PROJECT INFORMATION

Principal Investigator
Michelle Barker
Director, Research Software Alliance
Australia

Grantee Organization: Code for Science & Society
Amount Requested: $86,000
Requested Start Date: 1 Jan 2021
Requested End Date: 31 July 2021
Project URL: Researchsoft.org/taskforces

PROJECT GOAL

Research software is a fundamental and vital part of research worldwide, yet there remain significant challenges to software productivity, quality, reproducibility, and sustainability. Improving the practice of scholarship is a common goal of the open science, open source software and FAIR communities, but a focus on improving the management and sharing of research software is not yet a strong focus of the latter. This project aims to increase the FAIR community’s focus on improving software health through improved practices by applying the FAIR principles to research software. This project will achieve that goal by facilitating and driving the major work in this field, the FAIR For Research Software Working Group (FAIR4RS WG), to engage the research software community in the crucial first step of agreeing on the application of the FAIR principles to research software.

OBJECTIVES

This project seeks funding to ensure that the major FAIR4RS WG deliverable of a document developed with community support defining FAIR principles for research software is achieved by mid-2021. ReSA will provide the facilitation to ensure that this and preceding outputs are achieved in a timely manner, and with the extensive engagement with a broad range of research software community stakeholders that is necessary to ensure community support for the resulting outputs.

PROPOSED ACTIVITIES

The major activities will be the facilitation of FAIR4RS WG to develop outputs leading to and including the FAIR for research software principles, and webinars and conference sessions to engage the community with the development of these outputs.

EXPECTED PRODUCTS

This project will deliver a document summarising a community-agreed definition of the FAIR principles for research software by 30 June 2021, for broad dissemination (including publication).

EXPECTED OUTCOMES

The resulting adoption and implementation of FAIR software principles will create significant outcomes for many stakeholders, ranging from increased research reproducibility for research organisations, to clarity for funders around their own requirements for software investments, and guidelines for publishers on sharing requirements. This proposal aligns with the Alfred P. Sloan Foundation’s Data and Computational Analysis Research program in supporting the efficient management and sharing of research data and code from acquisition through analysis.
1. What is the main issue, problem, or subject and why is it important?

Improving the practice of scholarship is a common goal of the open science, open source software and FAIR communities, which each having slightly different aims: 1. Open science community: transparency, reproducibility, diversity and credit 2. Open source software community: collaboration, sustainability, diversity and credit 3. FAIR community: supporting open science goals by advancing research data practices. While the open science and open source communities have influenced each other (Tennant et al., 2020), the FAIR community has not been as involved in these interactions. There is significant potential to engage the FAIR community more fully in improving the management and sharing of research software. This project aims to achieve that by expanding application of the FAIR (Findable, Accessible, Interoperable, Reusable) principles beyond data into software, which is sufficiently different from data that new thinking and new work are needed. This work will tighten the alignment between open science and open source to increase the development and uptake of improved research software practices.

The FAIR principles have been applied extensively to research data since Wilkinson et al.’s publication in Nature in 2016 (Wilkinson et al., 2016), to maximize the discovery and reusability of research data to increase research impacts. With a recent OECD study showing that 25% of research produces new code (Bello & Galindo-Rueda, 2020, p. 22), it is increasingly urgent to engage the research software community in the crucial first step of agreeing on the application of the FAIR principles to research software, to create an agreed framework that the community can apply to improve research software practices. This project will achieve that goal by facilitating and driving the major work in this field, the FAIR For Research Software Working Group (FAIR4RS WG) (Research Data Alliance, 2020a).

The resulting adoption and implementation of FAIR principles for research software will create significant outcomes for many stakeholders, ranging from increased research reproducibility for research organizations, to clarity for funders around their own requirements for software investments, and guidelines for publishers on sharing requirements. FAIR4RS WG has been jointly convened as a Research Software Alliance (ReSA) Taskforce, Research Data Alliance (RDA) Working Group and FORCE11 Working Group, in recognition of the importance of this work for the advancement of the research sector. This project aligns with the Alfred P. Sloan Foundation’s Data and Computational Analysis Research program in supporting the efficient management and sharing of research data and code from acquisition through analysis.

