OPINION ARTICLE

Imagining tomorrow's university: open science and its impact [version 1; referees: awaiting peer review]

Adina Howe [1*], Michael D. Howe[2*], Amy L. Kaleita[1*], D. Raj Raman[1*]

[1]Department of Agricultural and Biosystems Engineering, Iowa State University, Ames, IA, 50011, USA
[2]Department of Management, Iowa State University, Ames, IA, 50011, USA

[*] Equal contributors

v1 **First published:** 31 Mar 2017, **6**:405 (doi: 10.12688/f1000research.11232.1)
Latest published: 31 Mar 2017, **6**:405 (doi: 10.12688/f1000research.11232.1)

Abstract

As part of a recent workshop entitled "Imagining Tomorrow's University", we were asked to visualize the future of universities as research becomes increasingly data- and computation-driven, and identify a set of principles characterizing pertinent opportunities and obstacles presented by this shift. In order to establish a holistic view, we take a multilevel approach and examine the impact of open science on individual scholars as well as on the university as a whole. At the university level, open science presents a double-edged sword: when well executed, open science can accelerate the rate of scientific inquiry across the institution and beyond; however, haphazard or half-hearted efforts are likely to squander valuable resources, diminish university productivity and prestige, and potentially do more harm than good. We present our perspective on the role of open science at the university.

Open Peer Review

Referee Status: *AWAITING PEER REVIEW*

Discuss this article

Comments (0)

Corresponding author: Adina Howe (adina@iastate.edu)

How to cite this article: Howe A, Howe MD, Kaleita AL and Raman DR. **Imagining tomorrow's university: open science and its impact [version 1; referees: awaiting peer review]** *F1000Research* 2017, **6**:405 (doi: 10.12688/f1000research.11232.1)

Grant information: The author(s) declared that no grants were involved in supporting this work.

Competing interests: No competing interests were disclosed.

First published: 31 Mar 2017, **6**:405 (doi: 10.12688/f1000research.11232.1)

Introduction

As part of a recent workshop entitled "Imagining Tomorrow's University", we were asked to visualize the future of universities as research becomes increasingly data- and computation-driven, and to identify a set of principles characterizing pertinent opportunities and obstacles presented by this shift. To establish a holistic view, we take a multilevel approach and examine the impact of open science on individual scholars as well as on the university as a whole. Generally, we agree that increased transparency in the scientific process can broaden and deepen scientific inquiry, understanding, and impact. However, the realization of these outcomes will require significant time, effort, and aptitude to convey the means by which data are transformed into knowledge. We propose that open science can most effectively enable this evolution when it is conceptualized as a multifaceted pathway that includes:

- The provision of accessible and well-described data, along with information about its context[1];
- The methodology and mechanisms necessary to reproduce data analyses;
- Training products that provide transparent understanding of how the data can be applied to answer questions.

Thus, impactful open science requires investments from individual researchers that are often greater than those that might be needed for "non-open" science. At the university level, open science represents a double-edged sword: when well executed, it can accelerate the rate of scientific inquiry across the institution and beyond; however, haphazard or half-hearted efforts are likely to squander valuable resources and diminish university productivity and prestige, potentially doing more harm than good. Here, we present our perspective on the varying roles of open science.

Open science enables low-barrier collaborations

For some university researchers, open science can be both powerful and transformative. Imagine a research program that generates not only publications but also develops code that can quickly reproduce each analysis and publishable figure with a minimal amount of manual intervention. This structure can provide continuity in a project and accelerate the research enterprise by allowing researchers to rapidly repeat the same analysis on new datasets, all while lowering training and other human capital investments. Included in a publication, this "research notebook" and accompanying datasets (e.g., 2), could be compiled into a tutorial for others in the field who could then repeat this work with their own data – all without the need for formal collaborations. Such approaches can benefit not only the initiating research group but also an entire scientific discipline.

