FAIR for Research Software

2020 RDA Working Group Case Statement

1. Working Group Charter

One of the major challenges of data-driven research is to facilitate knowledge discovery by assisting humans and machines in their discovery of, access to, integration and analysis of data and their associated research objects, e.g., algorithms, software, and workflows. To address this, an initial effort to define a "DATA FAIRPORT"[1] began in 2014 at the Lorentz workshop and transitioned into developing a set of FAIR Data Guiding Principles[2] in 2016. The details of the FAIR data principles strongly contribute to addressing this challenge with regard to research data. The principles, at a high level, are intended to apply to all research objects; both those used in research and those that form the outputs of research. Here we focus on the definition, adaptation and adoption of the FAIR principles for the case of research software.

Software has become essential for research. To improve the findability, accessibility, interoperability, and reuse of research software[3], it is desirable to develop and apply a set of FAIR Guiding Principles for software. Many of the high-level FAIR data principles can be directly applied to research software by treating software and data as similar digital research objects. However, specific characteristics of software — such as its executability, composite nature, and continuous evolution and versioning — make it necessary to revise and extend the original data principles.

Application of the FAIR principles to software will continue to advance the aims of the open science movement. The FAIR For Research Software Working Group (FAIR4RS WG) will be jointly convened as an RDA Working Group, FORCE11 Working Group, and Research Software Alliance (ReSA) Taskforce, in recognition of the importance of this work for the advancement of the research sector.

FAIR4RS WG will enable coordination and leverage existing community-led discussions on how to define and effectively apply FAIR principles to research software, to achieve adoption of these principles. The working group will deliver:

- A document developed with community support defining FAIR principles for research software
- A document providing guidelines on how to apply the FAIR principles for research software (based on existing frameworks)
- A document summarising the definition of the FAIR principles for research software, implementation guidelines and adoption examples.

[1] See also DTL, 2014; and Kok, 2014.
[2] See also Wilkinson et al., 2016.
[3] For further information refer to Clément-Fontaine et al., 2019.

2. Value Proposition

FAIR4RS WG will define FAIR principles for research software and identify practical means of achieving them.

This work will be of value to software project owners, researchers, users of research data and software, the scientific community, research software engineers, software developers who publish their software, software catalogue maintainers, repository managers, software preservation and archival experts, policymakers who are responsible for defining digital policies, and organisations that create, modify, manage, share, protect, and preserve research software, funders of research, and others with an interest in the FAIR principles for research software.

FAIR4RS WG will provide a range of ways for community members to engage. All community members will receive regular updates through the RDA email list. The email list will facilitate collaborations through invitations to webinars, collaborative documents, surveys, etc. All documentation produced by the group will be publicly accessible via collaborative documents. Community members are able to engage at any of the three levels:

- Steering committee. The steering committee is responsible for leadership of the WG.
- Working and feedback cohort. Community members can choose to engage with the WG by providing feedback at their preferred pace via the WG RDA space, subgroup activities and the WG GitHub repo.
- Advocates. Those who can play a key role in endorsing and promoting the outcomes of this group.

3. Engagement with existing work in the area

The FAIR4RS WG emerged from the "FAIR for research software" session at VP15, run by the Software Source Code IG along with authors of the position paper "Towards FAIR Principles for Research Software" https://doi.org/10.3233/DS-190026. This group brings together and builds on a wide range of work internationally; notably the SSC IG activity at P13 and events within the ELIXIR community. Other key community publications and events involving AGU, ARDC, CLARIAH, CODATA, DANS, ELIXIR, EOSC FAIR WG, FAIRsFAIR, FAIRsharing, GO-FAIR, Netherlands eScience Center, ReSA, SSI, Software Heritage and TIB are included in the following list[4]:

[4] Further details on these initiatives are contained in Software Source Code IG wiki VP15, 2020.

Members of these groups have been collaborating with FAIR4RS WG, and the adoption plan for the WG outcomes will continue this work. We will bring in other organisations as appropriate, such as US Research Software Sustainability Institute (URSSI), Workshop on Sustainable Software for Science: Practice and Experiences *(*WSSSPE), Software Preservation Network (SPN), Software Heritage, Software Sustainability Institute (SSI), Society of Research Software Engineering, CodeMeta, in addition to other national and disciplinary initiatives.

The significance of this work is illustrated by the convening of FAIR4RS WG across RDA, FORCE11 and ReSA. FORCE11 is a community of scholars, librarians, archivists, publishers and research funders that has arisen organically to help facilitate the change toward improved knowledge creation and sharing. ReSA's vision is to see research software recognised and valued as a fundamental and vital component of research worldwide. ReSA leads a number of taskforces that assist in achieving these aims. This new FAIR4RS WG will be jointly convened as an RDA WG, a FORCE11 WG, and a taskforce of ReSA.