2. What is the major related work in this field?

The FAIR principles are generally viewed as an effective tool for advancing research outcomes in all disciplines, and have resulted in numerous FAIR data investments, such as the European Commission’s €10 million, 3-year FAIRsFAIR initiative, and FAIRsharing (a community of 76 journals, funders, and databases. The 2019 European Union publication, Cost-benefit analysis for FAIR research data, estimates that the overall cost to the European economy of not having FAIR research data is €10.2 billion per year in Europe (European Commission et al., 2018, p.4), and FAIR lies at the heart of the European Commission’s European Open Science Cloud (EOSC) with their 2018 report, Turning FAIR into Reality, concluding that: “To take advantage of the digital revolution, to accelerate research ... while ensuring transparency, reproducibility and societal utility, data and other digital objects created by and used for research need to be FAIR” (European Commission, 2018, p. 8).
The EOSC Executive Board FAIR Working Group’s Six Recommendations for Implementation of FAIR Practice recognizes FAIR4RS WG as the community forum focused on application of the FAIR principles to research software (European Commission & EOSC Executive Board, 2020, p. 24). FAIR4RS WG has very strong community support with 135 members drawn from a diverse range of organizations internationally. FAIR4RS WG was convened to build on at least 19 works applying the FAIR principles to aspects of research software (Software Source Code IG, 2020) that had been initiated independently of each other, and consequently did not represent a coordinated body of work with broad community support. Without coordination of this work going forward, there would be significant potential for divergent renderings of the FAIR principles for software, duplication of effort, and gaps in consideration.

It is generally agreed that at the very least it will be necessary to revise and extend the original FAIR principles to apply them to research software, and Katz suggests that the goals for research software should go beyond FAIR to research software that is “open, FAIR, and citable” (Katz, 2017). The research challenges posed by COVID-19 is one of the many areas demonstrating the need to match the substantive efforts to increase open sharing of research data with similarly open research software, as part of the broader emphasis in increased use of open source software. As Lamprecht et al. and Katz’s analysis shows, there is clear alignment between the objectives of FAIR and open source software, but they are not the same (Katz, 2017; Lamprecht et al., 2020). Whilst the FAIR principles intentionally do not include open, the FAIR4RS WG focus on creating a consensus definition for FAIR for research software provides an opportunity to combine the push for FAIR software with a complementary effort for open source software, particularly because the advantages of making research software FAIR overlap with the benefits of increases in open source software.

3. Why is the proposer qualified to address the issue for which funds are being sought?

ReSA co-leads FAIR4RS WG in partnership with RDA and FORCE11. ReSA is well positioned to lead the broad community consultation required, with the ReSA mission being to bring research software communities together to collaborate on the advancement of research software. ReSA recently facilitated the software subgroup that made recommendations on software best practices to policy makers, funders, publishers and researchers as part of the RDA COVID-19 Recommendations and Guidelines for Data Sharing (Research Data Alliance, 2020b), coordinating 45 community members to work together in tight timelines. Application of FAIR to research software will represent a significant step forward in ReSA’s aims to support recognition and valuing of research software as a fundamental and vital component of research worldwide. ReSA is also leading a complementary project to roadmap a broad integration across the software landscape of the FAIR principles for software, once developed (Barker et al., 2020).

4. What is the approach being taken?

This project seeks funding to ensure that the major FAIR4RS WG deliverable of a document developed with community support defining FAIR principles for research software is achieved by

---

1 Pandemic response shines spotlight on coding in science (“Pandemic Response Shines Spotlight on Coding in Science,” n.d.); Call for transparency of COVID-19 models (Barton et al., 2020); Critiqued coronavirus simulation gets thumbs up from code-checking efforts (Chawla, 2020).

