The State of Open Data 2019 – What are the key issues in open data for researchers?

As mandates and policies encouraging open data are becoming more widely established and enforced, the use of and sharing of data is becoming more central to scholarly communication. This has resulted in data sharing becoming increasingly entangled with the prestige economy of academia. In this post, **Mark Hahnel** *presents findings from the largest continuous survey of academic attitudes to open data and suggests that as well promoting data sharing, it may also have inadvertently fed into the publish or perish culture of research.*

The State of Open Data, Figshare's (in partnership with Digital Science and Springer Nature) annual report was released in October during Open Access Week. The report (the biggest yet) is a summary of findings from over 8,500 researchers surveyed on a broad range of topics including attitudes to sharing and reusing data, national open data mandates, familiarity with initiatives and community standards, incentives and credit mechanisms. Overall, we saw an increase in the awareness, acceptance and adoption of open data, but there was also a notable frustration with the lack of credit given for good data practices and a lack of punishment for non-compliance with open data policies.

Policies and mandates

In the last six years we've seen an increasing number of memos, mandates and policies from governments, funders, publishers and institutions encouraging the sharing of research data.

To take just one, in 2013, the Obama Whitehouse announced the Office of Science and Technology Policy 2013 Memorandum stating:

> *"The Obama Administration is committed to the proposition that citizens deserve easy access to the results of research their tax dollars have paid for. That's why, in a policy memorandum released today, OSTP Director John Holdren has directed Federal agencies with more than $100M in R&D expenditures to develop plans to make the results of federally funded research freely available to the public—generally within one year of publication."*

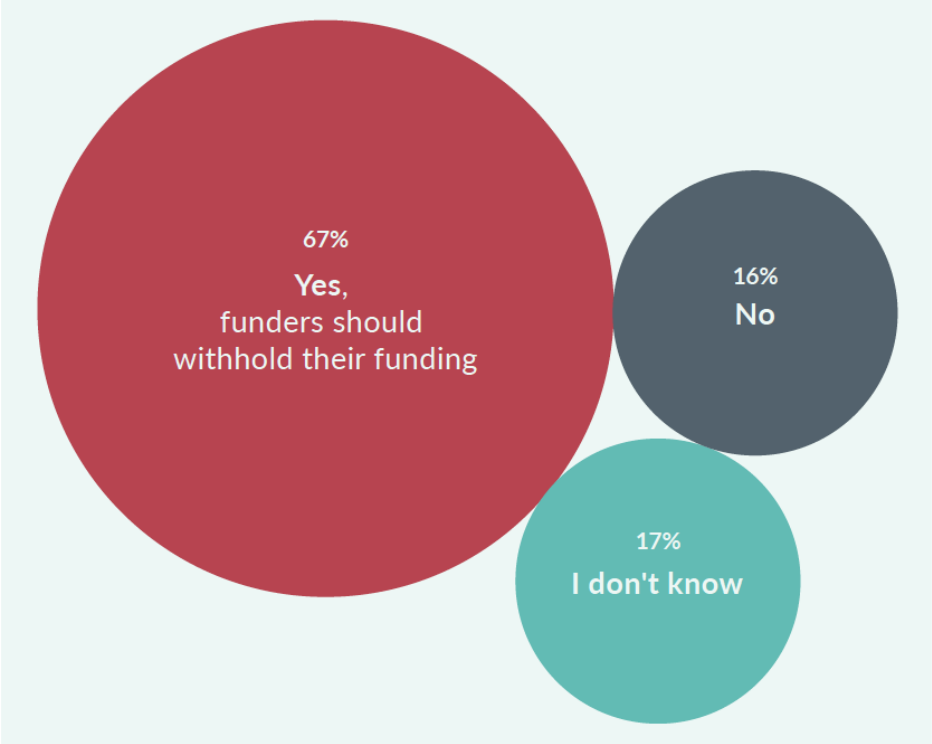
This was passed into law in January 2019 as the OPEN Government Data Act, which institutionalised the federal government's commitment to Open Data, and codified a policy of 'open by default' for all government data.