RDA groups

FAIR4RS WG will also engage with a range of other RDA WGs whose interests overlap, and ensure that the work of the WG aligns with the work of the other groups. We have a commitment to engage with co-chairs of relevant RDA groups and we would like to extend an invitation to the members of the following groups to check regularly our public updates via the mailing list archive.

- CURE FAIR WG - focusing on the reproducibility aspect of FAIR for data and code.
- SSC IG - provides a forum to discuss issues on management, sharing, discovery, archival and provenance of software source code. It pays special attention to source code that generates research data and plays an important role in scientific publications. FAIR4RS is proposed to be a WG of the SSC IG.
- SCID WG - bringing together a broad panel of stakeholders directly involved in software identification and planning an output with concrete recommendations for the academic community to ensure that the solutions that will be adopted by the

academic players are compatible with each other and especially with the software development practice of tens of millions of developers worldwide.
- FAIR data maturity model - working on RDA recommendations for a common set of core assessment criteria for FAIRness and a generic and expandable self-assessment model for measuring the maturity level of a dataset.

Other potential groups to coordinate with include: Go FAIR IG, Exposing Data Management Plans WG, Active Data Management Plans IG, Research Funders and Stakeholders on Open Research Data Management Policies and Practices IG, Research Metadata Schemas WG, and Global Open Research Commons IG.

4. Work Plan

The WG aims at defining FAIR for research software, and identifying guidelines on how to apply the FAIR principles for research software. Note that while the group is interested in FAIR metrics and indicators for research software, this is not in scope for this WG. We suggest this could be an activity for a follow-on group. We also like to emphasise that with the early definition of a 'steering committee' to coordinate the work the deliverables are attainable in a timeline of 18 months from endorsement.

	Milestone	*Deliverable*	*Due date*
1	Identify issues in the definition of FAIR for research software based on analysis of existing definitions and frameworks. Include definition of research software to delineate the scope that FAIR will be applied to. The draft would identify commonalities and differences, and thus key questions for the community to engage. Initiate consultation with the community (including identification of who is in this community, engagement with co-chairs of relevant RDA groups).	A review document identifying challenges in defining FAIR for research software	Prework: 0-4 months
2	Draft plan for development of guidelines for implementation, including identification of use cases, and community consultation on guidelines.	Reviewed work plan, collection of use cases and strategies for community consultation.	+4 months
3	Complete community consultation and finalise definition of guiding principles for research software.	Document developed with community support **defining FAIR principles for research software**	+12 months

4	Implement plan and develop guidelines	Document providing **guidelines** on how to apply the FAIR principles for research software	+12 months
5	Draft adoption plan, and coordinating community activities to collect adoption examples.	Plan for adoption of guidelines, including identification of adoption examples	+16 months
6	Implement plan for adoption of guidelines, which will continue after the cessation of this WG.	Document summarising definition of the FAIR principles for research software, implementation guidelines and adoption examples	+18 months

5. Adoption Plan

FAIR4RS WG will define principles for FAIR for Research Software, then guidelines and practices for their widespread adoption across the research software community at national, disciplinary, and international levels. The WG will organise dissemination about the activities and findings and gather community feedback regularly during all the phases of the work. To promote transparency and accessibility of work in progress, the group members will utilise a public Google Drive collaborative working folder (through ReSA) and a GitHub repository (through FORCE11) for WG collaborative documents. The following table details the engagement that will be undertaken towards relevant milestones, including final adoption.

	Milestone	*Engagement*
1	Identify issues in definition of FAIR for research software, based on analysis of existing definitions and frameworks.	The Steering Committee and the group members (i.e., people registered as part of the WG) will work together in multiple subgroups to analyse the state of the art and identify FAIR needs corresponding to the research software case.
2	Draft plan for development of guidelines for implementation, including identification of use cases, and community consultation on guidelines.	After gathering use cases and reviewing challenges regarding the implementation of the principles, the Steering Committee will produce a draft plan, which will be circulated and validated with the community for feedback.