2 For example, Jimenez et al.’s seminal paper on encouraging best practices in research software contains recommendations that are designed around open-source values (Jiménez et al., 2017).
mid-2021, in line with the FAIR4RS WG project plan (FAIR4RS WG, 2020), ReSA will provide the facilitation to ensure that this and preceding outputs are achieved in a timely manner, and through engagement with the broad range of research software community stakeholders necessary to ensure community support for the resulting outputs. Creation of these outputs will require extensive community consultation but remains a volunteer effort of 135+ community members led by the 9-member FAIR4RS Steering Committee. The first FAIR4RS WG output due in November 2020 is already facing delays as the coordination of four subgroups involving 50 members is lagging because some subgroup leaders struggle to prioritize this volunteer work. This project would provide paid staff to work with Steering Committee members to ensure deadlines are met.

Dr. Michelle Barker, ReSA Director, will function as the project director and a community manager will be recruited, to support the FAIR4RS WG Steering Committee in this work. Project director responsibilities will include project leadership, reporting and reviews of project progress, staff management, coordination of completion of FAIR4RS outputs (including subgroup work), and liaison for FAIR4RS community consultations that other organizations are facilitating. Community manager responsibilities will include coordination of community consultation events that the project is convening directly, and project management of FAIR4RS to ensure outputs are developed in time. Dr Barker will report to the ReSA Steering Committee, who will oversee her work. The ReSA Steering Committee members will play a key role in the strategic planning for this work:

- Neil Chue Hong, Director - Software Sustainability Institute, University of Edinburgh, UK
- Catherine Jones, Energy Data Center Lead, Science & Technology Facilities Council UK
- Daniel S. Katz, Chief Scientist, National Center for Supercomputing Applications (NCSA), University of Illinois, USA
- Chris Mentzel, Executive Director, Data Sciences, Stanford Data Science Initiative, Stanford University, USA
- Karthik Ram, US Research Software Sustainability Institute (URSSI) Principal Investigator, University of California, Berkeley, USA
- Andrew Treloar, Director, Platforms and Software, Australian Research Data Commons

5. What will be the output from the project?

This project will deliver the following key outputs:

- Engagement plan detailing webinars, workshops and conference sessions to engage the community in developing these outputs throughout 2021 - 28 March 2021
- Document summarizing a community-agreed definition of the FAIR principles for research software, for broad dissemination (including publication) - 30 July 2021

The FAIRsFAIR draft report on FAIRness of software strongly reinforces the importance of gaining community agreed application of the FAIR principles to research software, and includes as its third recommendation guidance how this should occur: “a large community forum MUST be consulted when writing the [FAIR4RS] principles. This community forum MUST include stakeholders from different disciplines and with different roles, looking at software in all its aspects: as a tool, as a research outcome and as the object of research” (Gruenpeter, Morane
et al., 2020, p. 34). The FAIR4RS WG is already engaging in more than a dozen community events in late 2020 to promote this work; this will increase when consultation begins on the principles in 2021. This will include both events that FAIR4RS WG convenes directly, and FAIR4RS themed events that other organizations will facilitate, including Collaborations Workshop 2021 (CW21) and 2021 Workshop on Sustainable Software Sustainability (WoSSS).

The outcomes of the community consultations will be measured in terms of success in engaging the research software community to provide input into the drafting of the FAIR4RS principles, including stakeholders from a range of disciplines and roles. A valuable by-product of the engagement will be increased awareness of the importance of research software sustainability and sharing, and ways to increase this. An engagement plan will be developed to identify which communities will be consulted for the development of the FAIR4RS principles, and to provide a basis for their dissemination in the second half of 2021 through application and adoption guidelines (not included in this proposal). FAIR4RS WG workshops in November 2020 at both RDA Plenary 16 and a Series of Online Research Software Events (SORSE) will be used to identify early adopters, and conversations are underway between ReSA and national Research Software Engineer (RSE) associations on how they could lead engagement in their local communities. Other champions are likely to include the organizations involved in some of the 19 earlier works applying FAIR to aspects of research software, which includes disciplinary efforts by the American Geophysical Union (AGU) and ELIXIR (life sciences), and national initiatives by the Netherlands eScience Centre and US National Academy of Sciences.

