Raising the Profile of Research Software: Recommendations for Funding Agencies and Research Institutions

Authors (alphabetically):

Anton Akhmerov
Maria Cruz
Niels Drost
Cees Hof
Tomas Knapen
Mateusz Kuzak
Carlos Martinez-Ortiz
Yasemin Turkyilmaz-van der Velden
Ben van Werkhoven

Disclaimer: All opinions expressed in this paper are the author's own opinions and do not necessarily reflect the official policy or position of their institutions.

Abstract

Research software is fundamental to contemporary research, yet it does not receive the recognition it rightfully deserves. This needs to change, particularly in the context of the discussions around open science and reproducibility. We argue that if open science is to truly lead towards better, more transparent, and reproducible research, then research software needs to be treated in equal footing to research data and publications at the policy level.

In this paper, we present a concrete followup to the DORA declaration in the form of recommendations for raising the profile of research software. We divide our recommendations into four categories: Software availability and quality, Software sustainability, Training, and Human capital.

These recommendations provide steps for achieving recognition for research software as a fundamental and vital component of research.

Introduction

"Without data it's difficult to validate results. But without code, we waste the opportunity to advance science."

Neil Chue Hong
Software Sustainability Institute

Research software is fundamental to contemporary research. In most fields of research it is impossible to conduct research without software.[1] The announcement in April 2019 of the first image of a black hole[2], making front page news around the world, is a good case in point. The landmark image in Figure 1 could not have been made without research software, both the standard packages already in use in the field as well as specially developed algorithms to process and analyse the data coming from several radio telescopes around the world.

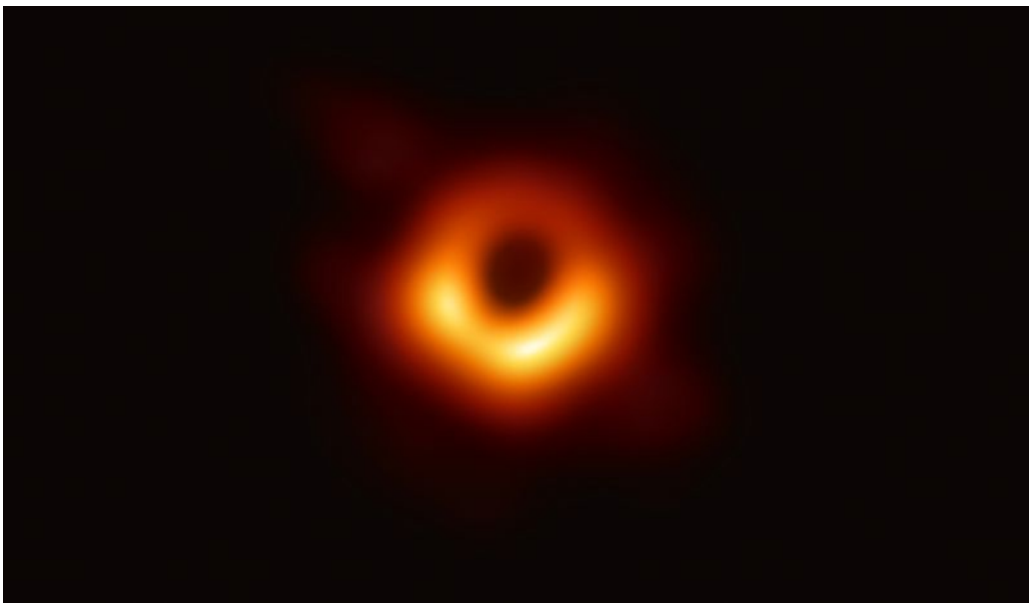

Figure 1: The black hole at the center of the galaxy Messier 87 as revealed by the Event Horizon Telescope. This image would not be possible without research software in its various guises. Image credit: Event Horizon Telescope Collaboration.

[1] "It's impossible to conduct research without software, say 7 out of 10 UK …," 4 Dec. 2014, https://www.software.ac.uk/blog/2014-12-04-its-impossible-conduct-research-without-software-say-7-out-10-uk-researchers. Accessed 27 Apr. 2019.
[2] "Focus on the First Event Horizon Telescope Results - IOPscience," Apr. 2019, https://iopscience.iop.org/journal/2041-8205/page/Focus_on_EHT. Accessed 27 Apr. 2019.

