David, R, et al. 2020. FAIRness Literacy: The Achilles' Heel of Applying FAIR Principles. *Data Science Journal,* 19: 32, pp. 1–11. DOI: https://doi.org/10.5334/dsj-2020-032

PRACTICE PAPER

FAIRness Literacy: The Achilles' Heel of Applying FAIR Principles

Romain David[1,2], Laurence Mabile[3], Alison Specht[4], Sarah Stryeck[5], Mogens Thomsen[3], Mohamed Yahia[6], Clement Jonquet[7], Laurent Dollé[8], Daniel Jacob[9], Daniele Bailo[10], Elena Bravo[11], Sophie Gachet[12], Hannah Gunderman[13], Jean-Eudes Hollebecq[14], Vassilios Ioannidis[15], Yvan Le Bras[16], Emilie Lerigoleur[17], Anne Cambon-Thomsen[3] and the Research Data Alliance – SHAring Reward and Credit (SHARC) Interest Group

[1] MISTEA, INRAE, Montpellier SupAgro, University of Montpellier, Montpellier, FR
[2] ERINHA (European Research Infrastructure on Highly Pathogenic Agents) AISBL, Paris, FR
[3] INSERM-Université Paul Sabatier Toulouse III, Toulouse, FR
[4] SEES-TERN, the University of Queensland, St Lucia, AU
[5] Graz University of Technology, Institute for Interactive Systems and Data Science, Graz, AT
[6] INIST-CNRS, Nancy, FR
[7] LIRMM, University of Montpellier, CNRS, Montpellier, FR
[8] Erasme Hospital, Brussels, BE
[9] INRAE, UMR BFP, Université de Bordeaux, Villenave d'Ornon, FR
[10] EPOS-ERIC/Istituto Nazionale di Geofisica e Vulcanologia, Rome, IT
[11] Istituto Superiore Sanità, Rome, IT
[12] Aix Marseille Université, CNRS, Avignon Université, IRD, IMBE, Marseille, FR
[13] Carnegie Mellon University Pittsburgh, Pittsburgh, US
[14] MISTEA, INRAE, Montpellier SupAgro, Université de Montpellier, Montpellier, FR
[15] SIB Swiss Institute of Bioinformatics, Lausanne, CH
[16] DGD-REVE, UMS PatriNat, MNHN, Paris, FR
[17] CNRS, UMR GEODE, Université Toulouse 2, Toulouse, FR
Corresponding author: Romain David (david.romain@gmail.com)

The SHARC Interest Group of the Research Data Alliance was established to improve research crediting and rewarding mechanisms for scientists who wish to organise their data (and material resources) for community sharing. This requires that data are findable and accessible on the Web, and comply with shared standards making them interoperable and reusable in alignment with the FAIR principles. It takes considerable time, energy, expertise and motivation. It is imperative to facilitate the processes to encourage scientists to share their data. To that aim, supporting FAIR principles compliance processes and increasing the human understanding of FAIRness criteria – i.e., promoting FAIRness literacy – and not only the machine-readability of the criteria, are critical steps in the data sharing process. Appropriate human-understandable criteria must be the first identified in the FAIRness assessment processes and roadmap. This paper reports on the lessons learned from the RDA SHARC Interest Group on identifying the processes required to prepare FAIR implementation in various communities not specifically data skilled, and on the procedures and training that must be deployed and adapted to each practice and level of understanding. These are essential milestones in developing adapted support and credit back mechanisms not yet in place.

Keywords: FAIR principles; FAIRness literacy; FAIR assessment, Research data sharing; FAIRification; Pre-FAIRification

Introduction

This paper reports on a work developed as part of the Research Data Alliance – SHAring Rewards and Credit Interest Group (RDA – SHARC IG), established in 2017 and gathering people from various disciplines (biomedicine, biodiversity, agriculture, geosciences, science information, semantic science) to move forward scientific crediting and supporting mechanisms for scientists who strive to organise their data (and material resources) for community sharing. The group's work led progressively to the notion that increasing the FAIRness literacy in the research field, as well as enabling the assessment of FAIRness by all scientific communities are prerequisite critical steps to make data & resources sharing truly happen. This paper aims at sharing this experience.

Making resources available for the community means ensuring that data (and related materials) are findable and accessible on the Web (Mabile et al. 2016), and that they comply with adopted international standards making them interoperable and reusable by others (Hansen *et al.* 2018, Lannom *et al.* 2019). This aligns with the FAIR data principles that have been proposed as part of the international initiative FORCE11,[1] to make data *Findable, Accessible, Interoperable and Reusable* (*FAIR data principles* from Wilkinson et al., 2016). Data should be made as FAIR as possible, even if it is not possible to make data open (Landi et al. 2019). However, the FAIR principles do not address the quality issues whereas they are a major concern for improving data sharing; they focus on mechanisms to facilitate and optimize data sharing and this does not preclude the responsibility in assessing the quality and appropriateness of the data reused.

In practice, making data FAIR requires considerable time, energy and expertise (Curty et al., 2017) and who should contribute to the process is still not clear. As a result, except in specific communities routinely dealing with big data (nuclear physics, astronomy, genomics, satellite imageries…), the FAIR principles are not considered a priority and not implemented rigorously or very heterogeneously across disciplines & countries.

This factual situation is persisting even though lots of efforts have been done over the last years to transform the FAIR principles into practice. Some technical solutions are now well described (Wilkinson *et al.* 2017) and many groups (including within RDA) are working to develop methods and tools. Multiple implementation networks have been created under the GO FAIR initiative to facilitate cross – discipline exchanges on FAIRification. Infrastructures for preserving, locating and reusing research data already exist (Wittenburg et al., 2019) such as the ones built in Europe in the context of the ESFRI roadmap[2] and the implementation of the European Open Science Cloud (EOSC).[3] In addition, worldwide validated certifications are now possible such as the ones delivered by the Core Trust Seal[4] which demands 16 trustworthy requirements for approved Core Trustworthy Data Repositories.