6. What is the justification for the amount of money requested?

$86,00 is requested for 9 months, to fund a part-time program director and community manager to facilitate the work of FAIR4RS WG. The budget breakdown is as follows:

<table>
<thead>
<tr>
<th>Budget item</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project director, part-time for 7 months, based on total compensation</td>
<td>$49,640</td>
</tr>
<tr>
<td>Community manager, part-time for 6 months, based on total compensation</td>
<td>$23,760</td>
</tr>
<tr>
<td>Engagement, including conference presentation costs (attending and/or convening virtual events, technology support, possibly travel)</td>
<td>$1,383</td>
</tr>
<tr>
<td><strong>Direct cost total</strong></td>
<td><strong>$74,783</strong></td>
</tr>
<tr>
<td>Indirect costs (overhead) for Code for Science and Society, 15% of direct cost total</td>
<td>$11,217</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>$86,000</strong></td>
</tr>
</tbody>
</table>
7. What other sources of support does the proposer have in hand or has he/she applied for to support the project?

ReSA has applied for funding through Code for Science & Society to widen geographic inclusion in the FAIR4RS consultation by translating key virtual workshops to engage Francophone Africa and Latin American communities. ReSA is also in discussion with the Wellcome Trust to support projects to advance ReSA aims across infrastructure, people and policy projects. FAIR4RS WG has significant community support, with many community individuals donating their time to FAIR4RS WG, including the FAIR4RS WG steering committee members:

- Daniel S. Katz, National Center for Supercomputing Applications (NCSA), University of Illinois, USA
- Neil Chue Hong, Software Sustainability Institute - University of Edinburgh, UK
- Fotis Posomopoulos, Institute of Applied Biosciences, ELIXIR-GRE, Greece
- Morane Gruenpeter, INRIS / Software Heritage, France
- Jen Harrow, ELIXIR-Hub, UK
- Leyla Garcia, ZBMed, Germany
- Paula Martinez, National Informatics Facility, Australia
- Carlos Martinez-Ortiz, Netherlands eScience Center, Netherlands
- Michelle Barker, ReSA

Appendix 1: List of Citations


European Commission, Directorate-General for Research and Innovation, & PwC EU Services. (2018). Cost-benefit analysis for FAIR research data: Cost of not having FAIR research data. 10.2777/02999


https://doi.org/10.5281/ZENODO.4095092

https://doi.org/10.12688/f1000research.11407.1

https://danielskatzblog.wordpress.com/2017/06/22/fair-is-not-fair-enough/


https://doi.org/10.1038/sdata.2016.18

Appendix 2: CV for Dr. Michelle Barker

Appendix 3: Conflict of Interest Statement

The potential for conflict of interest with the ReSA project is that ReSA Steering Committee members are involved in a number of advisory boards and steering committees for projects,
journals and funders in the research software community. This will not cause conflicts in performing the work or sharing the results. However, the project outcomes may advocate for policies that would result in funding for your steering committee members’ organisations.

To mitigate this, we will be developing a conflict of interest register for the ReSA Steering Committee as part of the maturity of our governance structures.

Appendix 4: Attention to Diversity Statement

Code for Science & Society (CS&S), ReSA, and all CS&S’ Sponsored Projects are committed to meaningful equity and inclusive organizational practices. Science and technology are human endeavours, subject to influence by the biases of people. In the pursuit of science, people have often replicated power structures that reinforce the disenfranchisement and exclusion communities. In the American context, this has historically centered on the exclusion of marginalized communities from work in scientific and technical communities. CS&S strives to be a community leader on issues around meaningfully inclusive public interest technology across domains. CS&S does this through work with our Sponsored Projects and Collaborative Communities Programs.

When working with sponsored projects, CS&S focuses on helping projects grow sustainable, inclusive culture through evolving governance, working transparently, encouraging open dialogue around building leadership skills. Our Collaborative Communities program focuses on building capacity on the open source ecosystem. The Open Source Alliance for Open Scholarship Handbook Project, includes a frequently referenced definition of Open Scholarship, where equity and inclusion are central. Ongoing work with inclusion professionals DeEtta Jones & Associates will focus on the challenging conversations about inclusion in programs and governance that are happening (and often not happening) in the open source and science space. Our work in 2019 with DeEtta Jones is summarized here, and ongoing work on governance will be released in 2021. We co-organize csv.conf, which brings together a diverse global data science community - read reactions to our 2019 event from Julia Lowndes and Soil Kenya. This broad community-level impact helps support our Sponsored Projects as they grow and implement inclusive internal and community-facing practices.

