

2017

Imagining tomorrow's university in an era of open science

Adina Howe

Iowa State University, adina@iastate.edu

Michael D. Howe

Iowa State University, howe@iastate.edu

Amy L. Kaleita

Iowa State University, kaleita@iastate.edu

D. Raj Raman

Iowa State University, rajraman@iastate.edu

Follow this and additional works at: http://lib.dr.iastate.edu/abe_eng_pubs

 Part of the [Higher Education Commons](#), [Science and Mathematics Education Commons](#), and the [Science and Technology Policy Commons](#)

The complete bibliographic information for this item can be found at http://lib.dr.iastate.edu/abe_eng_pubs/780. For information on how to cite this item, please visit <http://lib.dr.iastate.edu/howtocite.html>.

This Article is brought to you for free and open access by the Agricultural and Biosystems Engineering at Iowa State University Digital Repository. It has been accepted for inclusion in Agricultural and Biosystems Engineering Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

Imagining tomorrow's university in an era of open science

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 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.

Keywords

Ethical issues, Science and medical education, Science policy, Computational science

Disciplines

Higher Education | Science and Mathematics Education | Science and Technology Policy

Comments

This article is from *F1000Research* 2017, 6:405 (doi: [10.12688/f1000research.11232.2](https://doi.org/10.12688/f1000research.11232.2)). Posted with permission.

Creative Commons License



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).



OPINION ARTICLE

REVISED **Imagining tomorrow's university in an era of open science [version 2; referees: 3 approved]**

Imagining tomorrow's university: open science and its impact in an era of open science

Adina Howe ^{1*}, Michael Howe^{2*}, Amy L. Kaleita^{1*}, D. Raj Raman^{1*}

¹Department of Agricultural and Biosystems Engineering, Iowa State University, Ames, IA, 50011, USA

²Department of Management, Iowa State University, Ames, IA, 50011, USA

* Equal contributors

v2 **First published:** 31 Mar 2017, 6:405 (doi: [10.12688/f1000research.11232.1](https://doi.org/10.12688/f1000research.11232.1))
Latest published: 19 May 2017, 6:405 (doi: [10.12688/f1000research.11232.2](https://doi.org/10.12688/f1000research.11232.2))

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 and how this impacts 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:

	Invited Referees		
	1	2	3
REVISED			
version 2 published 19 May 2017		report	report
version 1 published 31 Mar 2017		?	?
	report	report	report

- 1 **May Khanna** , University of Arizona USA
- 2 **Nathan L. Vanderford** , University of Kentucky USA
- 3 **Marie-Claire Shanahan** , University of Calgary Canada

Discuss this article

Comments (0)

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

Competing interests: No competing interests were disclosed.

How to cite this article: Howe A, Howe M, Kaleita AL and Raman DR. **Imagining tomorrow's university in an era of open science [version 2; referees: 3 approved]** *F1000Research* 2017, **6**:405 (doi: [10.12688/f1000research.11232.2](https://doi.org/10.12688/f1000research.11232.2))

Copyright: © 2017 Howe A *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

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

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

REVISED Amendments from Version 1

Version 2 of this manuscript is modified to include improvements suggested by three reviewers. Revisions include a specific introduction to the definition of open science and modified citations to provide more specific examples of open science benefits and arguments. We have also modified the title to better reflect the content of this perspective.

See referee reports

Introduction

The mission of universities, specifically land-grant institutions originating from the Morrill Act of 1862, is to provide accessible education and scholarship to *all* people. In a similar vein, open science has emerged as an approach to minimize the barriers associated with traditional ways of sharing the outcomes of scholarship. As defined by the Open Definition (<https://okfn.org/>), open science embodies the notion that information is available for anyone to “freely access, use, modify, and share for any purpose”, regardless of personal or institutional resources. Fostered by increasingly data- and computation-driven research, universities are uniquely positioned to reimagine their role in knowledge dissemination vis-à-vis the principles of open science. As part of a recent workshop entitled “*Imagining Tomorrow’s University*”, we were asked to visualize the future of universities in an open, networked era and to 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. 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 successfully 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 context¹;
- 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 transformative². 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., 3), 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 imposes significant costs

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; and time to produce scripts with a sufficient level of robustness and documentation to be broadly useful to others⁴. 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 requires significant investment and does not provide the broad benefits of open science writ large. These investments include personnel, labor, computational infrastructure to store data, and publication costs to communicate findings in an open way (up to \$3900, 5). Consequently, a challenge for university engagement with open science is determining who should bear the financial burden of these costs. For example, a grant proposal that requires \$10,000 additional for open access publishing fees may not be viewed as competitively as one with these dollars allocated for direct research costs. Similarly, while universities can directly promote open science by subsidizing open access publication fees or providing cost-sharing opportunities, they too must decide where best to invest their limited resources.