According to the findings of our survey, researchers welcome mandates of this nature and hope to see them policed accordingly.

Date originally posted: 2019-12-13

Permalink: https://blogs.lse.ac.uk/impactofsocialsciences/2019/12/13/the-state-of-open-data-2019-what-are-the-key-issues-in-open-data-for-researchers/

Blog homepage: https://blogs.lse.ac.uk/impactofsocialsciences/



- 79% of 2019 respondents were supportive overall of a national mandate for making primary research openly available
- 67% of respondents think that funders should withhold funding from, or penalise in other ways, researchers who do not share their data if the funder has mandated that they do so
- 69% of respondents think that funders should make the sharing of research data part of their requirements for awarding grants

Demographics

This is the fourth year we have carried out the *The State of Open Data* survey making it the longest running longitudinal study on the subject. The open outputs of all years can be found in this Figshare collection.

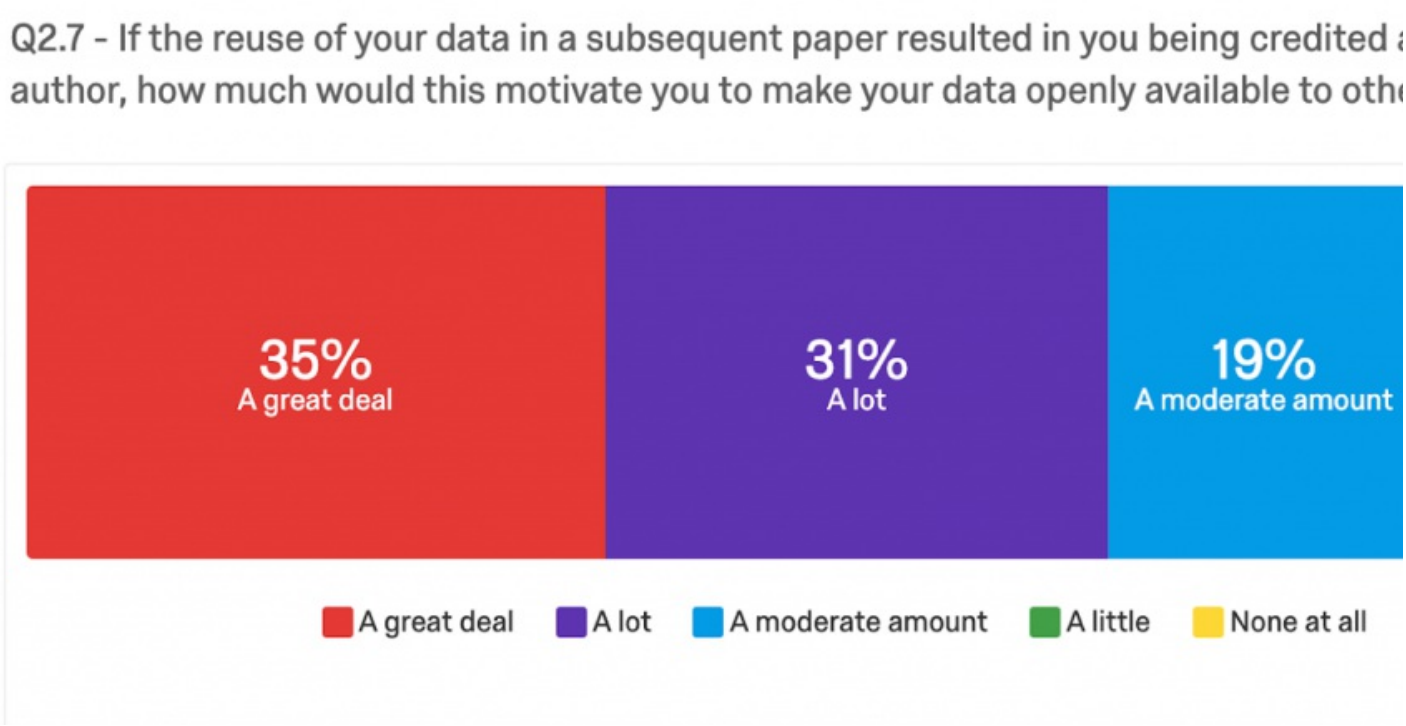
The report was born out of what was originally a survey of Figshare users on their attitudes to research data and open access. We used this data to inform our product development to prioritise the features that would help alleviate some of their biggest issues and concerns and refine the features we were developing. But, by only surveying Figshare users there was an implicit bias. We soon realised that the average Figshare user was familiar with the open research landscape, generally receptive to new technologies, new workflows and advocates of open data, which in 2015 was not a fair and accurate representation of the 'mainstream'.