As the nonprofit home of multiple sponsored projects, CS&S will oversee project leaders and teams, who will also collaborate with staff at outside organizations. By centering transparency and governance with our projects and leaning into organizational and community growth, we hope to continue to lead as a voice for meaningful inclusive practices in open source and science. Our direct plans, as well as examples of how we incorporate diverse representation our projects:

1. **Diversity in our Collaborators:** When collaborating with partners, we recruit diverse perspectives to participate, be it through user-centered design processes, governance, or to speak to our community.
2. **Equity and Representation in Staffing:** We offer structure and benefits to support our project’s abilities to recruit and retain talented people from all backgrounds.
3. **Centering Inclusion in the Project Governance:** CS&S has been actively engaged in supporting open source community to question, iterate, and mature governance models. For example, in 2020 we engaged DeEtta Jones & Associates to help us to develop and implement governance processes for ourselves and our sponsored projects that center meaningful equity and anti-racist values.
4. **Proactively Seeking Opportunities to Engage with the Global Community:** It can be hard for small projects on limited budgets to adjust to the needs of a growing community. There are unique funding opportunities for open projects to get expert diversity support. As an example, PREreview was selected to convene a working group of international experts at 2019 TriangleSci to focus on bringing equity and diversity to peer review, and a Wellcome Trust grant specifically focused on Diversity and Inclusion.

ReSA is already developing a global strategy through diverse representation on its Steering Committee. As the project develops, CS&S will support ReSA to implement inclusive practices. ReSA will also seek to ensure diversity in the stakeholders that will be consulted within the engagement strategy that will be completed for this project. For example, ReSA recently applied for funding through Code for Science & Society to widen geographic inclusion of the consultation by translating key virtual workshops to engage Francophone Africa and Latin American communities. This proposal was unsuccessful but will be resubmitted in the next round in 2021.

**Appendix 5: Information Products**

The Research Software Alliance (ReSA) will produce deliver the following key outputs:

- **Engagement plan** detailing webinars, workshops and conference sessions to engage the community in developing these outputs throughout 2021 - 28 Feb 2021.
- **Document summarizing** a community-agreed definition of the FAIR principles for research software, for broad dissemination (including publication) - 30 June 2021.

All materials will be made available under a CC-BY license. It is the stated goal of the ReSA project that these outputs be open to promote transparency and be of use to the wider community. Details on particular types of outputs are as follows.

**Articles and Working Papers**

- Articles and working papers will be produced. Publications that could be considered include F1000Research, PLOS ONE, or possibly Nature or Science. Pre-publication versions of articles will be made available on Zenodo as part of the consultation process, and available under a CC-BY license.
- Will a copy of the article be archived in an It is highly likely that a copy of the article will be archived in an institutional repository, and subject to an institutional mandate, but is dependent on the policies of individual authors’ institutions.

**Websites & Web Content**

- Project-related websites will be hosted by ReSA (www.researchsoft.org) and RDA (https://www.rd-alliance.org/groups/fair-4-research-software-fair4rs-wg). RDA is a partner in the FAIR4RS WG.
- These websites will be maintained for at least 12 months after the completion of the project. Their relevance will then be reviewed, as the project publications and working papers will be publicly available elsewhere, as described above. If they were not going to be maintained then the websites would be archived with archive.org. The project may also store some materials on Github, which is automatically archived by Software
Heritage.

- ReSA and RDA will be responsible for maintaining project-related websites.
- ReSA or RDA will own the intellectual property (IP) of any web content on their websites. CC-BY will be the license by which this is made available to others.

Other types of outputs

- It is not anticipated that any datasets, software, code or artistic work will be produced, as the project centers on development of community agreed standards.