

Overview of research software funding landscape

By [Michelle Barker](#) and [Daniel S. Katz](#)

DOI: <https://www.doi.org/10.5281/zenodo.6102487>

This report aims to provide a brief overview of the funding landscape for research software throughout its life cycle. This report explains the methodology, considers available information, and suggests future work.

Methodology

There are multiple ways in which analysis of research software funding could be undertaken, such as by type of funder¹, by country², research discipline³, type of output⁴, etc. Life cycle stages were chosen here as analyses have repeatedly identified challenges in obtaining funding for the life cycle stage focused on maintaining research software (Knowles, Mateen, and Yehudi 2021; Sufi et al. 2020), yet there is little international data to support this.

The following key terms are defined as follows:

- **Research software** - Source code files, algorithms, scripts, computational workflows and executables that were created in either of two categories:
 - A. Within a research project as a by-product to do the research, or
 - B. Through intentional development of a software product for general use in research by one or more projects.

Software components (e.g., operating systems, libraries, dependencies, packages, scripts, etc.) that are used in research but were not created during or with a clear research intent should be considered software in research and not research software (adapted from Gruenpeter et al. 2021).

- **Support** - Contributions of monetary support. This excludes non-monetary grants, gifts, support services, volunteers, etc.
- **Funder** - A supporter providing monetary support without any exchange of services or other tangible benefits (adapted from Dunks 2021b).

¹ Invest in Open Infrastructure are utilising this approach for their analysis of open infrastructure providers (Dunks 2021a).

² This approach is commonly used for analysing investment in research infrastructures (e.g., Barker et al. 2019; Ficarra et al. 2020).

³ This framework is often applied to examinations of research infrastructure funding (e.g., Asmi et al. 2019).

⁴ Open hardware funding can be classified as supporting outputs in four categories: community-related outputs, documentation, hardware and usage or adoption (GOSH 2021b).

- **Research software life cycle / Types of funding** - Comprises the following stages (adapted from Katz et al. 2019; GOSH 2021a; ESFRI 2018):
 1. Research and development: The initial period of taking a concept to create actual software that can be used.
 2. Maintenance and support: Once software exists, it must be maintained to prevent software collapse (responding to changes in underlying software and hardware) and to keep it useful (fixing bugs, adding new features). Most software also must provide some level of support to new and existing users, who otherwise would run into problems that would lead them to stop using it. (Note that while many funders view their work this way, supporting a combination of maintenance, support, and new development for established software, some separate this into two distinct areas: 1. maintenance and support of software to keep it useful for existing purposes, and 2. new development to expand its utility to new purposes (Anzt et al. 2021).)

Research software funding

This section identifies funding available for the two phases of the research software life cycle, and related funding programs.

Life cycle phase 1: Research and development

- Category A: Research software developed within a research project as a by-product to do the research

Most funding for research software fits in this category; however, it is difficult to quantify the amount of investment. The funding quantum is certainly significant, with research showing that ~20% of National Science Foundation (NSF) projects (totalling USD\$10b) over 11 years discussed software in their abstracts (Katz 2021), software-intensive projects are a majority of current publications (Nangia and Katz 2017), and 33% of research produces new code (Bello and Galindo-Rueda 2020, fig 3.4). General research funding does not usually name research software as a potential output, although there is increasing inclusion of research data, which include mention of tools for data analysis (which can include research software).

- Category B: Research software intentionally developed as a software product for general use in research by one or more projects.

Funding for research software through this mechanism is provided by some funders. However, it can be difficult to identify these as opportunities for research software as the funding is often provided within broader frameworks, such as digital infrastructure, technology innovation, and open science. Recent examples include:

- Australian Research Data Commons (ARDC): [Platforms co-investment program](#)

- Dutch Research Council (NWO): [Open science fund](#)
- European Commission: Horizon Europe [Research infrastructures](#)
- German Research Foundation (DFG): [Qualitätssicherung von Forschungssoftware durch ihre nachhaltige Nutzbarmachung](#)
- Netherlands eScience Center (NLeSC): [Open eScience](#)
- Nordic e-Infrastructure Collaboration (NeIC): [eInfrastructure collaboration](#)
- SAGE Publishing: [Concept grants](#)
- UK Research and Innovation (UKRI): [Transformative research technologies](#)
- US NSF: [Cyberinfrastructure for Sustained Scientific Innovation \(CSSI\)](#)
- US National Institutes of Health (NIH): [Software tools for open science](#)
- Wellcome Trust: [Data for science and health](#)