Yet research software does not receive the recognition it deserves. Research software is not adequately recognised in the scholarly record through citation.[3,4] There is a lack of funding and incentives for those who develop research software.[5] The acknowledgment of the importance of research software as an independent research output lags behind that of research data and research publications.[6] In terms of the incentives and rewards structure, both software and data contributions are considered second-rate relative to publications.[7]

This needs to change, particularly in the context of the discussions around open science and reproducibility.[8] These discussions, which gained much prominence in recent years, have all largely focused on open access publications and FAIR data.[9] As a result, open data and open access publishing are becoming mainstream in many fields of research.

We argue that **if open science is to truly lead towards better, more transparent, and reproducible research, then research software needs to be treated in equal footing to research data and publications at the policy level and in practice**. There cannot be fully open, transparent and reproducible research without open and sustainable research software. Such software requires, in turn, appropriate funding, incentives, recognition and rewards to be in place. Without these, It is difficult to sustain and maintain the algorithms and research software that are crucial for scientific advances.

The Declaration on Research Assessment (DORA) recognizes the need to improve scholarly research evaluation. It recognizes data and software as being just as important as research articles. By signing DORA NWO, ZonMW, KNAW and others have committed to taking steps in changing the scholarly research evaluation criteria.[10,11]

---

[3] J. Howison and J. Bullard, "Software in the scientific literature: Problems with seeing, finding, and using software mentioned in the biology literature," Journal of the Association for Information Science and Technology, vol. 67, no. 9, pp. 2137–2155, May 2015. https://doi.org/10.1002/asi.23538

[4] L. A. Barba et al., "Giving software its due through community-driven review and publication," Apr. 2019. https://doi.org/10.31219/osf.io/f4vx6

[5] A. Siepel, "Challenges in funding and developing genomic software: roots and remedies," Genome Biology, vol. 20, no. 1, Jul. 2019. https://doi.org/10.1186/s13059-019-1763-7

[6] "Making Software a First-Class Citizen in Research | Software …," 28 Nov. 2018, https://software.ac.uk/blog/2018-11-28-making-software-first-class-citizen-research. Accessed 27 Apr. 2019.

[7] "Why we need a hub for software in science: Research software …," 17 Dec. 2015, https://blogs.lse.ac.uk/impactofsocialsciences/2015/12/17/why-we-need-a-hub-for-software-in-science/. Accessed 28 Apr. 2019.

[8] "Improving Reproducibility in the Empirical Sciences, Koninklijke Nederlandse Akademie van Wetenschappen," 2018, ISBN 978-90-6984-720-7.

[9] M. D. Wilkinson et al., "The FAIR Guiding Principles for scientific data management and stewardship," Sci. Data , vol. 3, Mar. 2016. https://doi.org/10.1038/sdata.2016.18

[10] "KNAW, NWO and ZonMw to sign DORA declaration," 18 Apr. 2019, https://www.nwo.nl/en/news-and-events/news/2019/04/knaw-nwo-and-zonmw-to-sign-dora-declaration.html. Accessed 13 Jun. 2019.

[11] "Evolution or revolution?," April 2019, https://publicaties.zonmw.nl/evolution-or-revolution/, Accessed 13 Jun. 2019.

In this paper, we present a concrete follow-up to the DORA declaration in the form of recommendations for raising the profile of research software.

Recommendations and rationale

We provide recommendations which funding agencies and research institutions can directly implement. We divide our recommendations into four broad categories and provide the rationale which motivates the recommendations for each category.

Research policy is a complicated area with many different aspects to be considered. We focus on the four categories defined below because we consider that modest improvements in these areas would have a large impact on the research software landscape, and because these improvements can be achieved with relative ease.

Recommendations

Software availability and quality

- Require that software generated in research projects is archived and published in compliance with the FAIR principles, in repositories that facilitate citation, long term preservation and accessibility of the software.
- Encourage that all software produced in research projects is published under an open source license.
- Research and academic institutions should provide the necessary means to store and share software during the active research phase and provide clear guidelines for the long-term storage and preservation of software.
- Encourage the development of software which adheres to good practices

Software sustainability

- Implement pilot calls to fund the revitalization, dissemination long-term maintenance, as well as further development of existing open source software packages important to science.
- Stimulate community building and organisation of workshops around important software packages.
- Survey the research software community to identify its needs for maintaining specific software packages.
- Require grant proposals to provide a “software footprint”—the list of software used in the proposed research project.

Training

- Funding agencies should encourage the inclusion of budget in research proposals for training on digital skills. Such training activities could aim at knowledge development, skill development, research software best practices, data analysis skills, etc.
- Research and academic institutions should include basic digital competency skills in the core curricula of academic institutions at both undergraduate and graduate level.
- Research institutes should provide support and guidance on digital competency.
- Stimulate the training of specialists on digital competency skills (Research Software Engineers and Data Stewards).

