FAIR for Virtual Research Environments (VREs) Case Statement

1. Working Group Context

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 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 adaptation and adoption of the FAIR principles for VREs (Virtual Research Environments).

Digital objects such as data, software and workflows cannot be made FAIR in isolation - digital infrastructure is needed to store, manage, analyse and share the digital objects, and to make them discoverable.

Virtual research environments (VREs), also called science gateways, research platforms or virtual labs, are increasingly used as the vehicle for collecting or generating digital objects, processing, analysing, annotating and visualising these, then sharing the research outputs. VREs are defined here as a broad category of digital research infrastructure that consist of a set of online services, often with associated integration and/or orchestration functions and connections to specific data, software, workflow, storage and compute resources. One of the main drivers for creating VREs is to lower or remove the barrier for researchers in accessing datasets, performing complex analyses, and sharing their workflows to encourage reproducibility.

How infrastructure such as a VRE is developed, and the functions it supports, therefore have a large impact on the FAIRness of digital objects themselves. VREs should *enable* FAIRness in the digital objects that they create or produce, and at the very least should not make digital objects that they process less FAIR.

VREs themselves should also *be* FAIR, in that they should be easily discoverable and accessible; should interoperate with other digital research infrastructures; and their technical architecture, components and services should be reusable to improve development efficiency.

While some of the high-level FAIR principles as applied to data can be directly applied to VREs, others are not applicable. Likewise, the recently developed FAIR principles for

---

[1] See also DTL, 2014; and Kok, 2014.
[2] See also Wilkinson et al., 2016.

[Research Software](#) do not cover all of the aspects of VREs. The application of the FAIR principles is now also being considered for other areas, including workflows, machine learning, artificial intelligence (AI), and skills and training.

The FAIR4VREs WG will enable coordination between existing communities working with VREs, science gateways, platforms and virtual labs, to define what it means for VREs to *be* and *enable* FAIR, and provide guidance to VRE developers in achieving this.

The working group will:

1. Investigate how the existing application of the FAIR principles to data, software, workflows, computational notebooks, training materials, AI and machine learning enable VREs to enable FAIR digital objects, and themselves be FAIR, and identify any gaps in the existing work.
2. Produce guidance on and examples of how VRES can and should *be* FAIR.
3. Produce guidance on and examples of how VREs can and should *enable* FAIRness for other digital objects.

As the working group explores the FAIR ecosystem with respect to VREs, the exact format of the output(s) will be defined. Rather than producing a set of principles specifically for VREs, the output(s) will be another connecting piece in the FAIR toolkit between principles and implementation as they apply to VREs.

2. Value Proposition

The FAIR4VREs WG will identify how VREs can and should *be* FAIR and also *enable* FAIRness for other digital objects, and produce guidance for VRE developers on making their VRE both FAIR and FAIR-enabling.

This work will be of value to digital infrastructure developers and managers; researchers and other users of research data; research software engineers and researchers who develop tools; repository managers; policymakers who are responsible for defining digital policies; funders of research; and others with an interest in the FAIR principles for all research outputs.

3. Engagement with existing work in the area

Activities outside RDA

There are specific activities that are underway through the Science Gateways Community Institute (SGCI, US), the Australian Research Data Commons (ARDC, AU), and the European Open Science Cloud (EOSC, EU). All three of these organisations are working through the implications of how to apply the FAIR principles to a range of digital objects.

In Australia, the ARDC has developed, or is developing guidelines relating to the FAIRness of:

- Data
- Software
- Services
- Platforms (equivalent to VREs).

In the US, SGCI is developing guidelines relating to the FAIRness of:

- Science Gateways (equivalent to VREs)
- Research Infrastructures

In the case of software, the principles developed by the FAIR4RS WG will be adopted by the ARDC, as will the adoption guidelines and implementation plan being developed.

In the case of services, the relevance of the work undertaken by the FAIRsFAIR project (M2.10 Report on basic framework on FAIRness of services) will be considered as part of the development of the outputs of this WG.

In the case of VREs, the ARDC and SGCI propose to lead international guideline development, building on the successful approach taken by the FAIR4RS WG.

FAIRness of research infrastructures combines aspects of FAIRsFAIR project and the SGCI Tech Summit (https://sciencegateways.org/resources/sgci-tech-summit).

RDA groups

FAIR4VREs 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. There is a commitment to engage with co-chairs of relevant RDA groups and members of the following groups are invited to regularly check our public updates via the mailing list archive.

- VRE IG seeks to build the required technical bridges, skills and social communities that enable global sharing and processing of data across technologies, disciplines and countries through the creation of shared online virtual environments. As these individual VREs grow, inevitably they need to also connect with other major research infrastructures. This FAIR4VREs WG is sponsored by the VRE IG.
- FAIR4RS WG is enabling coordination of a range of existing community-led discussions on how to define and effectively apply FAIR principles to research software, to achieve adoption of these principles. FAIR4VREs will incorporate the outputs of FAIR4RS.
- FAIR data maturity model WG - 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.
- CURE FAIR WG - focusing on the reproducibility aspect of FAIR for data and code.