Open science requires significant investment

While the opportunities of open science practices hold promise, several costs and obstacles may prevent its realization and impact. A key cost of open science is time – time to format, annotate and publish data and associated metadata; time to learn new tools that allow for automated analysis and reproduction; time to produce scripts with a sufficient level of robustness and documentation to be useful to others[3], and so on. Of these, arguably, the least time-consuming step is simply providing access to data. While open data is an important component of open science, it is far from the whole enchilada, and does not provide the broad benefits of open science writ large.

It would be irresponsible to discuss open data and open science without acknowledging the risk posed to the anonymity that is so central to many human research studies. For example, to promote participant anonymity, data resulting from research currently conducted under the auspices of an IRB may be ineligible for distribution outside of the immediate research team. As multiple sources of open data become increasingly available, privacy concerns of this nature are likely to increase along with the prevalence of unintended participant identification[4,5]. In these cases, the benefits of open science may not stem from sharing data but rather reproducible analyses that may be more broadly useful, and the provision of open data does not in itself translate into our vision of open science. At the university level, the incentives to facilitate and expand open science at the university should not be monolithic (e.g., data-centric), but rather be selectively created and applied to maximize success and minimize unintended harm. Open science also presents unique challenges as universities and other research institutions turn increasingly to private sector funding, which comes with proprietary limitations on the dissemination of results.

The broader impact of open science is uncertain

It is possible that the increasing availability and transparency of scientific inquiry could ignite broader interest in research. The current publishing paradigm of most fields limits research availability to a relatively narrow audience, with paid access to scientific journals. Meanwhile, polling data from Gallup indicates a slow but relatively steady decline in Americans' trust of institutions in general since 2000[6], although Gallup does not include "universities" specifically in the poll. In one study that compared follow-on inventions from discoveries that were made simultaneously but separately at a university and at a corporate firm, the same discovery at a university was 20–30% less likely to be used in follow-up innovations[7,8]. This study also included open-ended interviews to shed light on this "Ivory Tower effect"; and a key driver appeared to be "considerable skepticism toward academic science." More openness in university science research may help to address this apparent skepticism.

Even though there are concerns associated with society's growing disconnect with the scientific enterprise and the accompanying devaluation of research, it should be noted that in general academics are still held in high regard and seen as reliable sources of information for a wide range of issues[9,10]. To maintain this esteem, it is important to realize that data without an understanding of what it entails or the questions it can answer can be considered useless and even dangerous when used improperly to influence decision-making and policy[11]. Thus, providing useful open data requires more thought on how this data can be translated into useful information. Mechanisms to reproduce analyses and communications that explain the complexities and intricacies of these tasks could be an important first step. While the peer-reviewed-publication paradigm currently provides an established, if not optimal, communication mechanism for conveying the results of scientific activities to our peers, no such standard currently exists to govern the creation and exchange of open science to our peers and beyond. Efforts at the

university level that encourage the rigorous construction of appropriate dissemination systems are laying the foundation for success in this endeavor.

A path forward: recognition, training and infrastructure

Universities have a moral responsibility to educate, and there are significant opportunities in the open science model to broaden the output of research with an eye towards education. Nevertheless, the current university promotion and tenure system is optimized for evaluating the traditional format of peer-reviewed journals as the only necessary and sufficient product of a research project. Given the "publish or perish" paradigm that currently pervades the academy, an accompanying lack of recognition for the time and effort put into facilitating open science is apt to dampen participation[12]. For example, utilizing openly available code for an analysis in a subsequent publication does not require a citation, and even if the code were to be highly cited, it does not carry the same weight as a peer-reviewed publication. Thus, universities have an opportunity to re-imagine what it means to contribute to research, specifically extending the definition to include more than a tally of peer reviewed publications. The development of robust, reliable, and transparent tools to track utilization of open science products may be one path forward to quantitatively measure the impact of faculty generated research outputs not currently tracked or rewarded, and both incentivize and acknowledge the resources required to effectively engage in open science.