Life cycle phase 2: Maintenance and support

Funding for maintenance and support of research software comprises a very short list:

- CANARIE, Canada: [Research software platforms](#)
- Chan Zuckerberg Initiative (CZI): [Essential Open Source Software for Science](#)
- German Research Foundation (DFG): [Research Software Sustainability](#)
- UKRI: [Software for research communities](#)
- US National Science Foundation (NSF): [CSSI Transition to sustainability](#)
- US NASA: [Support for Open Source Tools, Frameworks, and Libraries](#)

There are also occasional examples of programs that provide non-financial support, such as the NLeSC's [Small-Scale Initiatives in Software Performance Optimisation](#) which provides expertise.

Related funding

There are other types of funding that are relevant to the research life cycle, where the emphasis is on improving the environment and method for software development, not its actual creation. In Germany there have been calls for support for Infrastructure for developing, testing, validating, and benchmarking research software, and distributed versioning systems for collaborative software development (Anzt et al. 2021). Funding for information infrastructures for research software can support platforms, repositories and ecosystems to share and collaboratively develop research software as well as to make research software findable, accessible, interoperable and reusable (FAIR). This has overlaps with funding with a focus on addressing challenges and changing practices in software development, productivity and sustainability, e.g., US Department of Energy (DoE): [Interoperable Design of Extreme-scale Application Software \(IDEAS\)](#). This includes increasing investment in research issues that impact the development of exascale computing, e.g., UKRI: [Cross-cutting research for exascale software and algorithms](#).

There are a range of other funding types that support the broader ecosystem. Also important is funding for the people who develop and maintain research software, and the environment and culture in which research software is developed to change practices and norms. This not covered in this report. It should be noted that if software matures over time then the role of the personnel responsible can change from researchers and research software developers to include other IT professionals (Katerbow and Feulner 2018).

Discussion

One of the disadvantages of the fact that the majority of research and development funding is for category A is that there is a lack of focus on research software as an output. Consequently, policies to support the development of quality, reproducible, findable and sustainable research software are not applied. And because software is not an intentional output, the skills needed to develop good software are also not widely developed, or even considered when funding decisions are made. Those applicants who do have good skills and choose to develop good software can be at a disadvantage if these characteristics are not part of funding decisions. Finally, where software is not consciously considered, and recognition of software skills and knowledge does not exist, there is an added danger of duplicate software being developed because of relative ignorance of existing software, leading to a waste of scarce funds. Research and development funding for category B research software has the advantage that it improves recognition of the importance of research software, and usually requires that relevant policies are followed, such as version control, licensing, documentation, using demonstrated software engineering practices, code review, etc.

Whilst there are some funding programs in Category B to support the research and development phase, there is much more limited funding for maintenance and support. In part, this is due to an emphasis on novel research results over infrastructure that supports research. In this vein, the Technology Association of Grantmakers recommends that grantmakers reject the mindset of “technology as overhead” and increase investment in digital infrastructure for civil society (2021). The more limited funding for maintenance and support also reflects that funding is generally tied to short term cycles, not the long term maintenance needs of infrastructure. This is particularly true for software, which must be continuously invested in (though sometimes at a very low level) for it to continue to function. High profile examples of the problems that can occur without sufficient maintenance include the Heartbleed bug and Log4j (*BBC News* 2014; Saarinen 2021).

The significant gap in provision of the amount of funding for maintenance and support that current levels of research and development funding results in repeated funding cliffs for research software when development funding runs out. This finding is also similar to that of other digital infrastructure; analysis of open hardware funding found that most funders do not currently provide support for open hardware projects further than the research and development stage (GOSH 2021b). In comparison, high performance computing (HPC) systems often have a fairly short maximum lifetime of 5-10 years, and are typically fully supported during that period,

then not supported when they are no longer useful. Ongoing support is particularly important for research software, which has a stronger need than research data infrastructure because of software collapse (Hinsen 2019). And where funding for maintenance and support does exist, it is still usually funded for periods that do not match the timescales of the software's use. It also may be provided by a more limited set of funders: analysis of the UK research infrastructure landscape found that 30% of research infrastructures relied on public sources for at least 90% of their operational funding, with an additional 17% having a 71-90% reliance (UKRI 2020, fig 6.5).

