

New gold standard established for open and reproducible research

May 3 2015

A group of Cambridge computer scientists have set a new gold standard for openness and reproducibility in research by sharing the more than 200GB of data and 20,000 lines of code behind their latest results - an unprecedented degree of openness in a peer-reviewed publication. The researchers hope that this new gold standard will be adopted by other fields, increasing the reliability of research results, especially for work which is publicly funded.

The researchers are presenting their results at a talk today at the 12th USENIX Symposium on Networked Systems Design and Implementation (NSDI) in Oakland, California.

In recent years there's been a great deal of discussion about so-called 'open access' publications - the idea that [research publications](#), particularly those funded by [public](#) money, should be made publicly available.

Computer science has embraced open access more than many disciplines, with some publishers sub-licensing publications and allowing authors to publish them in open archives. However, as more and more corporations publish their research in academic journals, and as academics find themselves in a 'publish or perish' culture, the reliability of [research results](#) has come into question.

"Open access isn't as open as you think, especially when there are corporate interests involved," said Matthew Grosvenor, a PhD student

from the University's Computer Laboratory, and the paper's lead author. "Due to commercial sensitivities, corporations are reluctant to make their code and [data sets](#) available when they publish in peer-reviewed journals. But without the code or data sets, the results are irrelevant - we can't know whether an experiment is the same if we try to recreate it."

Beyond [computer science](#), a number of high-profile incidents of errors, fraud or misconduct have called quality standards in research into question. This has thrown the issue of reproducibility - that a result can be reliably repeated given the same conditions - into the spotlight.

"If a result cannot be reliably repeated, then how can we trust it?" said Grosvenor. "If you try to reproduce other people's work from the paper alone, you often end up with different numbers. Unless you have access to everything, it's useless to call a piece of research open source. It's either open source or it's not - you can't open source just a little bit."

With their most recent publication, Grosvenor and his colleagues have gone several steps beyond typical [open access](#) standards - setting a new gold standard for open and reproducible research. All of the experimental figures and tables in the award-winning final version of their paper, which describes a new method of making data centres more efficient, are clickable.

By clicking on any of the figures or tables in the paper, readers are taken to a website where the researchers have produced technically detailed descriptions of the methods for every one of their experiments. These descriptions include the original data sets and tools that were used to produce the figures as well as free and open source access to all of the source code that they wrote and modified.

In the past this might not have been possible, but thanks to cheap cloud storage, the researchers have put nearly 200GB of data and 20,000 lines

of code on to the internet and made it freely available to all under a permissive [open-source](#) license.

"It now should be possible for anyone with a collection of computers to follow our instructions and produce our exact graphs," said Grosvenor.

"We think that this is the way forward for all scientific publications and so we've put our money where our mouth is and done it."

More information: Queues don't matter when you can JUMP them! the 12th USENIX Symposium on Networked Systems Design and Implementation (NSDI), 2015.

Provided by University of Cambridge

Citation: New gold standard established for open and reproducible research (2015, May 3)
retrieved 24 April 2024 from <https://phys.org/news/2015-05-gold-standard.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--