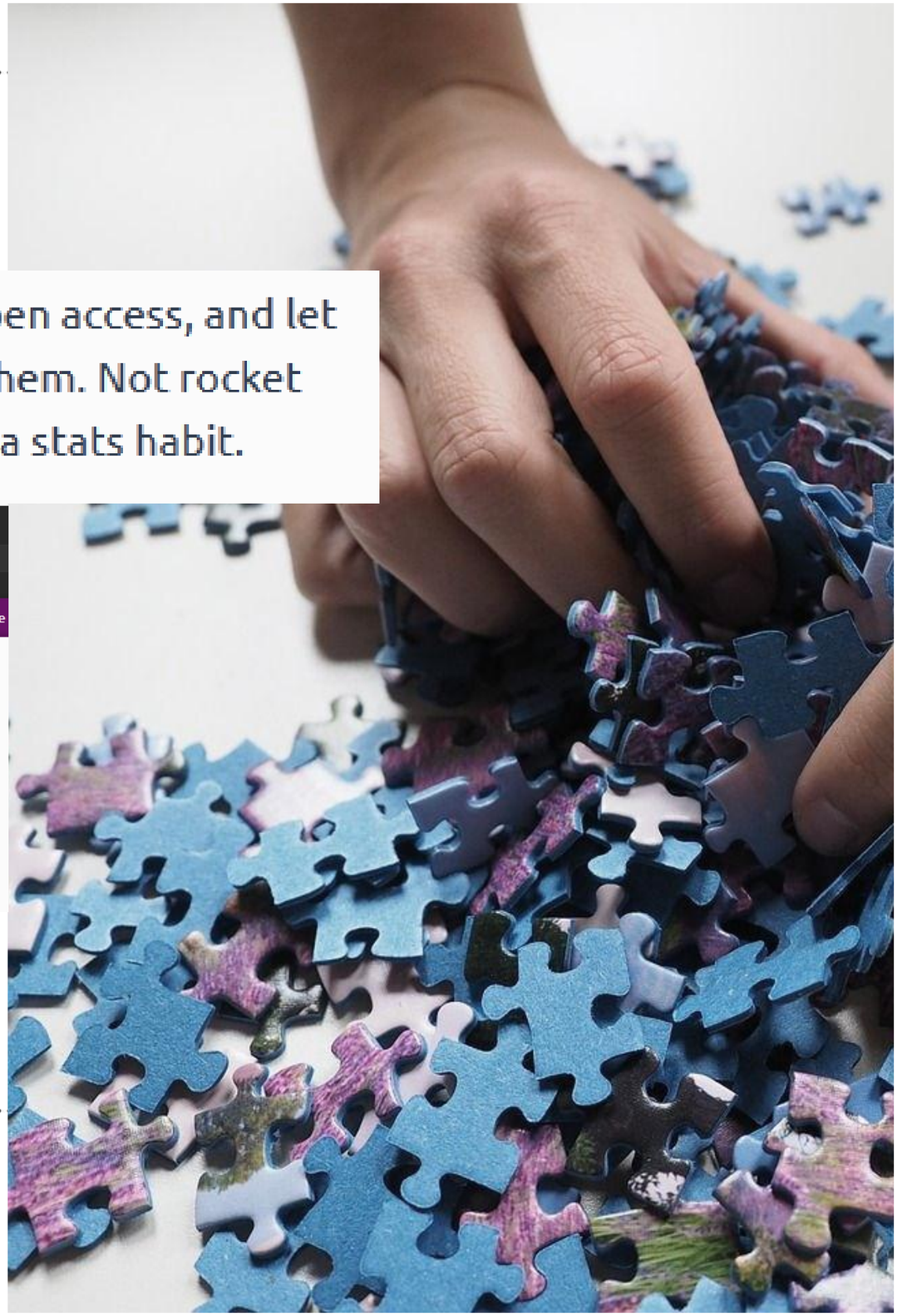
Open Science Career Assessment Matrix (OS-CAM)	
Open Science activities	*Possible evaluation criteria*
RESEARCH OUTPUT	
Research activity	Pushing forward the boundaries of open science as a research topic
Publications	Publishing in open access journals Self-archiving in open access repositories
Datasets and research results	Using the FAIR data principles Adopting quality standards in open data management and open datasets Making use of open data from other researchers
Open source	Using open source software and other open tools Developing new software and tools that are open to other users
Funding	Securing funding for open science activities
RESEARCH PROCESS	
Stakeholder engagement / citizen science	Actively engaging society and research users in the research process Sharing provisional research results with stakeholders through open platforms (e.g. Arxiv, Figshare) Involving stakeholders in peer review processes
Collaboration and Interdisciplinarity	Widening participation in research through open collaborative projects Engaging in team science through diverse cross-disciplinary teams
Research integrity	Being aware of the ethical and legal issues relating to data sharing, confidentiality, attribution and environmental impact of open science activities Fully recognizing the contribution of others in research projects, including collaborators, co-authors, citizens, open data providers
Risk management	Taking account of the risks involved in open science
SERVICE AND LEADERSHIP	
Leadership	Developing a vision and strategy on how to integrate OS practices in the normal practice of doing research Driving policy and practice in open science Being a role model in practicing open science
Academic standing	Developing an international or national profile for open science activities

O'Carroll, C., Rentier, B., Cabello Valdès, C., Esposito, F., Kaunismaa, E., Maas, K., ... & Lossau, N. (2017). Evaluation of Research Careers fully acknowledging Open Science Practices-Rewards, incentives and/or recognition for researchers practicing Open Science. Publication Office of the European Union.
http://ec.europa.eu/research/openscience/pdf/os_rewards_wgreport_final.pdf

Open science….

…is what you make of it!

Ergo, if you want people to read your papers, make them open access, and let the community know (via blogs, twitter, etc.) where to get them. Not rocket science. But worth spending time doing. Just don't develop a stats habit.



ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Thank you!

Professor Dr. Isabella Peters
ZBW – Leibniz Information Center for Economics & Kiel University
i.peters@zbw.eu

Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Leibniz-Gemeinschaft

C A U