Future work

There are a range of areas where future work could facilitate improved research software investments.

1. Increase awareness of the need to resource all life cycle stages

It would be useful to increase understanding of the need for resourcing maintenance and funding across a wider group of stakeholders, to assist in enabling sustainability of the research software enabled by research and development funding.

2. Improve understanding of the gaps in the investment landscape

More detailed data is needed to allow quantitative analysis of levels and trends in research software funding could support understanding of how to improve funding approaches. The current quantum of investment in research and development resulting in new research software is needed to understand the amount needed to provide maintenance and support. Efforts to quantify investment data in similar areas that could inform this include Invest in Open Infrastructure's mapping of the costs of open infrastructure (Dunks 2021a), and Simply Secure's work to identify open source software funders (Huerta 2021).

3. Identify a range of models for enabling sustainability

Some of the challenges in funding research software reflect the general issues faced by digital infrastructures that are public goods; many of the resources are developed collaboratively, and as the community continues to widen it is difficult to identify where ongoing support should come from. Open source scientific software is particularly complex, as it is potentially supported by significant unpaid volunteer labour, as is the case for open source software (Eghbal 2016). Research institutions are only just beginning to recognize research software as an asset that needs management. Research data management has faced a similar challenge, and it has been suggested that 5% of overall research costs should go towards data stewardship (Mons 2020). The [Critical Digital Infrastructure Research](#) grants provided by Ford Foundation, Sloan Foundation, Mozilla, Omidyar Network and Open Society Foundations aim to fill gaps in understanding of how digital infrastructure is built, maintained, and sustained; and similar understanding would be useful for research software.

Acknowledgements

This work was supported by the Alfred P. Sloan Foundation, grant G-2021-17001.

References

- Anzt, Hartwig, Felix Bach, Stephan Druskat, Frank Löffler, Axel Loewe, Bernhard Y. Renard, Gunnar Seemann, et al. 2021. "An Environment for Sustainable Research Software in Germany and beyond: Current State, Open Challenges, and Call for Action." *F1000Research* 9 (January): 295. <https://doi.org/10.12688/f1000research.23224.2>.
- Asmi, Ari, Lorna Ryan, Emmanuel Salmon, Christine Kubiak, Serena Battaglia, Miriam Förster, Julie Dupré, et al. 2019. "International Research Infrastructure Landscape 2019." <https://doi.org/10.5281/ZENODO.3539254>.
- Barker, Michelle, Silvia Delgado Olabarriaga, Nancy Wilkins-Diehr, Sandra Gesing, Daniel S. Katz, Shayan Shahand, Scott Henwood, et al. 2019. "The Global Impact of Science Gateways, Virtual Research Environments and Virtual Laboratories." *Future Generation Computer Systems* 95 (June): 240–48. <https://doi.org/10.1016/j.future.2018.12.026>.
- BBC News*. 2014. "Tech Giants Spend Millions to Stop Another Heartbleed," April 25, 2014, sec. Technology. <https://www.bbc.com/news/technology-27155946>.
- Bello, Michael, and Fernanda Galindo-Rueda. 2020. "Charting the Digital Transformation of Science: Findings from the 2018 OECD International Survey of Scientific Authors (ISSA2)." OECD Science, Technology and Industry Working Papers 2020/03. <http://www.oecd.org/digital/charting-the-digital-transformation-of-science-1b06c47c-en.htm>.
- Dunks, Richard. 2021a. "Exploring Costs & Characteristics of Open Infrastructure Providers." Invest in Open Infrastructure. November 1, 2021. <https://investinopen.org/blog/costs-characteristics-oi-providers/>.
- . 2021b. "Funding Open Infrastructure: Key Terms and Concepts in Our Analysis." Invest in Open Infrastructure. November 5, 2021. <https://investinopen.org/blog/funding-open-infrastructure-key-terms-and-concepts-in-our-analysis/>.
- Eghbal, Nadia. 2016. "Roads and Bridges: The Unseen Labor Behind Our Digital Infrastructure." Ford Foundation. 2016. <https://www.fordfoundation.org/work/learning/research-reports/roads-and-bridges-the-unseen-labor-behind-our-digital-infrastructure/>.
- ESFRI. 2018. "Roadmap 2018: Strategy Report on Research Infrastructures." <http://roadmap2018.esfri.eu/>.
- Ficarra, Victoria, Mattia Fosci, Andrea Chiarelli, Bianca Kramer, and Vanessa Proudman. 2020. "Scoping the Open Science Infrastructure Landscape in Europe." Zenodo. <https://doi.org/10.5281/ZENODO.4159838>.
- GOSH. 2021a. "Funding Open Hardware: Institutional Support." 2021. <https://drive.google.com/file/d/1H6YtXsojx9oFmSVbIAiq7J87TI7xIM1V/view>.
- . 2021b. "Funding Open Hardware: Outputs and Impacts." 2021. https://drive.google.com/file/d/1QWEJ1hSZ4dwrgZKRhgyp3HTAvZ4i0YFa/view?usp=sharing&usp=embed_facebook.
- Gruenpeter, Morane, Katz, Daniel S., Lamprecht, Anna-Lena, Honeyman, Tom, Garijo, Daniel, Struck, Alexander, Niehues, Anna, et al. 2021. "Defining Research Software: A