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, FAIR in AI BOF (exploring an IG), and Global Open Research Commons IG.

4. Work Plan

The FAIR4VREs WG will identify how VREs can and should *be* FAIR and also *enable* FAIRness for other digital objects by investigating how the existing applications of the FAIR principles to digital objects translate to VREs, and identifying any gaps in the existing work.

Guidance will then be developed for VRE developers on making their VRE both FAIR and FAIR-enabling.

Note that while the group is interested in FAIR metrics and indicators for VREs, 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 co-chairs to coordinate the work we believe the deliverables are attainable in a timeline of 18 months from endorsement.

4.1 Milestones and Deliverables

Milestone and description of work	**Deliverable**	**Due date**
Define scope of work Identify issues in the application of FAIR to VREs based on preliminary analysis of existing definitions and frameworks. The scope would identify commonalities and differences, and thus key questions for the community to engage.	**A review document outlining the issues that need to be addressed in defining FAIR for VREs**	0-4 months
Initiate consultation with the community (including identification of who is in this community, engagement with co-chairs of relevant RDA groups).		0-4 months
Finalise work plan for identifying any gaps in the existing FAIR work, including identification of use cases, and community consultation.	**Revised work plan**, collection of use cases and strategies for community consultation.	+4 months
Finalise analysis of gaps in existing FAIR principles for VREs.	**Document identifying how VREs can and should be FAIR, and enable FAIRness for other digital objects**	+9 months

Crowdsource case studies on how VREs contribute to FAIRness of other digital objects and examples of guidance materials	**Collection of case studies** on how to achieve a FAIR and FAIR-enabling VRE	+12 months
Review work plan to determine if further guidance materials for VRE developers are required, and create as needed.	**Updated work plan**	+12 months
Draft adoption plan, and coordinate community activities to collect adoption examples and test guidance materials.	**Plan for adoption of guidance**, including identification of adoption examples	+12 months
Presentation of Recommendation at Plenary	**Presentation** summarising application of the FAIR principles for digital research objects to VRE development and use, including implemented examples	+18 months

4.2 Working Group Operations

In addition to meeting in-person or virtually at Plenaries, we will have two or more formal calls in between the Plenaries and share information via a mailing list.

Documents will be created and made public through Google docs and GitHub. This allows for collaborative work and also serves as a form of communication. Those individuals actively working on outputs will have ad-hoc virtual meetings as needed (at least monthly). Trello and Github will be used for planning and tracking group deliverables.

4.3 Addressing Consensus and Conflicts

The WG will adhere to the stated [RDA Code of Conduct](#) and will work towards consensus, which will be achieved primarily through mailing list discussions and online meetings, where opposing views will be openly discussed and debated amongst members of the group. If consensus cannot be achieved in this manner, the group co-chairs will make the final decision on how to proceed.

The co-chairs will keep the working group on track by reviewing progress relative to the deliverables. Any new ideas about deliverables or work that the co-chairs deem to be

outside the scope of the WG defined here will be referred back to the VRE IG to determine if a new WG should be formed.

4.4 Community engagement

The FAIR4VREs WG will provide a range of ways for community members to engage at any of the three levels:

- Co-chairs. The Co-chairs are 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.

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. The WG will organise dissemination about the activities and findings and gather community feedback regularly during all the phases of the work. All documentation produced by the group will be publicly accessible via collaborative documents.

5. Adoption Plan

The WG will create an adoption plan for distributing and maintaining the deliverables. A specific plan will be developed to facilitate adoption of the WG Recommendation within the organisations and programs represented by WG members. This will include strategies for adoption more broadly within the global VRE community.

The WG will aim to recruit other potential adopting organisations early in the development process and ensure that their perspectives inform the guidance, and adoptions would ideally start within the 18 month timeframe before the WG is complete.

6. Initial Membership

The initial membership of this group will be drawn from the RDA VRE IG. Communications around the formation of this WG will continue to promote membership to the wider community. Active participation as a co-chair or working group member will be encouraged as the work plan is further refined.

The initial co-chairs of the RDA FAIR4VREs WG are listed below.

First name(s)	Last name(s)	Gender	Institution(s)	Country
Kerry	Levett	Female	Australian Research Data Commons	Australia

Sandra	Gesing	Female	SGCI, Discovery Partner Institute	US
Kheeran	Dharmawardena	Male	Cytrax	Australia
Javier	Sánchez Mondragón	Male	INOAE	Mexico
Leyla Jael	G. Castro	Female	ZB MED – Information Centre for Life Sciences	Germany

Contributions to the case statement

Kerry Levett, Andrew Treloar, Sandra Gesing, Michelle Barker.

References

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

Koers, Hylke; Herterich, Patricia; Hooft, Rob; Gruenpeter, Morane; Aalto, Tero 2019. "M2.10 Report on basic framework on FAIRness of services" from Fostering FAIR Data Practices in Europe project (FAIRsFAIR). 10.5281/zenodo.4292599

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

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