On the funders' side, handling of research data according to the FAIR principles as part of Data Management Plans (DMPs) is now strongly encouraged by most funders and even required in certain funding schemes. This holds true for many applications to European Commission calls including H2020 and for some national funding agencies (e.g., French ANR,[5] US NIH,[6] UK Wellcome Trust,[7] Austrian FWF[8]). However, projects that opt out are still encouraged to submit a FAIR compliant DMP on a voluntary basis (Jones *et al.* 2019). The RDA has several groups working on specifying DMPs, establishing recommendations for their format and building mechanisms to facilitate their creation and use.

Although essential, the methods and tools developed support machine actionability – capability of computational systems to use services on data without human intervention – in their focus (Schwardmann *et al.*, 2020) making difficult their usability and understanding beyond the data scientist communities which limits a lot their use. Implementation of services and tools required to get FAIR compliant data can only be possible if scientists who are leaders of projects and non-experts in data science are able to understand the whole of the FAIRification processes, the stakes and the steps required. This lies on the need to reach community-approved choices of necessary tools, services and standards. The willingness of non data experts to comply to FAIR principles is a prerequisite in such an approach. Thus, to deal with the increasing number of tools, standards & services released, sufficient and appropriate guidance and support are needed. To that

[1] https://www.force11.org/.
[2] https://www.esfri.eu/esfri-roadmap.
[3] https://www.eosc-portal.eu.
[4] https://www.coretrustseal.org.
[5] The ANR introduces a Data Management Plan for projects funded in 2019 onwards.
[6] NIH Data Sharing Policy and Implementation Guidance.
[7] https://wellcome.ac.uk/funding/guidance/how-complete-outputs-management-plan.
[8] https://www.fwf.ac.at/en/research-funding/open-access-policy/research-data-management/.

aim, appropriate criteria must be clearly defined to enable FAIRness assessment and methods, and processes must be developed to enable FAIRification. With the new age of Open Science, the *European Commission Working Group on Rewards under Open Science*[9] delivered a matrix to help implement and assess FAIR criteria during the elaboration of DMPs (EC DGRI, 2017). Based on this matrix, Reymonet *et al.* (2018) made a detailed plan of quality criteria required during the three phases of development of DMPs; meanwhile Wilkinson *et al.* (2018) displayed the FAIR criteria according to their ability to be validated automatically.

We believe that these criteria are not easily put in practice by researchers and are difficult to interpret by evaluators as they rely mainly on machine-actionability properties that may not be accessible to every scientist. In this respect, many recent collaborative works have started to propose ways to implement, adapt and evaluate FAIR principles in several communities (e.g., Herschel *et al.* 2017, Doorn, 2018, Doorn *&* Timmermann, 2018, Federer *et al.* 2018, Mons *et al.*, 2017, Stall *et al.*, 2018, de Miranda Azevedo *&* Dumontier, 2019, Erdmann *et al.*, 2019, Sansone *et al.*, 2019). However, we are also reaching a moment where the FAIR principles now need cross-community convergence and *consensus* (EC DGRI, 2016; EC DGRI, 2018; Jacobsen *et al.*, 2019; Sustkova *et al.*, 2019; Thompson *et al.*, 2019; Wilkinson *et al.*, 2019). The work on FAIR data standards, repositories and policies is already ongoing as very well illustrated by the FAIRsharing.org platform which gathered more than 2800 registered standards, databases and policies (McQuilton *et al.*, 2019, Sansone *et al.*, 2019), by various RDA – WG (*e.g.*, FAIR data maturity model) or by the international GO FAIR initiative (Schultes *et al.* 2018). There is a need for an implementation of the FAIR principles in a progressive and community-oriented way, consolidated within existing practices to ensure that they evolve without interruption and in a way that is acceptable to the various actors. With this approach, everyone's ability to assess FAIRness is at the heart of the process.

In consideration of the tools available and on the basis of discussions at workshops (Breakout sessions at RDA 13th and 14th plenaries), an online survey and multiple discussions in weekly teleconferences, the SHARC Interest Group developed and discussed a tool for FAIRification that was more attuned to the broader research community to enable FAIRness assessment either by external data evaluators or by the researchers themselves. Doing so, we addressed the building of pre-FAIRification processes which are needed by the community to better understand the requirements and stakes in FAIRification. This paper describes many of the matters that came to light as we worked towards the creation of such a tool. This whole experience is here reported.

Developing and proposing an interdisciplinary language

Because of the interdisciplinary nature of our IG's work, various interpretations of underlying concepts have arisen, implying a need for clear definitions. A glossary – designed to evolve – has been created using as many community-approved references as possible (available in "Glossary" tab in the SHARC tool – David *et al.* 2020), as well as other complementary terms.

Converging towards consensual human-readable, understandable & assessable criteria

Encouraging implementation of efficient data sharing methods requires that they can be assessed easily qualitatively and quantitatively by any scientist. In particular, the FAIR data principles must be intelligible for any research scientist, even if he/she is not a data scientist. FAIRness assessment must be realistic and pragmatic – what should be measured, and how to explicitly find the information needed.

Getting such human-understandable FAIRness assessment criteria would help to guide scientists to follow the FAIR data principles and would help evaluators to objectively achieve their task.