3	Complete community consultation and finalise definition of guiding principles for research software.	Undertake consultation with authors of other work done to date and broader research software community on the issues identified in milestone 1. To achieve a finalised document defining FAIR for research software. Engage relevant RDA IGs and WGs to gather feedback and identify areas of overlap as well as the ReSA and FORCE11 communities. Likely tools: webinars, google docs, surveys, comments on posts and github issues.
4	Implement plan and develop guidelines	Engage the community (including other RDA WG/IGs and those institutions and projects of interest listed in section 3) to get input on guidelines for application of definition of FAIR for research software. Likely tools: webinars, google docs, surveys.
5	Draft plan for adoption of guidelines, including identification of adoption examples	After gathering adoption examples from the community, the Steering Committee will draft a set of adoption guidelines, which will be circulated and validated with the community.
6	Implement plan for adoption of guidelines, which will continue after the cessation of this WG.	The adoption plan will address how to work with different stakeholders, including: ● those that will endorse and promote the guidelines ● those that will provide training on the guidelines ● users of the guidelines ● and will include suggestions on follow-up work

The convening of this WG across RDA, FORCE11, and ReSA will support usage of the outcomes across those communities. RDA, FORCE11, and ReSA will systematically promote the outcomes, aiming to raise awareness and facilitate a wider adoption of the Working Group outcomes by existing and emerging initiatives. Organisations with a focus on FAIR will also be engaged, to encourage promotion of the application of FAIR to research outputs other than data.

6. Initial Membership

During the virtual Plenary 15 (March 2020), 50 participants joined the two SSC IG sessions that helped conveening this group. They were invited to indicate their interest in joining FAIR4RS WG. This provided a list of 34 initial members. Communications around the

formation of this WG will continue to promote membership to the wider community, which is intended to extend beyond RDA members. The fact that the WG is established jointly with FORCE11 and ReSA is another guarantee of the breadth of representation. To date more than a 100 members have joined the FAIR4RS WG.

Steering Committee

The co-chairs of the RDA FAIR4RS WG are Michelle Barker, Leyla Garcia, Daniel S. Katz, Paula Andrea Martinez and Neil Chue Hong. They will liaise between RDA and the Steering Committee leading FAIR4RS.

First name/s	Surname	Gender	Institution/s	Country
Michelle	Barker*	Female	ReSA/ University of Melbourne	Australia
Leyla	Garcia*	Female	ZB MED Information centre for life sciences	Germany
Daniel S	Katz*	Male	University of Illinois at Urbana-Champaign	US
Neil	Chue Hong*	Male	Software Sustainability Institute / EPCC, University of Edinburgh	UK
Paula Andrea	Martinez*	Female	The University of Queensland / National Imaging Facility	Australia
Morane	Gruenpeter	Female	Inria/ Software Heritage	France
Fotis	Psomopoulos	Male	Institute of Applied Biosciences (INAB), Centre for Research and Technology Hellas (CERTH)	Greece
Jen	Harrow	Female	ELIXIR-Hub	UK
Mateusz	Kuzak	Male	Netherlands eScience Center	Netherlands

* denotes FAIR4RS co-chairs

Active participation in the steering committee will remain open. It is possible for people to easily join or leave the committee, with ongoing continuous disclosure of membership changes.

Contributions to the case statement

Paula Andrea Martinez, Michelle Barker, Daniel S. Katz, Leyla Garcia, Neil Chue Hong, Morane Gruenpeter, Fotis Psomopoulos, Jennifer Harrow, Mateusz Kuzak, Peter McQuilton, Andrew Treloar representing the RDA TAB review, Esther Plomp and Carlos Martinez on behalf of the Netherlands eScience center.

References

Clément-Fontaine, Mélanie, Roberto Di Cosmo, Bastien Guerry, Patrick Moreau, and François Pellegrini. 2019. “Encouraging a Wider Usage of Software Derived from Research.” Research Report. Committee for Open Science’s Free Software and Open Source Project Group. https://hal.archives-ouvertes.fr/hal-02545142.

DTL. “Data FAIRPORT”. 2014. Available at https://www.dtls.nl/2014/01/20/jointly-designing-data-fairport/ https://www.datafairport.org/ https://www.czelo.cz/files/Akce-2016/FAIRPORT-report-final.pdf

Kok, R. 2014. “Data Fairport: enabling global exchange of research data,” International Innovation, no. 131, pp. 98–101, 25 Feb. Available at https://www.dtls.nl/wp-content/uploads/2014/08/FAIRPORT.pdf

Lamprecht, A.L., Garcia, L., Kuzak, M., *et al.* Towards FAIR Principles for Research Software. Data Science 3 (2020) 37–59 37. IOS Press https://doi.org/10.3233/DS-190026

Software Source Code IG. 2020. FAIR for software reading materials. VP15 session wiki. Available at https://www.rd-alliance.org/group/software-source-code-ig/wiki/fair4software-reading-materials

Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3,** 160018 (2016). https://doi.org/10.1038/sdata.2016.18