Transitioning "Open Data" from a NOUN to a VERB

Varma D Aadi Narayana[1] *   Sheevendra Sharma[2]

1, Head-Product & Innovation at Profeza. Faridabad, Haryana, India 121003

2, Head-Strategy & Market outreach at Profeza. Faridabad, Haryana, India 121003

* Corresponding author: [aadi@profeza.com]

Abstract

Research assessment is the process or a metric which aims to assess the impact of the research study. The assessment may include the process that aims in evaluating the quality or intellect of a researcher given the notion that qualified scientists are more productive and may drive quality research in the process of Scholarly Communications. Over time, we have become used to equating the quality of the research with the quality or performance of the researcher. The emphasis over publications may encourage unethical practices, Which may be extrapolated to the evolution of problems like, Irreproducibility, Scientific fraud.

Over the past century, a myriad of activities has been undertaken or are still being taken to improve the ways by which research can be assessed. Beginning with the first evolution of the Impact Factor and more recently other, Citation metrics, Altmetrics, etc. have resulted from this work. In this article, we discuss around the myriad of strategies that may play a significant role in the cultural transition of Science and Scientists that is still ongoing. And also highlight the reasons why we should not only look at research assessment but should also be keen on researcher evaluations and differentiate them from one another. Reflecting this, the title of the article signals how the strategies that researchers may need to consider might impact the way they interact with the Open Data movement.

Introduction

There are multiple problems faced by scholarly ecosystem and swift momentum towards open data is a drive to curb the problems. Importantly the bigger one is Irreproducibility. A daunting report has appeared in the past which states that irreproducible research may costs upto 28bn$ per a year to governments. This may stand a reason which may have resulted in cutting the funds for research grants from the last few years. Primary objective driving open-data momentum is to enhance reproducibility and reusability of scientific research. Scholarly studies (discovery) is closely related to innovation. Higgs found his boson in a research lab at an academic university. A multitude of great discoveries happened in Academic institutes which then translated into Corporates benefitting the society. For instance, Louis Pasteur is serving at

the University of Strasburg when he discovered Vaccine.   However, even though, academic research is the birthplace of innovation, it appears that cultural shift within academia moves at a far slower pace. In this opinion article, we discuss the reasons that may slow the pace of cultural change still further and discuss some key strategies which may be applied to accelerate the process.
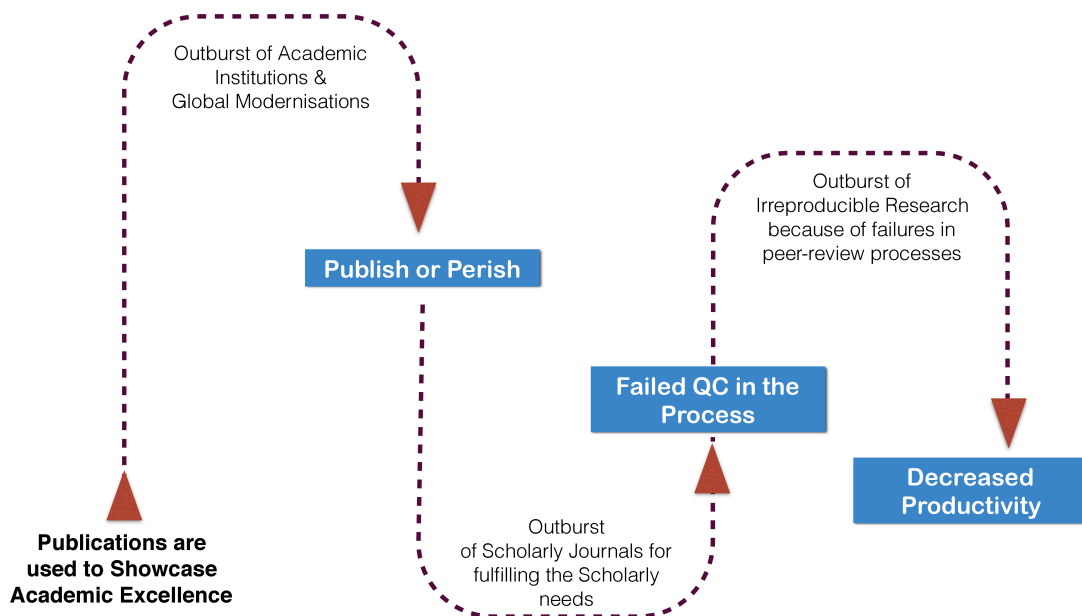
Evolution of Problems in Scientific Research

Fig:1 *Evolution of Problems in Scientific Research*

All these events might have occurred in during the years 1920-2005.

It all started when researchers began to equate academic excellence with scientific publications, mainly appraising the researchers who were able to publish in high impact journals (that is, journals with high impact factors [9]). During the same period, there was an outburst of new academic institutions due to global modernization, increased population and the realization that academic research improves competitiveness. This outburst, in turn, has led to need to publish, to sustain the tenure-track positions. Giving rise to a problem of peer-pressure in the scholarly ecosystem known as 'Publish or Perish' and might also decrease research quality or productivity.