Further, 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^{6,7}. 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 potential for broader impacts with open science

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 2000⁸, 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 innovations⁹. This study also included open-ended interviews to shed light on this "Ivory Tower effect", and a 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^{10,11}. 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 policy¹². 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

Recognize open science impacts. 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 participation¹². 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.

Train best practices and provide infrastructure to broaden participation. A notable effort to define the characteristics of open science products are the FAIR Data Principles¹³, 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 inherent in open science. In many respects, the products of open science are available to benefit by all that require support infrastructure to share data, tools, and training to broaden participation and limit exploitation. 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. McKiernan EC, Bourne PE, Brown CT, *et al.*: **How open science helps researchers succeed.** *eLife.* eLife Sciences Publications Limited. 2016; **5**: pii: e16800.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
3. Howe A, Chain P: **Example of a reproducible IPython Notebook for Analysis [Internet].**
[Reference Source](#)
4. Barnes N: **Publish your computer code: it is good enough.** *Nature.* Nature Publishing Group. 2010; **467**(7317): 753.
[PubMed Abstract](#) | [Publisher Full Text](#)
5. Van Noorden R: **Open access: The true cost of science publishing.** *Nature.* 2013; **495**(7442): 426–429.
[PubMed Abstract](#) | [Publisher Full Text](#)
6. Sweeney L: **Simple Demographics Often Identify People Uniquely.** Tech Rep LIDAP-WP. 2000.
[Reference Source](#)
7. 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](#)
8. Poll G: **Americans' Confidence in Institutions Stays Low | Gallup [Internet].** [cited 1 Feb 2017].
[Reference Source](#)
9. Bikard M: **Peer-Based Knowledge Validation: A Hurdle to the Flow of Academic Science to Inventors.** 2012.
[Publisher Full Text](#)
10. 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](#)
11. 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](#)
12. Longo DL, Drazen JM: **Data Sharing.** *N Engl J Med.* Massachusetts Medical Society, 2016; **374**(3): 276–277.
[PubMed Abstract](#) | [Publisher Full Text](#)

Open Peer Review

Current Referee Status:   

Version 2

Referee Report 26 May 2017

doi:[10.5256/f1000research.12634.r22928](https://doi.org/10.5256/f1000research.12634.r22928)



Marie-Claire Shanahan

University of Calgary, Calgary, AB, Canada

Thank you for the revisions made to the paper and to the attention given to the comments from all three reviewers. The argumentation has been substantially clarified and the issues with references have been corrected.

Competing Interests: I contributed to a discussion paper for the same workshop (Imagining Tomorrow's University) but the papers were not in competition with one another, evaluated relative to one another, or written in collaboration with each other in any way.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Referee Report 22 May 2017

doi:[10.5256/f1000research.12634.r22930](https://doi.org/10.5256/f1000research.12634.r22930)



Nathan L. Vanderford

Markey Cancer Center, Department of Toxicology and Cancer Biology, College of Medicine, University of Kentucky, Lexington, KY, USA

Howe et al. have been responsive to my previous review as well as to the review of the other two reviewers. The article is improved and I find it to be acceptable at this stage.

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Referee Report 15 May 2017

doi:10.5256/f1000research.12118.r22286



Marie-Claire Shanahan

University of Calgary, Calgary, AB, Canada

This is a timely and important topic, and it is clear that the authors have first-hand experience with the area of open science related to creating open and shared analysis code (e.g., Citation #2). They have a very valuable perspective to add to the scientific community's conversations about open science. And while I definitely support the eventual indexing of this article, I feel that there are some areas in which the argument should be strengthened and clarified first.