On the basis of the initial criteria developed by Wilkinson et al. (2016; 2018) and Reymonet et al. (2018) in particular, the SHARC IG's worked at providing a FAIR assessment template drawn on a new classification of FAIR criteria that is aligned with the questions researchers should address while elaborating DMPs. This template is meant to be adaptable to the needs of various communities and is designed to train and support their data FAIRification improvement. The sets of FAIR criteria and their relations have been summarised as mind-maps to present a quick overview of the FAIR aspects (**Figures 1–4**). The assessment tool is available at http://doi.org/10.5281/zenodo.3922069.

[9] Working Group on Rewards|Open Science – Research and Innovation – European Commission.

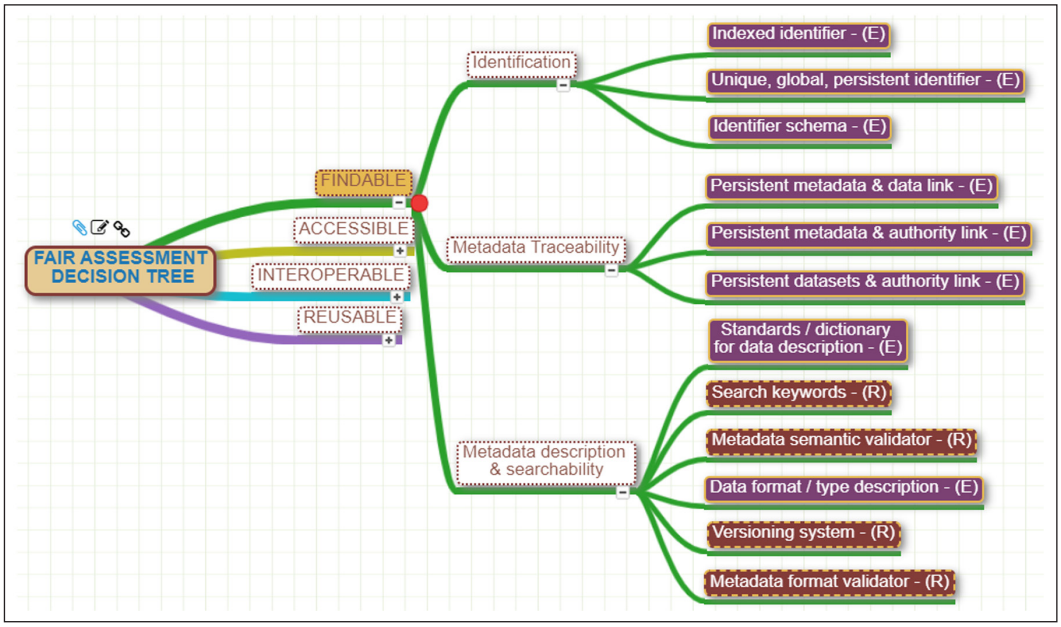

Figure 1: Mind mapping FAIRness assessment criteria. The mind map shows the extensive set of FINDABLE criteria of the FAIR principles (12 criteria). The level of importance of each criteria is divided into three categories (illustrated by three colours), Essential (purple)/Recommended (brown)/Desirable (red).

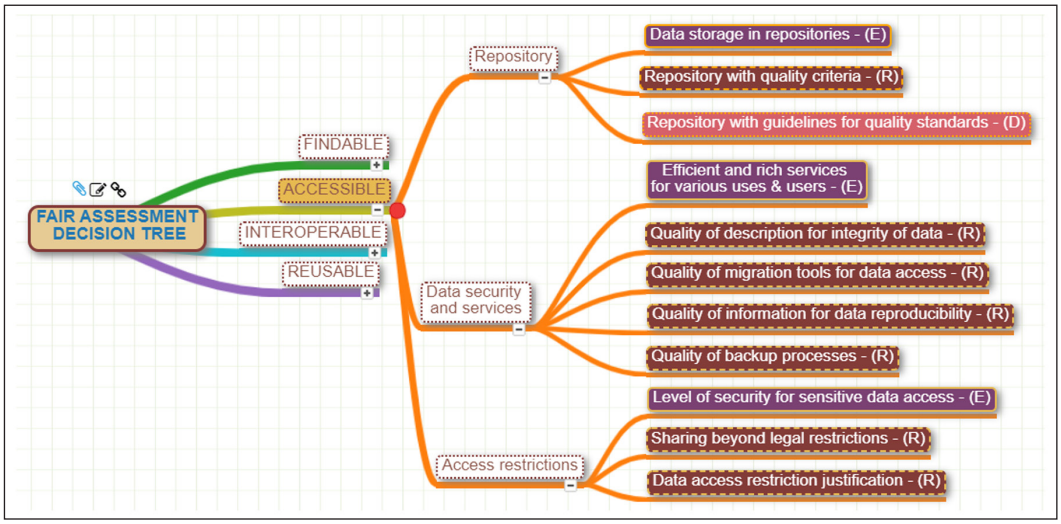

Figure 2: Mind mapping FAIRness assessment criteria. The mind map shows the extensive set of ACCESSIBLE criteria of the FAIR principles (11 criteria). The level of importance of each criteria is divided into three categories (illustrated by three colours), Essential (purple)/Recommended (brown)/Desirable (red).

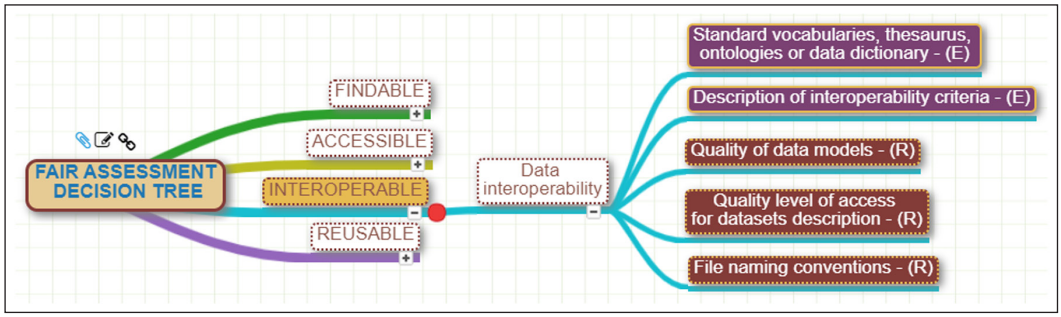

Figure 3: Mind mapping FAIRness assessment criteria. The mind map shows the extensive set of INTEROPERABLE criteria of the FAIR principles (5 criteria). The level of importance of each criteria is divided into three categories (illustrated by three colours), Essential (purple)/Recommended (brown)/Desirable (red).