Human capital

- Stimulate the formalisation of Research Software Engineer and Data Steward roles.
- Stimulate the recognition and rewarding of all research outputs (including datasets and software) in addition to research publications by requiring to list software (including software maintenance, contributing to software communities, etc.) in CVs, alongside other research outputs, when evaluating research grants.

Rationale

Software availability and quality

For research to be truly reproducible, it should be possible to re-run scientific experiments using the same data and the same software to produce the same results.

First and foremost this means the software should be available. Furthermore, software must be developed in such a way that it can be reused in the future. Following good software development practices helps with building software which is easier to preserve.

Funding agencies should encourage projects which develop software to adhere to good practices and point to the guidelines for such practices. These should be community agreed sets of recommendations which make sense for individual communities. Funding agencies should work together with specialist knowledge centers (Netherlands eScience Center, DLR, SSI, ReSA, CESSDA, etc.) in developing such guidelines and should facilitate scientists to work together with specialist knowledge centers in following these guidelines.

Existing guidelines:

- DLR guide
- eScience Center Software Development Guide
- SSI
    - Software Evaluation Guide

    - Online sustainability evaluation
- CLARIAH Software quality guidelines
- CESSDA's Software Maturity Model

Software sustainability

Due to its nature, software requires continuous maintenance in order to function properly. With technology constantly evolving—new hardware, new operating systems, new versions of libraries, etc.—software needs to be updated periodically to ensure that it still operates as expected, and to fix compatibility issues which might emerge. Frequently, these changes require very specific knowledge for that particular piece of software. However this type of maintenance is seldom taken into account during the initial development of the software and it is typically funded in an ad-hoc fashion. Not all software developed is worth maintaining in the long term, research communities have the implicit knowledge of which software packages are more frequently reused and therefore are worth the effort of keeping them alive.

Funding mechanisms for continuous maintenance of relevant software packages are necessary. Efforts for building communities around these software packages should be stimulated, as this has been shown to be an effective way to keep software alive.

Training

Programming has become an integral research tool in all branches of science. Its importance is comparable to that of the skills required in an experimental laboratory (e.g. pipetting or making dilutions). Programming skills are often concentrated in a few individuals with specialist training. Future researchers should acquire these skills early on as part of their training, and practicing researchers should have the opportunity to acquire these skills as part of their career development. For these skills to become more commonplace, it is important to stimulate knowledge and skills exchange activities (study groups/hacky hours etc.) inside and outside the institutions.

Within research projects it should be possible to spend time on knowledge and skills exchange activities -- there should be a possibility to have a WP dedicated to such activities (e.g. you should be able to assign a budget / number of hours for these activities). At the end of a project, these activities should be recognisable as valuable outputs of a research project (e.g. individuals should be able to put them on their CV). Grant reviewing process should take into account participation in training and mentoring activities as evaluation criteria.

Human capital

The Netherlands has been one of the leading European countries in training, hiring and embedding Data Stewards in research and academic institutions. Considering that data and

software are intrinsically linked[12] and that software is fundamental to research, having data skills and competencies alone are not enough. Research institutions need Research Software Engineers (RSEs)[13] working alongside Data Stewards. RSEs "closely collaborate with researchers to understand the challenges they face, and then develop research software to provide the answers"[14], "combine expertise in programming with an intricate understanding of research"[15] and "apply the skills and practices of software development to research to create a more robust, manageable, and sustainable research software."[16]

Data Stewards support researchers in activities ranging from daily research data management to compliance with funder and publisher requirements on data availability, as well as legal requirements related to GDPR and intellectual property rights. It is becoming clear for many research institutions in the Netherlands "that data stewards are indispensable"[17], and this has resulted in the establishment of data stewardship programmes, such as the one at TU Delft.[18,19] However, there is still no common job description and a universally accepted overview of the knowledge, skills and competencies.[20] Additionally, currently there is no UFO profile for Data Stewards, which is part of the function ordering system used at Dutch universities.[21] This is why there are ongoing efforts to define the tasks and roles of Data Stewards in Dutch research institutes. For example, a ZonMw funded project[22] has specifically focused on professionalising the Data Steward function within the life-sciences domain, by developing a function matrix with necessary tasks and responsibilities.[23,24]

---

[12] P. K. Doorn and P. Aerts, "A Conceptual Approach To Data Stewardship and Software Sustainability: Scientists in charge, with a little help from their friends," Data Archiving and Networked Services (DANS), 2016. http://hdl.handle.net/20.500.11755/59c24848-9cf7-437c-b2d5-e943e9e4a35e

[13] M. Teperek and A. Dunning, "Data (and code) roles of the future at TU Delft," 2019. https://doi.org/10.5281/zenodo.3256575

[14] "What is a Research Software Engineer?," 13 Jun. 2017, http://nl-rse.org/2017/06/13/what-is-rse.html, Accessed 3 Jun. 2019.

