Open and Reproducible Research with community-driven Research Software metrics

Emmy Tsang @emmy_ft
Yo Yehudi @yoyehudi
Mateusz Kuzak @matkuzak
How is open source software used in research and science?
Open source software in research

The first image of a black hole!

Open source software in research

https://setiathome.berkeley.edu/
How many of you have worked in science or research before?
Metrics in research
Metrics in research

As well as all of the metrics you might see in other open source orgs or companies, research has other pressures
A traditional output: The research paper

- Researchers typically write up their results in peer-reviewed papers, published in research journals.
- This is a format well-suited to writing up a lab experiments (and terrible for summing up/describing software)!
A traditional output: The research paper

- Software doesn’t look much like a paper, but we end up shoehorning it in to papers anyway.
- What happens after we’ve updated our software? More papers about software updates?
A traditional output: The research paper

- It’s possible to publish a paper about software without showing the software

“For this review, unfortunately I did not have access to a computer with the correct operating system, so I assume that the application functions as described”
A traditional metric: Citations

**Citation:** The number of times a given paper has been cited by other papers.

**Problem**

People often don’t think/remember/know how to cite software - leaving research software engineers worse off, due to their non-traditional outputs.
Other traditional metrics

Journal / conference prestige
# of papers

⛔ Problem ⛔

Can these research paper-based metrics accurately reflect the quality of the research software?
Another metric: Research Excellence!

In the UK, the Research Excellence Framework asserts:

“In assessing outputs, the sub-panels will look for evidence of originality, significance and rigour”

So, what does “originality, significance and rigour” mean in software?
CHAOSS + Research metrics - how do we measure?

Many/most of the CHA OSS metrics apply to research quite clearly, but some are company/organisation specific and may not be aware of grant funding models (as opposed to corporate sponsorship) and research environment pressures.
In groups of 2-4, discuss one or both of:
- Which CHA OSS metrics could be useful for research software
- Which other metrics might be useful, even if not CHA OSS-ey?

Go to http://sli.do (enter code chaoscon2020) - enter ideas / thoughts real-time!
Results

Go to http://sli.do (enter code chaosscon2020) - enter ideas / thoughts real-time!

- To follow up or view results, visit http://bit.ly/CHA OSScon-2020-research
Thank You

Go to http://sli.do (enter code chaosscon2020) - enter ideas / thoughts real-time!

View sli.do results & get involved by going to http://bit.ly/CHAOSScon-2020-research
Mateusz Kuzak, Research Software Community Manager, the Netherlands eScience Center, t @matkuzak g @mkuzak m.kuzak@esciencecenter.nl

Emmy Tsang, Innovation Community Manager @ eLife t @emmy_ft / @eLifeInnovation g @emmyft e.tsang@elifesciences.org

Yo Yehudi, Software Developer, University of Cambridge, & EngD student, University of Manchester t @yoyehudi g @yochannah yy406@cam.ac.uk