As reviewer Vanderford notes, a clear definition of open science is needed early on. This would also be very helpful for strengthening the arguments that develop in the middle and final sections that open databases are not sufficient for open science and that universities need better infrastructure for recognizing shared code as an academic contribution. A clear definition would also help place the open analysis code argument more clearly. Is sharing analysis programs and code sufficient for open science (i.e., is it synonymous with open science) or is it instead an under-recognized but important element or type of open science that the authors wish to highlight?

The paper seems to settle in on open analysis code as the central argument later on, but in the opening sections the argument seems overly broad for the examples and support that are given. The only example of benefits that is given is of the "open notebook" (a good and valuable example) but it is not sufficient to support broad claims about open science as a double-edged sword (where neither the broads benefits or potential downfalls are explained in detail or supported with evidence from the literature). I think focusing the argument and placing their perspective more clearly within the broader field of open science early on would create a more cohesive argument and one that can be better supported with the experiences and examples the authors provide.

Is the topic of the opinion article discussed accurately in the context of the current literature?

Are all factual statements correct and adequately supported by citations?

Are arguments sufficiently supported by evidence from the published literature?

In addition to the example above there are a few places where support for statements and accuracy with relation to the literature could be improved. Defining open science clearly and placing the authors' perspectives related to open analysis code within that larger definition would help improve connections to the literature. A few examples are below that might be helpful.

Later on, citations 7 and 8 are used to support discussion of take up of academic discoveries. These are both citations of the same study though, with 7 being the study itself and 8 being a popular media report on the study. Using both seems to suggest that there are two independent sources to support this claim. There is also a published version of the study that might be a preferred citation¹. And I think it might be helpful to note that the study does not find mistrust to be the main reason for lack of uptake but says it is secondary to the natural competitiveness of industrial science, where they are constantly monitoring competitors and therefore likely to notice discoveries that competitors make.

"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]. This is a strong claim, and it could represent important reservations about open science, but no support is provided. The citation does not seem to be related at

all (Title: "Electrically conductive bacterial nanowires produced by *Shewanella oneidensis* strain MR-1 and other microorganisms.") Are the authors referring to some controversy surrounding the inappropriate use of that data for decision making and policy? If so, this needs to be explained and supported explicitly. Otherwise, other citations should be found to support this claim.

Referring to Hardin's "Tragedy of the commons" [Citation 14] also does not seem like a closely related source for the use of the term "common good" as the authors have used it. Some clarification would be helpful there as well. Neilsen's book², similarly talks about "knowledge commons" specifically in relation to open science in a way that might be more relevant to the arguments made here.

Are the conclusions drawn balanced and justified on the basis of the presented arguments?

In the end, the authors make an important and valid argument about university supports and infrastructure, but the points leading up to that conclusion could be more clearly explained and better supported and connected. From the section "A path forward" and onward, the examples lead nicely towards the conclusion. Given the authors expertise I feel that the could be expanded a bit and to explore and support the conclusion, and the earlier paragraphs could focus more specifically on the issues of open analysis code to build towards that conclusion. For example, in discussing the Ivory Tower effect earlier on, the original study that is cited explores the publish or perish system as one of the reasons for distrust. The argument made about publish or perish later on could be more meaningful if that connection had been made in the previous section.

Overall, the authors have an important contribution to make to discussions of open science and important expertise in the practice of open science. With some clarification to the supports and the argument, this paper will be a valuable and interesting piece of that conversation.

References

1. Bikard M: Is Academic Science Trapped Inside the Ivory Tower? Universities and the Diffusion of New Knowledge. *Academy of Management Proceedings*. 2012; **2012** (1). [Publisher Full Text](#)
2. Nielsen M: Reinventing discovery: the new era of networked science. *Princeton University Press*. 2012.
3. Nosek BA, Alter G, Banks GC, Borsboom D, Bowman SD, Breckler SJ, Buck S, Chambers CD, Chin G, Christensen G, Contestabile M, Dafoe A, Eich E, Freese J, Glennerster R, Goroff D, Green DP, Hesse B, Humphreys M, Ishiyama J, Karlan D, Kraut A, Lupia A, Mabry P, Madon TA, Malhotra N, Mayo-Wilson E, McNutt M, Miguel E, Paluck EL, Simonsohn U, Soderberg C, Spellman BA, Turitto J, VandenBos G, Vazire S, Wagenmakers EJ, Wilson R, Yarkoni T: SCIENTIFIC STANDARDS. Promoting an open research culture. *Science*. 2015; **348** (6242): 1422-5 [PubMed Abstract](#) | [Publisher Full Text](#)
4. Easterbrook S: Open code for open science?. *Nature Geoscience*. 2014; **7** (11): 779-781 [Publisher Full Text](#)

Is the topic of the opinion article discussed accurately in the context of the current literature?