[15] "What is an RSE?," https://rse.ac.uk/what-is-an-rse/, Accessed 13 Jun. 2019.

[16] "What is an RSE?," https://us-rse.org/what-is-an-rse/, Accessed 13 Jun. 2019.

[17] I. Verheul, M. Imming, J. Ringerma, A. Mordant, J.-L. van der Ploeg, and M. Pronk, "Data Stewardship on the map: A study of tasks and roles in Dutch research institutes," 2019. https://doi.org/10.5281/zenodo.2669149

[18] "Role of Data Stewards and Data Stewardship Community," 4 January 2017, https://openworking.wordpress.com/2017/01/04/role-of-data-stewards-and-data-stewardship-community/ Accessed 13 Jun. 2019.

[19] M. Teperek, M. J. Cruz, E. Verbakel, J. Böhmer, and A. Dunning, "Data Stewardship addressing disciplinary data management needs," International Journal of Digital Curation, vol. 13, no. 1, pp. 141–149, Dec. 2018. https://doi.org/10.2218/ijdc.v13i1.604

[20] I. Verheul, M. Imming, J. Ringerma, A. Mordant, J.-L. van der Ploeg, and M. Pronk, "Data Stewardship on the map: A study of tasks and roles in Dutch research institutes," 2019. https://doi.org/10.5281/zenodo.2669149

[21] "Functie_ordeningsystem_UFO - VSNU," https://www.vsnu.nl/functie_ordeningsystem_ufo.html. Accessed 3 Jun. 2019.

[22] "Towards a community-endorsed data steward profession ... - Zenodo," 25 Jan. 2019, https://zenodo.org/communities/nl-ds-pd-ls/. Accessed 3 Jun. 2019.

[23] S. Scholtens et al., "Life sciences data steward function matrix," 2019. https://doi.org/10.5281/zenodo.2561722

Efforts of professionalising the data stewardship function is certainly promising, considering the significant demand for this role in the changing research and funding landscape, which requires an increasing amount of skills and capacities. It is essential that similar efforts for career professionalization are also spent for RSEs. As put in a recent report titled “Data stewardship on the map”[25], “Professional data stewardship needs professional data stewards”. The same is true for research software development, management and sustainability.

Final remarks

Software has become a crucial part of research, but it still does not receive the same recognition as other research outputs. There is a need to fully acknowledge that research software is as important as research data and scientific publications, as expressed in DORA. We share the Research Software Alliance’s vision that research software should be “recognised and valued as a fundamental and vital component of research.”[26]

We have provided recommendations which funding agencies and research institutions can implement to achieve this goal. These recommendations do not offer a full solution to addressing this issue, but provide a further step in the direction of achieving recognition for research software as a fundamental and vital component of research.

We believe that even minor policy improvements in the domain of research software will lead to visible improvement in science and suggest to funding and research institutions to adopt at least a part of our suggestions.

Acknowledgements

This paper follows on from a meeting the authors had with representatives from NWO (The Netherlands Organisation for Scientific Research) on 28 March 2019.[27] Our aim with that meeting was to highlight the importance of research software in contemporary research and its relationship to research data, open science, and reproducibility in research. We thank NWO for receiving us on 28 March 2019 and for the discussions we had then, which led to the creation of this paper.

---

[24] S. Scholtens et al., “Towards a community-endorsed data steward profession description for life science research - Poster: Function Matrix,” 2019. https://doi.org/10.5281/zenodo.2616476
[25] I. Verheul, M. Imming, J. Ringerma, A. Mordant, J.-L. van der Ploeg, and M. Pronk, “Data Stewardship on the map: A study of tasks and roles in Dutch research institutes,” 2019. https://doi.org/10.5281/zenodo.2669149
[26] “Research Software Alliance,” https://www.researchsoft.org/hestia-front/, Accessed 3 Jun. 2019.
[27] A. Akhmerov et al., “Making Research Software a First-Class Citizen in Research,” 2019. http://doi.org/10.5281/zenodo.2647436