- Controversial Discussion.” Zenodo. <https://doi.org/10.5281/ZENODO.5504016>.
- Hinsen, Konrad. 2019. “Dealing With Software Collapse.” *Computing in Science & Engineering* 21 (3): 104–8. <https://doi.org/10/gf2dh9>.
- Huerta, Melissa. 2021. “Building a Toolkit for Funders to Grow Their Digital Infrastructure Portfolio.” 2021. <https://simplysecure.org/blog/building-a-toolkit-for-funders-to-grow-their-digital-infrastructure-portfolio/>.
- Katerbow, Matthias, and Georg Feulner. 2018. “Recommendations On The Development, Use And Provision Of Research Software,” March. <https://doi.org/10.5281/ZENODO.1172988>.
- Katz, Daniel S. 2021. “Towards Sustainable Research Software.” December 1. <https://doi.org/10.5281/zenodo.5748175>.
- Katz, Daniel S., Stephan Druskat, Robert Haines, Caroline Jay, and Alexander Struck. 2019. “The State of Sustainable Research Software: Results from the Workshop on Sustainable Software for Science: Practice and Experiences (WSSSP5.1).” *Journal of Open Research Software* 7 (April): 11. <https://doi.org/10.5334/jors.242>.
- Knowles, Rebecca, Bilal A. Mateen, and Yo Yehudi. 2021. “We Need to Talk about the Lack of Investment in Digital Research Infrastructure.” *Nature Computational Science* 1 (3): 169–71. <https://doi.org/10.1038/s43588-021-00048-5>.
- Mons, Barend. 2020. “Invest 5% of Research Funds in Ensuring Data Are Reusable.” *Nature* 578 (7796): 491–491. <https://doi.org/10.1038/d41586-020-00505-7>.
- Nangia, Udit, and Daniel S. Katz. 2017. “Understanding Software in Research: Initial Results from Examining Nature and a Call for Collaboration.” In *2017 IEEE 13th International Conference on E-Science (e-Science)*, 486–87. <https://doi.org/10/ggfkvb>.
- Saarinen, Juha. 2021. “Log4j’s Project Sponsorship Skyrockets after Critical Bug Exploitation.” ITnews. 2021. <https://www.itnews.com.au/news/log4js-project-sponsorship-skyrockets-after-critical-bug-exploitation-573914>.
- Sufi, Shoaib, Carlos Martinez Ortiz, Cees Hof, Patrick Aerts, Adriaan Klinkenberg, Anna-Lena Lambrecht, Barbara Sierman, et al. 2020. “Report on the Workshop on Sustainable Software Sustainability 2019 (WOSS19).” Zenodo. <https://doi.org/10.5281/ZENODO.3922155>.
- Technology Association of Grantmakers. 2021. “A Responsibility to Rebuild: Investing in Digital Infrastructure for Civil Society.” https://cdn.ymaws.com/www.tagtech.org/resource/resmgr/digital_infrastructure/digital_infrastructure-repor.pdf.
- UKRI. 2020. “The UK’s Research and Innovation Infrastructure: Landscape Analysis.” <https://www.ukri.org/wp-content/uploads/2020/10/UKRI-201020-LandscapeAnalysis-FINAL.pdf>.