Partly

Are all factual statements correct and adequately supported by citations?

Partly

Are arguments sufficiently supported by evidence from the published literature?

Partly

Are the conclusions drawn balanced and justified on the basis of the presented arguments?

Partly

Competing Interests: I contributed to a discussion paper for the same workshop (Imagining Tomorrow's University) but the papers were not in competition with one another, evaluated relative to one another, or written in collaboration with each other in any way.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 17 May 2017

Adina Howe, Iowa State University, USA

1. As reviewer Vanderford notes, a clear definition of open science is needed early on. This would also be very helpful for strengthening the arguments that develop in the middle and final sections that open databases are not sufficient for open science and that universities need better infrastructure for recognizing shared code as an academic contribution. A clear definition would also help place the open analysis code argument more clearly. Is sharing analysis programs and code sufficient for open science (i.e., is it synonymous with open science) or is it instead an under-recognized but important element or type of open science that the authors wish to highlight?

Thank you for this suggestion. We have now included a clearer open science definition within the introduction to help clarify our perspective and define "sharing" data as an important element within open access themes of "access, use, modify, and sharing".

2. The paper seems to settle in on open analysis code as the central argument later on, but in the opening sections the argument seems overly broad for the examples and support that are given. The only example of benefits that is given is of the "open notebook" (a good and valuable example) but it is not sufficient to support broad claims about open science as a double-edged sword (where neither the broad benefits or potential downfalls are explained in detail or supported with evidence from the literature). I think focusing the argument and placing their perspective more clearly within the broader field of open science early on would create a more cohesive argument and one that can be better supported with the experiences and examples the authors provide.

In response to reviewer #2, more specific examples, of the benefits and barriers to open science have been included to represent the broader field of open science (e.g. publishing cost). Further, we believe that the modification of the title and introduction revisions help to focus the central topic of this effort on the impacts of a university in an open era. Finally, for more specific examples, we have also cited a McKiernan et al. 2016 which represents a more data-centric approach of the benefits of open science.

3. In addition to the example above there are a few places where support for statements and accuracy with relation to the literature could be improved. Defining open science clearly and placing the authors' perspectives related to open analysis code within that larger definition would help improve connections to the literature. A few examples are below that might be helpful.

Later on, citations 7 and 8 are used to support discussion of take up of academic discoveries. These are both citations of the same study though, with 7 being the study itself and 8 being a popular media report on the study. Using both seems to suggest that there are two independent sources to support this claim. There is also a published version of the study that might be a preferred citation¹. And I think it might be helpful to note that the study does not find mistrust to be the main reason for lack of uptake but says it is secondary to the natural competitiveness of industrial science, where they are constantly monitoring competitors and therefore likely to notice discoveries that competitors make.

Thank you for these comments – we agree that two citations here is misleading and have removed the citation #8 and replaced the original preprint with the suggested citation provided more recent published study (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2333413). The general results of the cited study indicate that “the results suggest that the peer-based knowledge validation process in academia creates uncertainty about the reliability and relevance of academic science as a map for technology development.” [Bikard, 2015]. We also modified the reason inventors draw on knowledge from firms rather than academics to be associated (e.g., a driver vs a key driver) with skepticism toward academic science.

4. "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]. This is a strong claim, and it could represent important reservations about open science, but no support is provided. The citation does not seem to be related at all (Title: "Electrically conductive bacterial nanowires produced by *Shewanella oneidensis* strain MR-1 and other microorganisms.") Are the authors referring to some controversy surrounding the inappropriate use of that data for decision making and policy? If so, this needs to be explained and supported explicitly. Otherwise, other citations should be found to support this claim.

Thank you for this observation – we've corrected the citation.

5. Referring to Hardin's "Tragedy of the commons" [Citation 14] also does not seem like a closely related source for the use of the term "common good" as the authors have used it. Some clarification would be helpful there as well. Neilsen's book², similarly talks about "knowledge commons" specifically in relation to open science in a way that might be more relevant to the arguments made here.