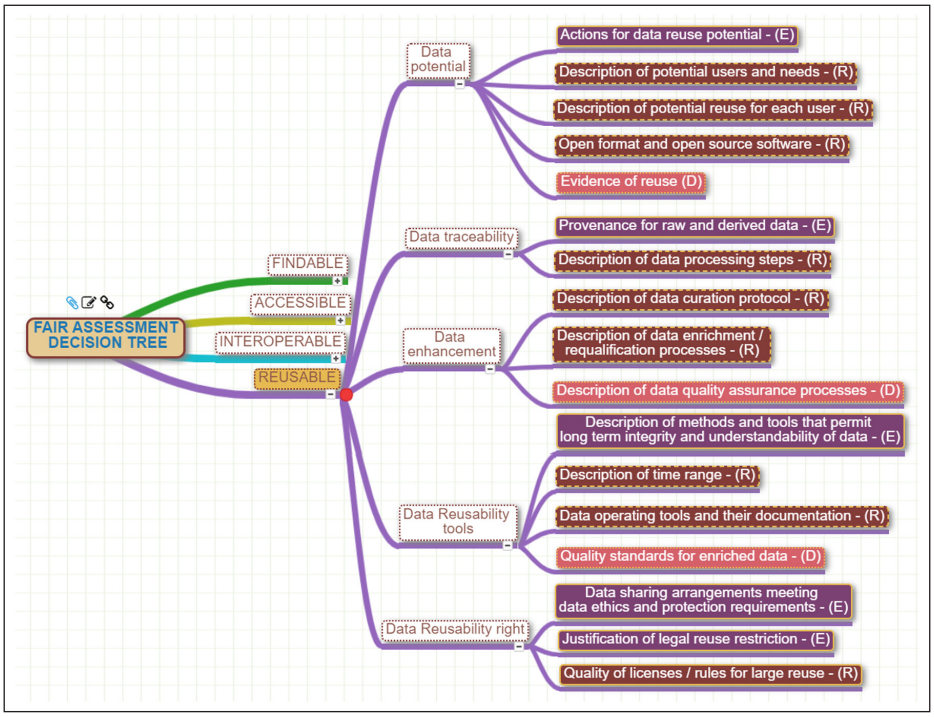

Figure 4: Mind mapping FAIRness assessment criteria. The mind map shows the extensive set of REUSABLE criteria of the FAIR principles (17 criteria). The level of importance of each criteria is divided into three categories (illustrated by three colours), Essential (purple)/Recommended (brown)/Desirable (red).

Lessons learned: need for a gradual implementation of FAIR criteria

While collectively designing and discussing the FAIR assessment tool, a major concern emerged: the need for a gradual implementation of the FAIR criteria, where at each stage the levels of achievement are made explicit. The collective feedback was analysed and formalised as a step-by-step iterative process for a FAIR data sharing, as described in parts A to C and further discussed in parts D to F.

A– Fostering pre-FAIRification decisions: when shared interoperability needs become community approved vocabularies goals

A better perception of the gain and return on investment from the use of controlled vocabularies (especially in the form of ontologies) is critical as it triggers the need to prepare FAIRification and ensure interoperability.

Each of the four categories of FAIR criteria refers to the use of community-approved formats, standards and controlled vocabularies (Kryger Hansen *et al.* 2018). Common naming conventions for classifying standards for reporting and sharing data or other objects appear favourable, assuming that those standards are registered and therefore reusable and mappable. FAIR criteria as defined by Wilkinson refer to FAIR compliant vocabulary used to ensure interoperability. The use of these reference vocabularies or ontologies also improves indexing and findability of data and sometimes makes it possible to cite an author for a selection of data only. This can help give additional impact to original studies by generating more citations. Standard vocabularies and ontologies also allow inclusion in many catalogues formatted with these recommendations. It also offers a greater ability to maintain data on the long term, allowing much greater backward compatibility of new software and new innovative processing chains of old data. Further, the use of vocabularies and ontologies increases the number of links and therefore possibilities of access and reuse of all or parts of a dataset.

However, the use or implementation of controlled vocabularies or ontologies may seem complex and time-consuming, especially since the means to find and explore them are still not fully harmonized (i.e., in the biomedical domain, an example in terms of number of semantic resources developed, four/five ontology repositories or libraries are available). It is crucial that the numerous potential benefits are better explained to each actor of a given scientific community. The responsibilities lie, among others, with repositories and standard providers. They have to outreach to the community via targeted education, publicity at events and reporting success stories.

Nonetheless, the first level of interoperability can be easily reached, with substantial return on investment by using metadata vocabularies recommended by the W3C to make data of different domains and disciplines compatible with each other. Semantic Web standards such as RDF, OWL and SKOS recommendations from the W3C, help metadata and data machine-readability.

B– Dealing with the unequal understanding of FAIR criteria

Even within a particular scientific community, definitions and interpretations of concepts covered by the FAIR evaluation criteria are manifold. The diversity of perception of minimum work to be achieved in the short and long term often leads to either a deep underestimation of the means to be implemented or even worse, to the fact that when significant resources are allocated to FAIRification, the solutions provided do not meet the prerequisites and create additional locks.