Researchers have a habit of swiftly addressing a problem by seeking a solution. And they concluded that Impact Factors is the primary culprit for this peer-pressure and hence started

initiatives to curb the use of Impact Factors in academics. San Fransisco Declaration on Research Assessment (DORA) is an outcome of this process which encourages the community to not rely or depend on Journal Impact factors to assess researchers performance.

Simultaneously there is an outburst of Scholarly journals in the ecosystem to fulfill the needs of scholars to publish happened, which then worsened the QC or Peer-review process which resulted in polluting the scientific literature with fabricated or Irreproducible science [ref].

The myriad of events during the course, importantly

Article-level Metrics: These are quantifiable measures that document many ways in which both scientists (commonly through citations) and the general public engage (through social media channels) with published research. Traditional metrics like citations and journal impact factor captures a narrow view of the work's value and do so only after the accumulation of citations in academic literature(7).

Altmetrics: In scholarly and scientific publishing, **altmetrics** are non-traditional metrics proposed as an alternative to more traditional citation impact metrics, such as impact factor and h-index. They are scholarly impact measures, based on activity in web-based environments(10). There is still an ongoing debate in the community whether altmetrics be used in quantifying the quality of research articles.

Open peer-review: Any scholarly review mechanism which was providing disclosure of author and referee identities to one another at any point during the peer review or publication process, and this can also be a continuous process after publication. It aims to provide the scholarly community an insight into author/referee conversations during the review process. Surfacing these conversations provides readers an expanded contextual discussion of the subject at hand, and enriches science communication for all stakeholders(11,12).

Pre-prints: it is a version of a scholarly or scientific paper that precedes publication in a peer-reviewed academic or scientific journal. The preprint may persist, often as a non-typeset version available free, after an article is published in a journal. The immediate distribution of preprints allows authors to receive early feedback from their peers, which may be helpful in revising and preparing articles for submission(14).

Both open peer-review and pre-prints initiatives add to the faster dissemination of scientific achievements as the publication of manuscripts in a peer-reviewed journal often takes months or even years from the time of initial submission, owing to the time required by editors and reviewers to evaluate and critique manuscripts, and the time taken by authors to address these critiques(14). The need to quickly circulate current results within a scholarly community has been a key driver towards their growing acceptance.

We tried to compile important strategic plans and thoughts researchers may need to keep in their discussions and initiatives before they dream of a more beneficial cultural shift for the benefitting both science and humankind.

This is not a One-Man Show!

Scholarly communications is an ecosystem, with Funders/Governments providing the revenues and funds for feeding the ecosystem. Fig: 2 depicts different stakeholders in the scholarly ecosystem. It's been clear and apparent that despite having different priorities for various stakeholders the juice/motto behind their priorities flows to improvise the quality of the scholarly output.

Fig: 2 *Individual stakeholders and their priorities in the scholarly communication process*

The initiatives or the steps that involve a cultural shift should be able to satisfy all the priorities of different stakeholders. One reason why the open-access initiative is not disruptive is that there is priority mismatch between funders and publishers. Funders needed science to be accessible freely, but the majority of publishers are reluctant because they don't see or apprehensive about a sustainable business in the drive. This is evidenced by the fact that funders are now working with publishers to figure out a sustainable business model for them and also by supporting initiatives that aim at archiving scholarly outputs.

High Priority Areas of Research Assessments: "Are we there Yet?"

Researchers are the key players in the scholarly ecosystem. From funding acquisitions to conducting, and communicating their research is all done bared by the academicians. With a single expectation in mind that all of their efforts be added up to their academic reputation and also to prove to the funding agencies, societies and governments that we did something with the money that has invested in us.

4

Figure 1: vision of a set of metrics across the research workflow to support optimal decision-making
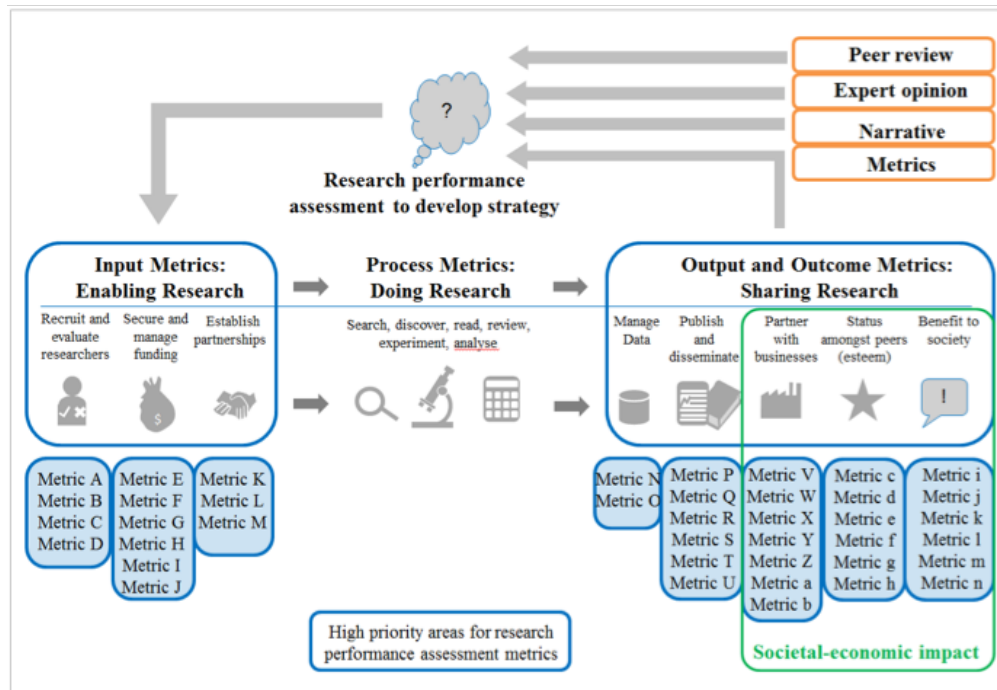