To clarify this phrase, we modified “common good” to explicitly state “available to benefit by all”.

6. In the end, the authors make an important and valid argument about university supports and infrastructure, but the points leading up to that conclusion could be more clearly explained and better supported and connected. From the section "A path forward" and onward, the examples lead nicely towards the conclusion. Given the authors expertise I feel that the could be expanded a bit and to explore and support the conclusion, and the earlier paragraphs could focus more specifically on the issues of open analysis code to build towards that conclusion. For example, in

discussing the Ivory Tower effect earlier on, the original study that is cited explores the publish or perish system as one of the reasons for distrust. The argument made about publish or perish later on could be more meaningful if that connection had been made in the previous section.

Thank you for your suggestions. With these suggestions in mind, we have edited the text for a more natural flow, using both additional examples and citations and specific headers.

Competing Interests: No competing interests were disclosed.

Referee Report 27 April 2017

doi:10.5256/f1000research.12118.r22285



Nathan L. Vanderford

Markey Cancer Center, Department of Toxicology and Cancer Biology, College of Medicine, University of Kentucky, Lexington, KY, USA

This article by Howe et al. is a generally well-written commentary on the importance of open science as well as some of the current barriers to its most effective implementation and use. This is a very timely topic that certainly deserves discussion of its value to the scientific community and of how it can be improved.

I believe there are several noteworthy areas in which the article could be improved prior to it being approved for indexing.

1. The authors should step back at the very onset of the article and define open science. It would also be helpful to provide more context on the history of open science and why/ how that history is important to the current article.
2. While this is a well-written article in general, there are a few phrases used that should be reconsidered. For example, the use of the phrase “the whole enchilada” should be re-written in a more professional phrasing.
3. I would argue that the most significant “costs” or “barriers” to open science are the financial costs (of which personnel labor/ time and electronic/ computer storage space issues are perhaps the biggest components). It is a missed opportunity to not mention the financial costs of open science. As part of that discussion, it would be interesting for the authors to discuss who bears the financial costs and whether there is room for improvement. For example, could funders/sponsors do more to support the costs of open science? And, are there opportunities for new policies (at the level of funders/sponsors) to be developed that could further support the wider implementation of open science?
4. The title of the article refers to the “impact” of open science but the current content of the article falls short on convincing the reader of the current and potential impact of open science. It may be

worth providing some specific examples of how open science has been used to make impactful discoveries, etc. Providing such examples could drive home the point of why it is so important to support and improve open science.

In summary, this is a timely article discussing a very important topic. There are limitations of the current version of the article that dampen this reviewer's enthusiasm at this time regarding giving a full approval. As such, I look forward to reviewing a revised version of the article.

Is the topic of the opinion article discussed accurately in the context of the current literature?

Yes

Are all factual statements correct and adequately supported by citations?

Yes

Are arguments sufficiently supported by evidence from the published literature?

Yes

Are the conclusions drawn balanced and justified on the basis of the presented arguments?

Yes

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 17 May 2017

Adina Howe, Iowa State University, USA

1. The authors should step back at the very onset of the article and define open science. It would also be helpful to provide more context on the history of open science and why/ how that history is important to the current article.

Great suggestion. We agree that this was missing and we now provide a context for universities and their role in open science as an introduction.

2. While this is a well-written article in general, there are a few phrases used that should be reconsidered. For example, the use of the phrase "the whole enchilada" should be re-written in a more professional phrasing.

Modified.

3. I would argue that the most significant "costs" or "barriers" to open science are the financial costs (of which personnel labor/ time and electronic/ computer storage space issues are perhaps the biggest components). It is a missed opportunity to not mention the financial costs of open science. As part of that discussion, it would be interesting for the authors to discuss who bears the financial costs and whether there is room for improvement. For example, could funders/sponsors do more to support the costs of open science? And, are there opportunities for new policies (at the level of funders/sponsors) to be developed that could further support the wider implementation of open science?

Thank you for pointing out this opportunity in this piece. We have added a discussion to bring up the topic of financial costs associated with open science and the complexities of determining incentives at the university level. While there is opportunity for funders/sponsors to help bear these financial burdens, the scope of this effort is what a university can do, and we have limited our discussion to this topic.