Two impediments relate to specific types of communities:

- In more mature scientific communities there may be a reluctance to change familiar tools and entrenched habits. FAIRification will probably require a great deal of effort, a change in mind-set and training in new skills.
- Emerging science communities are often embryonic; the complexity of the topics covered and the emergence of jargons (particularly in interdisciplinary fields) increases the difficulty of organizing complex processes around data at the interfaces of different scientific cultures.

These two impediments can be removed by detailing concrete actions and associated means for pre-FAIRification. Such actions must explain the meaning and the interest of the respect of each criterion within the framework of the community. In addition, these actions will not only have to be approved by the domain community as bringing a real improvement in the FAIR quality of data, but also meet all criteria for sustainable reuses beyond the community.

Furthermore, convincing all stakeholders that FAIR criteria compliance can foster data quality improvement, not only in terms of data management for reuse but also for data initial uses, will be a good asset.

Moreover, increasing global FAIRness compliance requires absolutely to raise the level of FAIRness literacy for each actor, and particularly that of team managers and project leader, who have the role of arbitrating between recruitment of several human profiles.

C– Planning pre-FAIRification training and support

Essential criteria (i.e., FAIR criteria that would block the FAIRification process if not applied) are not always understandable without specific education or training. Our collective discussions underlined frequently that implementation of some criteria is thought to be time consuming and may need sufficient and appropriate technical support, adapted to the capabilities of each actors (e.g., skills, language, availability).

The adoption of the FAIR data principles requires a cultural change. The first desirable and desired step for researchers is to better understand the efforts required to meet each FAIR criterion in a suitable manner that optimizes costs (time, human resources, skills). This good comprehension of the necessary efforts is a *sine qua non* condition for the researchers that can help them choose the less costly strategies to FAIRify their data. These efforts can relate to:

- Policies: e.g., the choice of the perimeter of the community in which this FAIRification process will be implemented;
- Strategic and tactical steps: technological choices, and implementation steps for wide acceptance in the predefined community;
- Pedagogical steps: e.g. providing educational kits adapted to different levels of skills up to 'qualified training' where 'trainees' knowledge is evaluated;
- Human resources: enabling researchers, engineers and technicians to acquire, at different time scales, the necessary skills;
- Governance: the training of managers to improve the relevance of their arbitrations and improve the planning of the means necessary for FAIRification in the short and long term.

It is also necessary that all the benefits of the FAIRification are clearly explained and demonstrated. Among the benefits of FAIRification is the impact of research, where data is a new entry point. It is also time saving

for setting up new data processing chain, manipulating data, cleaning up data, reconstructing missing data, or feeding more rigorously machine learning processes. FAIRification of data or services is also expected to build Virtual Research Environments (VRE) that are more reliable and allow for more reproducible research.

The strategic and tactical choices for implementing FAIRification must allow for a gradual implementation. Each step must be understood by each of the stakeholders, the decision process about criteria prioritisation should be defined (this may depend on the area of application) and possibly the list of criteria that can be neglected if the means are not sufficient. For each criteria, it will be useful to show the return on investment, especially in terms of scope, speed, quality and richness of shared treatments. All these elements must be included in training and supports.

D– Planning a step-by-step pre-FAIRification process

Considering the diversity of actors whose contribution is necessary for explaining, training and supporting FAIRification, it is imperative to rely on planning tools and on structuring the organization of the actors and on prioritizing the steps required with respect to each one of these actors, and according to the means and skills available.

This pre-FAIRification process is a *sine qua non* step for community acceptance of the efforts required for later FAIRification; especially with an efficient long-term effect and considering interoperability between heterogeneous systems.

Pre-FAIRification involves different stakeholders, such as funders, policy makers, publishers (to make FAIR data a requirement), institutions (to provide infrastructure, training, support and policies at their departments e.g., in the library), and disciplinary communities (to create community standards). This illustrates that researchers have to be supported throughout the entire life cycle of research data starting from their generation until final sharing and archival.

To empower scientists for these efforts, it is necessary to co-construct the FAIRification planning tools by taking into account the resources of each actor (e.g., typically DMPs in a research project are under the responsibility of the project coordinator). It is all the more important as future research work-programmes project (e.g., European Commission's Horizon Europe are announced to be granted only if the FAIRification processes are properly detailed and correctly sized in the proposal (human and technical resources in particular).

Identification of community milestones to organise FAIRification as an iterative and quality process resulted from our discussions (described below and illustrated as the wheel of FAIRification in **Figure 5**). FAIRification should include four distinct steps in an iterative process:

1. Preparing: FAIRification for a specific scientific community requires first that what is meant by FAIRification is carefully explained and that constraints and advantages in the short, medium and long term are described.
2. Training: to improve FAIRness literacy; it permits to convince stakeholders in the whole community. Note that this literacy should be maintained during the planning for FAIRification, especially when this is done before funding.
3. Pre-FAIRifying: this stage must be feasible for all actors. It encompasses the largest common denominator of objectives achieved by all and its success is crucial to empower the whole community further in the FAIRification process. Pre-FAIRifying can be divided in the following iterative steps:
    - 3a. Defining the community: a set of actors agreeing on the way they delimit their own community and the subjects related to it.
    - 3b. Defining objects and variables to be FAIRified
    - 3c. Selecting what to identify and index (data, actors, objects, processes, etc.)
    - 3d. Analyzing which are the common denominators
    - 3e. Reducing the explicit needs and expectations related to this first step in order to ensure the achievement of a common objective for all stakeholders (downward levelling)
4. FAIRifying by applying the prepared plan, check if results are compliant with the community approved plan, and adjust if necessary.