A notable effort to define the characteristics of open science products are the FAIR Data Principles[13], which emphasize that scholarly products should be findable, accessible, interoperable, and reusable and that good data management is not a goal in itself but can catalyze knowledge discovery and innovation. At the university, training for sustainable data management best practices would deepen the overall understanding of the opportunities of open science. In many respects, the products of open science are a common good resource[14], but require support infrastructure to share data, tools, and training to broaden participation. This infrastructure could also be re-imagined to include metrics to quantify impact, supporting the need to acknowledge contributions.

In conclusion, open science is a significant opportunity for universities, but a one-size-fits-all approach is sub-optimal. Executing open science in a way that facilitates meaningful advances requires a personal investment of time, both upfront to develop relevant capabilities, and ongoing for execution expenses. As such, it is important that universities develop infrastructure and training to support, measure, and reward efforts that deliver on the promise of open science, focusing on domains best positioned to further scientific understanding.

A preprint of this article can be found on PeerJ (https://doi.org/10.7287/peerj.preprints.2781v1).

Author contributions

AH, MH, AK, RR contributed equally in the preparation of this manuscript and have agreed to the final content.

Competing interests

No competing interests were disclosed.

Grant information

The author(s) declared that no grants were involved in supporting this work.

References

1. Faniel IM, Jacobsen TE: **Reusing Scientific Data: How Earthquake Engineering Researchers Assess the Reusability of Colleagues' Data.** *Comput Support Coop Work.* Springer Netherlands. 2010; **19**(3): 355–375.
Publisher Full Text
2. Howe A, Chain P: **Example of a reproducible IPython Notebook for Analysis [Internet].**
Reference Source
3. Barnes N: **Publish your computer code: it is good enough.** *Nature.* Nature Publishing Group. 2010; **467**(7317): 753.
PubMed Abstract | Publisher Full Text
4. Sweeney L: **Simple Demographics Often Identify People Uniquely.** Tech Rep LIDAP-WP. 2000.
Reference Source
5. Tonidandel S, King EB, Cortina JM: **Big Data Methods: Leveraging Modern Data Analytic Techniques to Build Organizational Science.** *Organ Res Methods.* SAGE Publications Sage CA: Los Angeles, CA; 2016.
Publisher Full Text
6. Poll G: **Americans' Confidence in Institutions Stays Low | Gallup [Internet].** [cited 1 Feb 2017].
Reference Source
7. Bikard M: **Is Knowledge Trapped Inside the Ivory Tower? Technology Spawning and the Genesis of New Science-Based Inventions.** 2012.
Reference Source
8. Vermuelen F: **Why Firms Don't Trust Universities - Business Insider [Internet].** 2013. [cited 30 Jan 2017].
Reference Source
9. Nisbet MC, Kotcher JE: **A Two-Step Flow of Influence?: Opinion-Leader Campaigns on Climate Change.** *Sci Commun.* SAGE Publications Sage CA: Los Angeles, CA. 2009; **30**(3): 328–354.
Publisher Full Text
10. Leiserowitz A, Maibach EW, Roser-Renouf C, *et al.*: **Climate Change in the American Mind: Americans' Global Warming Beliefs and Attitudes in April 2013.** *SSRN Electron J.* 2013.
Publisher Full Text
11. Gorby YA, Yanina S, McLean JS, *et al.*: **Electrically conductive bacterial nanowires produced by *Shewanella oneidensis* strain MR-1 and other microorganisms.** *Proc Natl Acad Sci U S A.* 2006; **103**(30): 11358–11363.
PubMed Abstract | Publisher Full Text | Free Full Text
12. de Rond M, Miller AN: **Publish or Perish: Bane or Boon of Academic Life?** *J Manag Inq.* 2005; **14**(4): 321–329.
Publisher Full Text
13. Wilkinson MD, Dumontier M, Aalbersberg IJ, *et al.*: **The FAIR Guiding Principles for scientific data management and stewardship.** *Sci Data.* Nature Publishing Group; 2016; **3**: 160018.
PubMed Abstract | Publisher Full Text | Free Full Text
14. Hardin G: **The tragedy of the commons.** *Science.* 1968; **162**(3589): 1243–8.
PubMed Abstract | Publisher Full Text