Fig: 3 Vision of a set of Metrics across the research workflow (re-use permission awaited).

In response to HEFCE's call for evidence, Elsevier conducted an independent review of the role of metrics in research assessment, and the report shows the myriad of parameters that may be available for research assessment. Fig:3 depicts the review over the parameters that are available which may be used for evaluation of research. Unfortunately, we don't have any metrics that aims to identify the way that the researchers conduct the research. And even till today we indirectly rely on the impact of research to assess the researchers.

A critical problem with the way that metrics that are currently used is the fact that the impact of research is usually calculated using the end result of a researcher's work. So this, and every other metrics like Impact Factors, Citation matrices, is biased towards scientists whose papers have been out for extended periods of time, and hence may not be ideal for assessing younger researchers.

Young researchers, including doctoral students and postdocs, are the main workhorses in producing research data and a myriad of other stakeholders are involved in other steps in the research workflow. CrediT CASRAI has defined a contribution role taxonomy depicts possible roles of different contributors to the research process and output workflow.

Campaigns and workshops that aim to educate researchers about ethical work habits and explain the consequences of data fabrication or falsified research will in the long term contribute to the cultural shift, but a major drive would or can only happen if we stop ourselves, from assessing

5

just the end result of the research. An ideal assessment method should aim at identifying or recognizing all the steps involved in research.

Conclusion

The following points summarize the factors that may have a significant role in increasing the Open Data movement's momentum and in clarifying its real purpose.

1. Satisfy the priorities of all the stakeholders involved in scholarly processes and use it as a primer to initiate open-data drive. This can be achievable that is inter-operable and aim to 'follow the data' through the research cycle which also try and align various stakeholders without affecting their priorities (Fig:2).

2. Researchers should idolize a method or metrics that aims in differentiating Researcher's assessments with that of Research assessment.

3. Researchers should not just focus solely on the END result. Of course, it's necessary to ascertain research productivity. But research involves endless optimizations, and hence cultural shift towards better practices will be fool-proof if we also start to focus ourselves on all the steps involved in research workflow.

   The possible consequences of focusing solely on the end result may be extrapolated to **longer durations of conducting research,** as researchers always worry about how they can create a story around the results that can be published (narration). Well to address this already Science-matters has started an initiative to encourage researchers to stop worrying about the storytelling and keep publishing results and narrate on the go. Other consequences may be that researchers suffer from peer pressures to publish and hence may be encouraged to adopt unethical standards of skewing and improper repetitions to confirm data and results which finally may have lead to rising of problems like Irreproducibility.

   Like recent developments in this regard taken by one of the funders Wellcome Open Research platform to make the all the associated outputs of research openly available & Digital Science innovation towards making research community work smarter & discover more. Perhaps the work by Kramer & Bosman suggests that it's about connecting the dots together & it's more about people, practice & interoperability than the tools. So working in this direction & going a step further we are developing open workflow (Profeza) embracing the priority of all the individual stakeholder in the scholarly ecosystem so that Open Science & research becomes a reality. Of course, there will be further studies to test this hypothesis, building collaborations with partners in this space & investigate sustainability/business model.

6

Acknowledgements

References:

1. http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1002165

2. http://faseb.org/Science-Policy-and-Advocacy/Federal-Funding-Data/NIH-Research-Funding-Trends.aspx

3. https://en.wikipedia.org/wiki/Louis_Pasteur

4. http://libereurope.eu/blog/2017/03/27/report-towards-competitive-sustainable-oa-market-europe/

5. https://www.acu.ac.uk/research-information-network/finch-report

6. https://www.elsevier.com/research-intelligence/resource-library/response-to-hefces

7. https://elife.elifesciences.org/content/2/e00855

8. http://docs.casrai.org/CRediT

9. https://doi.org/10.1038%2Fnature.2016.20224

10. https://en.wikipedia.org/wiki/Altmetrics

11. https://works.bepress.com/ian_mccullough/15/download/

12. http://pdxscholar.library.pdx.edu/cgi/viewcontent.cgi?article=1186&context=ulib_fac

13. https://en.wikipedia.org/wiki/Pre-print