Whenever a common goal is reached, the community can redefine a new objective (and ideally enlarge it). Planning the second iteration with a new statement of expectations (back to step 3b or 3a if the community has evolved, which is very often the case) requires before so "Preparing" and "Training" steps.

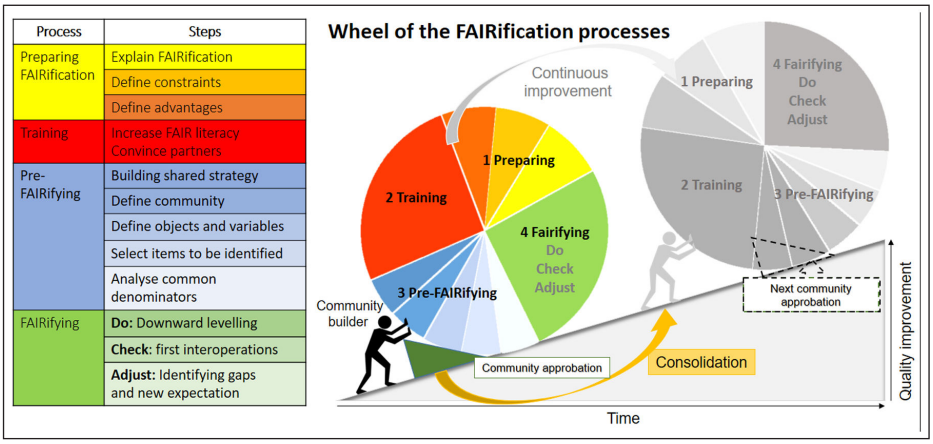

Figure 5: FAIRification can be schematized as a wheel describing iterative quality steps that need to be approved by the community throughout the process. This schema displays the "preparing" and "training" phases as conditions of pre-FAIRification. The pre-FAIRification processes must be community-approved at each iteration. The FAIRification steps 'check' and 'adjust' implementation must be approved by the community before a new iteration.

E– Maintaining sustainability

The complex concepts behind FAIRification keep evolving quickly: because of constant knowledge and process improvement, information systems for research data are fed by constantly renewed protocols. FAIRifying new and old data requires different iterative processes based on a consultation of all participants each time a new concept is needed. Both situations demand some differentiated human, technical or financial resources and organisations.

New data produced are increasingly heterogeneous and multi-source, sometimes even in exotic formats. Each new set of data, resulting from an exploration of a new research subject and/or the implementation of an experimental process based on emerging technologies is still rarely based on a fixed scheme allowing an immediate match with the FAIR principles. Human and technical support over time are critical for implementing them. For old data, unless it is a research subject on its own, FAIRification can not be funded currently out of responses to calls for research projects re-using these data.

Whether these data are old or recent, FAIRification is only possible if it occupies a central place in the research project, and therefore is understood and adhered to by all stakeholders as a long term issue. As it is clear, even today in the call for projects on FAIRification, that the level of understanding of the FAIR issues and of the needs for FAIRifying in a sustainable way is still very uneven, specific efforts are absolutely required for each type of actors.

F– Increasing good research data sharing during pre-FAIRification processes by rewarding and crediting

Among the obstacles to data sharing identified by the RDA SHARC ig group is the lack of recognition. This is even more true when considering the efforts required for FAIRification which is critical for efficient data sharing.

Until now, research institutions do not take into account FAIR principles implementation when evaluating researchers and themselves are not evaluated on the extent to which they support their researchers to do so. To that aim, it is essential that the FAIRification activity be assessed in research evaluation schemes (policy issues at institutional, national and supranational levels). Specific mechanisms necessary for such evaluation scheme (e.g., the use of identifiers for researchers and their data to enable credit back) have to be included early in the FAIRification process. Those are first aimed to link persistently and unambiguously shared data within the global web of data by associating their creators/contributors and institutions (with relevant metadata) so that shared data paternity is unequivocally attributed. Therefore perennial and unambiguous identification processes are needed for data on one side and persons (creators, contributors) or research institutions on the other side. The idea came out from many discussions that perennisation will be guaranteed only if communities are able to create, choose and administrate consensual data authorities over time.

Various types of identifiers currently exist, for research outputs such as DOI, ARKs, Handles, URIs etc., or for researchers (ORCID, ResearcherID, Scopus IDs, etc.) and for research organisations (ROR IDs, GRID, ISNI).

With CrossRef part of it, the current ecosystem makes it possible to interconnect unambiguously scientists with their deposited datasets, publications, other professional contributions or activities (e.g., open peer reviews), research organisations, funders and so forth. Online generated metrics can also help crediting the work and a different kind of reward may follow if this is taken into account in the research evaluation scheme (e.g., project awarding; dedicated financial support; career promotion).

Conclusion

The RDA – SHARC IG has established that the FAIR data principles need to be adequately explained from the very beginning of a research project design and that training needs to be provided as early as possible. To help implement this, the SHARC IG is developing a tool to support the assessment of FAIRness literacy and therefore enable measurement of progress towards compliance.

When developing this tool we identified a step-by-step process which will help teams organise the various actions towards the achievement of FAIR data, (i) by pre-FAIRifying, and (ii) by anticipating heterogeneity in FAIR literacy and sustainability. Even if the FAIR goal is not reached in one step, it can make FAIRification more understandable and improve its acceptability.

We highlight that researchers should be supported by data management professionals (not only data stewards), organised in networks and embedded in institutions. In order to enhance treatment of data according to the FAIR principles, we suggest that organisations should be assessed on the basis of how well they support their researchers in becoming FAIR advocates.

This shared experience provides arguments to motivate researchers and institutions to invest sufficient means (human or financial) for the needs of structured and long term FAIRification processes.