In 2019, we were able to reach over 8500 researchers across different geographies, disciplines and at various levels of their career, creating a more representative data set through which to assess researcher attitudes to open data.

Motivations and Credit

One of the biggest fundamental issues in academia is the credit and reward system, only 11% of researchers surveyed believe they get enough credit for sharing data. Despite the diverse outputs being published such as datasets, code, media and models and the varying roles of academics from the traditional lab technician to computer scientists – the paper is still the primary output, citation counts the leading currency and journal impact factor reigns supreme.

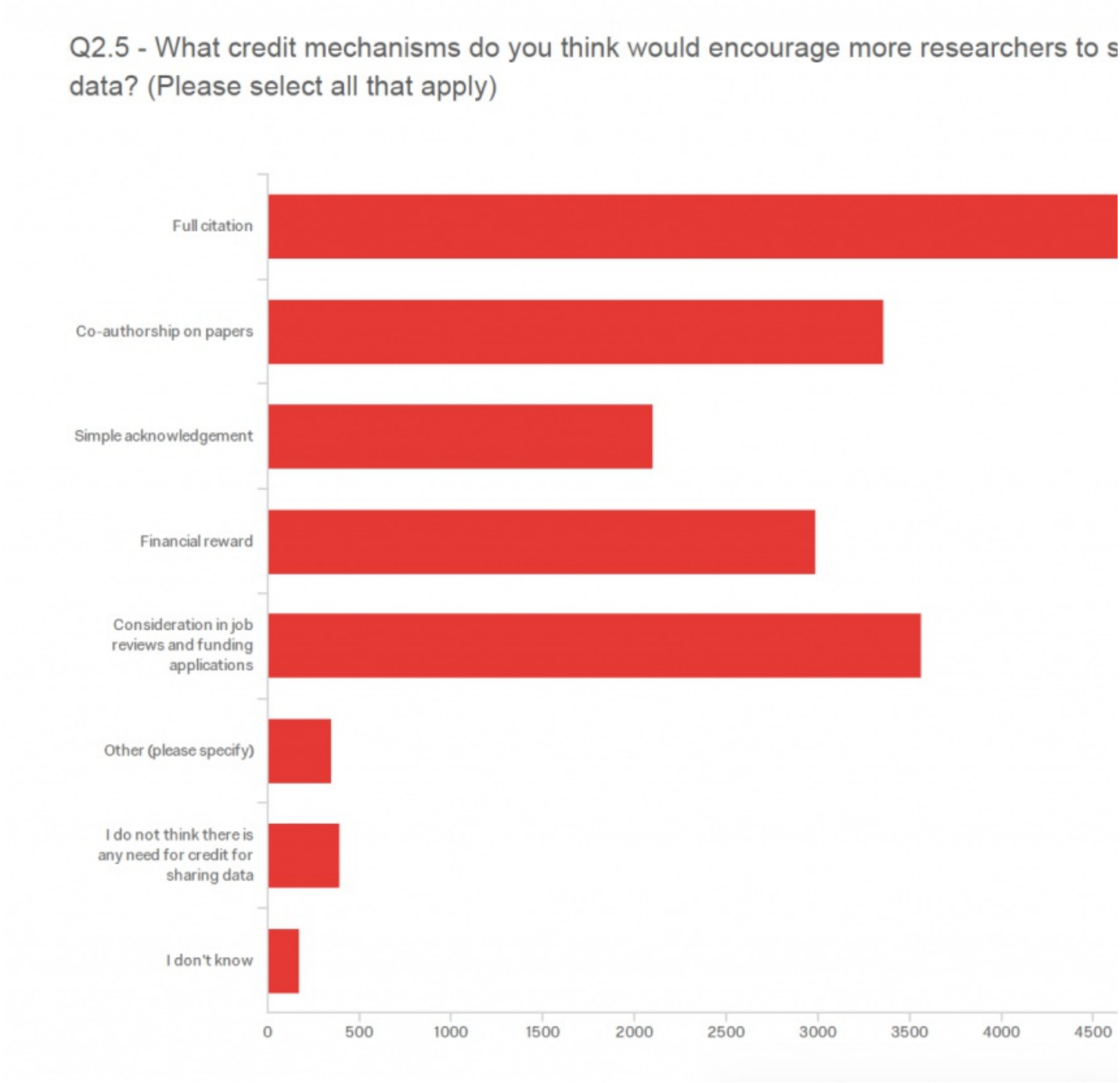
There is growing evidence that making data openly available increases the citation counts on publications, but the *State of Open Data 2019* showed researchers are using data to boost the number of papers they feature on, by leveraging the use of this data in return for authorship on papers, thus further buying into this broken system. Eight per cent of researchers stated that getting a co-authorship on a paper wouldn't motivate them to share their data. Holding data to ransom to get a co-authorship or financial reward, is a disturbing new trend, better credit mechanisms could help stem this behaviour and promote a more intrinsic positive culture around data sharing.