4. The title of the article refers to the “impact” of open science but the current content of the article falls short on convincing the reader of the current and potential impact of open science. It may be worth providing some specific examples of how open science has been used to make impactful discoveries, etc. Providing such examples could drive home the point of why it is so important to support and improve open science.

We agree that the title of this article was not a good fit for the content of the piece and have adjusted it accordingly. We have provided citations to one peer reviewed article that identifies examples and a perspective of how open science have provided impact. A challenge in providing specific examples is that the impact of open science is still debated and difficult to quantify. Most articles are often personal perspectives of specific authors and not necessarily data-driven approaches to study the positive impacts of open science (which are currently rather limited). Further, there are also examples of negative impacts of open science, <http://www.nature.com/news/open-data-contest-unearths-scientific-gems-and-controversy-1.21572?W>. We hope to present a balanced perspective on the complexities of open science at universities and guide thoughts on the most impactful path forward in this diverse environment.

Competing Interests: No competing interests were disclosed.

Referee Report 21 April 2017

doi:10.5256/f1000research.12118.r21433



May Khanna

Department of Pharmacology, College of Medicine, University of Arizona, Tucson, AZ, USA

This was a very well-written manuscript and quite relevant to our current times since there is a definite push for open science in academia. This is becoming increasingly necessary as we continually acquire larger data sets. I have a few comments and suggestions.

I would replace the words “the whole enchilada” under the headline “Open science requires significant Investment”. I would also replace “so on” in the same section.

Instead of the headline “Broader impact of open science is uncertain” may I suggest something along the line of: “Open and broad communication could impact open science”.

In the section “Open Science requires significant investment”, it is suggested that a key cost of open science is time, which is reasonable. However, one of the points is that “time to produce scripts with a sufficient level of robustness and documentation to be useful to others”; this point is less reasonable. I believe with or without open science, this should be an integral requirement of all scientists and so this last point should be omitted.

Your three first points to propose to enable open science more effectively are valid, however, you don't circle back to these in your final discussion. May I suggest using instead the points that you end with, such as:

1. Finding a way to properly cite open codes or data available through open sources
2. Developing a reliable, robust tool to track utilization of open science (note: similar idea to 1, but goes one step further)
3. Universities need to support infrastructure to implement FAIR data principles

This way you end with developing the points that should enable open science.

All these are suggestions. Should the authors decide to ignore them, it will not dampen my enthusiasm for this well-written manuscript.

Is the topic of the opinion article discussed accurately in the context of the current literature?

Yes

Are all factual statements correct and adequately supported by citations?

Yes

Are arguments sufficiently supported by evidence from the published literature?

Yes

Are the conclusions drawn balanced and justified on the basis of the presented arguments?

Yes

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 17 May 2017

Adina Howe, Iowa State University, USA

I would replace the words "the whole enchilada" under the headline "Open science requires significant investment". I would also replace "so on" in the same section.

Modified as suggested.

Instead of the headline "Broader impact of open science is uncertain" may I suggest something along the line of: "Open and broad communication could impact open science".

Modified the heading to "The Potential for Broader Impacts with Open Science"

In the section "Open Science requires significant investment", it is suggested that a key cost of open science is time, which is reasonable. However, one of the points is that "time to produce scripts with a sufficient level of robustness and documentation to be useful to others"; this point is less reasonable. I believe with or without open science, this should be an integral requirement of all scientists and so this last point should be omitted.

Modified this sentence to “time to produce scripts with a sufficient level of robustness and documentation to be **broadly** useful to others”. We agree that there is a minimal requirement of reproducibility for all scripts. However, our aim with this sentence was to convey that the impact of this e.g, documentation or code and its ability to be reproduced by a broad audience (e.g., the public vs. domain experts) requires significant and accountable requirements on time. We feel like the inclusion of the “broadly useful” more appropriately captures our intent – thank you for your suggestion!

Your three first points to propose to enable open science more effectively are valid, however, you don't circle back to these in your final discussion. May I suggest using instead the points that you end with, such as:

1. Finding a way to properly cite open codes or data available through open sources
2. Developing a reliable, robust tool to track utilization of open science (note: similar idea to 1, but goes one step further)
3. Universities need to support infrastructure to implement FAIR data principles

Great suggestion. Given the length of this article, we've provided some more directed heading to guide the conclusion, which are aligned with your suggestions.

This way you end with developing the points that should enable open science.

Competing Interests: No competing interests were disclosed.