Nevertheless, FAIR compliance will certainly not be enough to enhance real and high data sharing and reuse; other mechanisms and qualities should be considered in future sharing processes (e.g., data veracity, quality, publicity, large indexation, curation, support). Developing a strategy to orchestrate efforts across the variety of communities should be a first priority in order to avoid the dispersed attempts of standards adoption and of compliance to FAIR.

Acknowledgements

This work was partly supported by the Research Data Alliance, especially the "RDA Europe 4.0" project (H2020 grant N°777388), the EPPN2020 project (H2020 grant N°731013), the 'Infrastructure Biologie Sante' PHENOME-EMPHASIS project funded by the French National Research Agency (ANR-11-INBS-0012) and the 'Programme d'Investissements d'Avenir'.

This research is also a product of the PARSEC group funded by the Belmont Forum as part of its Collaborative Research Action on Science-Driven e-Infrastructures Innovation (SEI2018) and by the French National Research Agency (ANR-18-BELM-0002-02).

Complementary support was provided through the FAIRplus project that has also received funding from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No 802750. The JU receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA". During his ERINHA AISBL involvement, Romain David was supported by the EOSC-Life european program under grant agreement N°824087 and ERINHA-Advance european program under grant agreement N°824061. Clement Jonquet was supported in part by the French National research Agency D2KAB project (ANR-18-CE23-0017). Special acknowledgement to Julien Lecubin from OSU Pytheas, France for technical support in the survey and to all the participants. Laurent Dollé is grateful to project support by Innoviris.

Competing Interests

The authors have no competing interests to declare.

Author Contributions

- Writing Paper: RD, LM, SS, LD, DJ, YLB, CJ
- Reviewing & editing paper: ACT, JEH, MT, MY, SG, AS, HG, EB, EL, VI
- Discussions about criteria implementation processes: SHARC IG members, especially during RDA 13th and 14th plenary meeting
- Key criteria conception: RD, LM, ACT, MT, MY
- Survey conception: RD, LM
- Survey review: SG, AS, ACT, MT, MY
- Survey completion: JEH, DB, HG, VI, 4 anonymous
- Survey analyses: RD

References

Curty, RG, Crowston, K, Specht, A, *et al.* 2017. Attitudes and norms affecting scientists' data reuse. *PLOS ONE*, 12: e0189288. DOI: https://doi.org/10.1371/journal.pone.0189288

David, R, Mabile, L, Specht, A, *et al.* 2020. Templates for FAIRness evaluation criteria – RDA-SHARC ig (Version 1.1) [*Data set*]. *Zenodo*. DOI: http://doi.org/10.5281/zenodo.3922069

de Miranda Azevedo, R and **Dumontier, M.** 2019. Considerations for the Conduction and Interpretation of FAIRness Evaluations. *Data Intelligence*, 285–292. DOI: https://doi.org/10.1162/dint_a_00051

Doorn, P and **Science Europe.** 2018. Science Europe Guidance. Presenting a Framework for Discipline-specific Research Data Management. [*WWW Document*]. URL http://www.scienceeurope.org/wp-content/uploads/2018/01/SE_Guidance_Document_RDMPs.pdf [Accessed January 07, 2020].

Doorn, P and **Timmermann, M.** 2018. Towards Domain Protocols for Research Data Management (IG Domain Repositories RDA 9th Plenary meeting Community-driven Research Data Management). *Paper presented at the 9. Plenary meeting Community-driven Research Data Management*, Barcelona. https://www.rd-alliance.org/sites/default/files/attachment/RDA%20DRIG%20Domain%20Protocols%20V3%20Barcelona%20April%202017%20-%20DoornAerts.pptx [Accessed January 07, 2020].

Erdmann, C, Simons, N, Otsuji, R, *et al.* 2019. Top 10 FAIR Data & Software Things. [*WWW Document*]. [Accessed January 07, 2020]. DOI: https://doi.org/10.5281/zenodo.2555498

European Commission Directorate General for Research and Innovation (EC DGRI). 2016. E.U. H2020 Programme Guidelines on FAIR Data Management in Horizon 2020, Version 3.0. Luxembourg: Publications Office of the EU. [*WWW Document*]. URL https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf [Accessed January 07, 2020].

European Commission Directorate General for Research and Innovation (EC DGRI). 2017. Evaluation of Research Careers fully acknowledging Open Science Practices; Rewards, incentives and/or recognition for researchers practicing Open Science. Luxembourg: Publications Office of the EU. [*WWW Document*]. URL https://ec.europa.eu/research/openscience/pdf/os_rewards_wgreport_final.pdf [Accessed January 07, 2020].

European Commission Directorate General for Research and Innovation (EC DGRI). 2018. Turning FAIR into reality: final report and action plan from the European Commission expert group on FAIR data. Luxembourg: Publications Office of the EU. [*WWW Document*]. URL https://op.europa.eu:443/en/publication-detail/-/publication/7769a148-f1f6-11e8-9982-01aa75ed71a1/language-en/format-PDF [Accessed January 07, 2020].