Date originally posted: 2019-12-13

Permalink: https://blogs.lse.ac.uk/impactofsocialsciences/2019/12/13/the-state-of-open-data-2019-what-are-the-key-issues-in-open-data-for-researchers/

Blog homepage: https://blogs.lse.ac.uk/impactofsocialsciences/



One encouraging final thought is that co-authorship is not the biggest rewarding factor for academics sharing their research data. For the fourth year in a row (ie. since the *State of Open Data* survey and report was initiated), citations are seen as the holy grail in terms of reward. Once again, we see that there are more citations to datasets by Open Access Week in October 2019 than there were in the whole of 2018. The rewards are growing, the incentives are increasing, and the mandates are expanding. It is an exciting time to start investigating what re-use of data looks like, and investigating the hypothesis that better described data leads to more re-use, which in turn, leads to more rewards for academics, and ultimately more efficient research for humanity. Who does that curation is an open question, and one that I hope we will have a much clearer answer to by 2020.

Date originally posted: 2019-12-13

Permalink: https://blogs.lse.ac.uk/impactofsocialsciences/2019/12/13/the-state-of-open-data-2019-what-are-the-key-issues-in-open-data-for-researchers/

Blog homepage: https://blogs.lse.ac.uk/impactofsocialsciences/

The State of Open Data report is available to view and download on Figshare along with the raw data and visualisation.

About the author

Mark Hahnel *is founder and CEO of Figshare. He is passionate about open science and the potential it has to revolutionise the research community. He is on twitter* @MarkHahnel

Note: This article gives the views of the authors, and not the position of the LSE Impact Blog, nor of the London School of Economics. Please review our comments policy if you have any concerns on posting a comment below.

Featured Image Credit adapted from, Federico Beccari via Unsplash (Licensed under a CC0 1.0 licence)

Date originally posted: 2019-12-13

Permalink: https://blogs.lse.ac.uk/impactofsocialsciences/2019/12/13/the-state-of-open-data-2019-what-are-the-key-issues-in-open-data-for-researchers/

Blog homepage: https://blogs.lse.ac.uk/impactofsocialsciences/