Federer, LM, Belter, CW, Joubert, DJ, *et al.* 2018. Data sharing in PLOS ONE: An analysis of Data Availability Statements. *PLOS ONE*, 13: e0194768. DOI: https://doi.org/10.1371/journal.pone.0194768

Hansen, KK, Buss, M and **Sztuk Haahr, L.** 2018. A FAIRy tale. *Zenodo*. [*WWW Document*]. DOI: https://doi.org/10.5281/zenodo.2248200

Herschel, M, Diestelkämper, R and **Ben Lahmar, H.** 2017. A survey on provenance: What for? What form? What from? *The VLDB Journal*, 26: 881–906. DOI: https://doi.org/10.1007/s00778-017-0486-1

Jacobsen, A, de Miranda Azevedo, R, Juty, N, *et al.* 2019. FAIR Principles: Interpretations and Implementation Considerations. *Data Intelligence*, 10–29. DOI: https://doi.org/10.1162/dint_r_00024

Jones, S, Pergl, R, Hooft, R, *et al.* 2019. Data Management Planning: How Requirements and Solutions are Beginning to Converge. *Data Intelligence*, 208–219. DOI: https://doi.org/10.1162/dint_a_00043

Landi, A, Thompson, M, Giannuzzi, V, *et al.* 2019. The "A" of FAIR – As Open as Possible, as Closed as Necessary. *Data Intelligence*, 47–55. DOI: https://doi.org/10.1162/dint_a_00027

Lannom, L, Koureas, D and **Hardisty, AR.** 2019. FAIR Data and Services in Biodiversity Science and Geoscience. *Data Intelligence*, 122–130. DOI: https://doi.org/10.1162/dint_a_00034

Mabile, L, De Castro, P, Bravo, E, *et al.* 2016. Towards new tools for bioresource use and sharing. *Information Services & Use*, 36: 133–146. DOI: https://doi.org/10.3233/ISU-160811

McQuilton, P, Batista, D, Beyan, O, *et al.* 2019. Helping the Consumers and Producers of Standards, Repositories and Policies to Enable FAIR Data. *Data Intelligence*, 151–157. DOI: https://doi.org/10.1162/dint_a_00037

Mons, B, Neylon, C, Velterop, J, *et al.* 2017. Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud. *Information Services & Use*, 37: 49–56. DOI: https://doi.org/10.3233/ISU-170824

Reymonet, N, Moysan, M, Cartier, A, *et al.* 2018. Réaliser un plan de gestion de données « FAIR » : modèle. [*WWW Document*]. URL https://archivesic.ccsd.cnrs.fr/sic_01690547/document [Accessed January 10, 2020].

Sansone, S-A, McQuilton, P, Rocca-Serra, P, *et al.* 2019. FAIRsharing as a community approach to standards, repositories and policies. *Nat Biotechnol*, 37: 358–367. DOI: https://doi.org/10.1038/s41587-019-0080-8

Schultes, E, Strawn, G and **Mons, B.** 2018. Ready, Set, GO FAIR: Accelerating Convergence to an Internet of FAIR Data and Services. *DAMDID/RCDL*. http://ceur-ws.org/Vol-2277/paper07.pdf.

Schwardmann, U. 2020. Digital Objects – FAIR Digital Objects: Which Services Are Required? *Data Science Journal*, 19(1): 15. DOI: https://doi.org/10.5334/dsj-2020-015

Stall, S, Cruse, P, Cousijn, H, *et al.* 2018. Data Sharing and Citations: New Author Guidelines Promoting Open and FAIR Data in the Earth, Space, and Environmental Sciences. *Science Editor*, 41: 83–87. https://www.csescienceeditor.org/wp-content/uploads/2018/11/CSEv41n3_text_83-87.pdf [Accessed January 07, 2020].

Sustkova, HP, Hettne, KM, Wittenburg, P, *et al.* 2019. FAIR Convergence Matrix: Optimizing the Reuse of Existing FAIR-Related Resources. *Data Intelligence*, 158–170. DOI: https://doi.org/10.1162/dint_a_00038

Thompson, M, Burger, K, Kaliyaperumal, R, *et al.* 2019. Making FAIR Easy with FAIR Tools: From Creolization to Convergence. *Data Intelligence*, 87–95. DOI: https://doi.org/10.1162/dint_a_00031

Wilkinson, MD, Dumontier, M, Aalbersberg, IjJ, *et al.* 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3: 160018. DOI: https://doi.org/10.1038/sdata.2016.18

Wilkinson, MD, Dumontier, M, Sansone, S-A, et al. 2019. Evaluating FAIR maturity through a scalable, automated, community-governed framework. *Scientific Data*, 6: 1–12. DOI: https://doi.org/10.1038/s41597-019-0184-5

Wilkinson, MD, Sansone, S-A, Schultes, E, *et al.* 2018. A design framework and exemplar metrics for FAIRness. *Scientific Data* 5. DOI: https://doi.org/10.1038/sdata.2018.118

Wilkinson, MD, Verborgh, R, da Silva Santos, LOB, *et al.* 2017. Interoperability and FAIRness through a novel combination of Web technologies (No. e2522v2). *PeerJ Inc.* DOI: https://doi.org/10.7287/peerj.preprints.2522v2

Wittenburg, P, Sustkova, HP, Montesanti, A, *et al.* 2019. The FAIR Funder pilot programme to make it easy for funders to require and for grantees to produce FAIR Data. [*WWW Document*]. URL https://arxiv.org/abs/1902.11162? [Accessed January 07, 2020].

How to cite this article: David, R, Mabile, L, Specht, A, Stryeck, S, Thomsen, M, Yahia, M, Jonquet, C, Dollé, L, Jacob, D, Bailo, D, Bravo, E, Gachet, S, Gunderman, H, Hollebecq, J-E, Ioannidis, V, Bras, YL, Lerigoleur, E, Cambon-Thomsen, A and The Research Data Alliance – SHAring Reward and Credit (SHARC) Interest Group. 2020. FAIRness Literacy: The Achilles' Heel of Applying FAIR Principles. *Data Science Journal,* 19: 32, pp. 1–11. DOI: https://doi.org/10.5334/dsj-2020-032

Submitted: 03 February 2020 **Accepted:** 27 July 2020 **Published:** 11 August 2020

Copyright: © 2020 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See http://creativecommons.org/licenses/by/4.0/.

]u[*Data Science Journal* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS