

Survey on bridging the gap between funders and communities – perspectives on benefits and challenges of FAIR assessments V1.0

Christophe Bahim, Makx Dekkers, Edit Herczog, Keith Russell, Shelley Stall

RDA FAIR Data Maturity Model Working Group

Abstract:

This report provides a consolidated view of the answers collected during a survey conducted in October 2020. This survey took place in the context of the work of the RDA FAIR Data Maturity Model Working Group and aimed at investigating the differences of perspectives on benefits and challenges of the FAIR assessments between funders and research communities.

Published: 14 January 2021

Contents

Executive summary.....	3
1 Introduction	4
1.1 Background	4
1.2 Approach.....	4
2 Survey responses	4
2.1 Baseline views.....	4
2.2 Future work	8
3 Conclusion	11

Executive summary

This report provides a consolidated view of the answers collected during the survey conducted in October 2020. The survey took place in the context of the work of the RDA FAIR Data Maturity Model Working Group and aimed at investigating the differences of perspectives on benefits and challenges of the FAIR assessments between funders and research communities. The survey was divided in three main sections (i) objective, (ii) challenges and (iii) future of the FAIR data maturity model (FDMM). Analysis of the answers showed interesting views, as listed below:

1. FAIR is seen as a means to tackle societal challenges through reuse of research data
2. Cooperation between communities can be enhanced by sharing best practices and tools
3. Developing standards within and across communities is seen as a way to improve reuse
4. Standards setting should be done at a global level while the implementation and adoption should be done at regional level.
5. Leading organisations within communities should work together towards creating standards
6. Funding is needed to develop further FAIR standards and agreements.
7. Assessments of FAIRness across communities are not seen as useful, tailored guidance on the other hand is recommended.
8. Increased collaboration and cooperation will come through use cases and awareness raising around these use cases.
9. Scoring is necessary to evaluate the FAIRness of digital resources
10. Scoring should be focused on improving the FAIRness of digital resources rather than comparing resources (across disciplines)
11. A FAIR baseline, per discipline, from which there is a trajectory to gradually improve the FAIRness of a digital resource should be set.
12. Feedback loops, case studies, success stories, regular updates to the FDMM are the key components to increase FAIRness.
13. Priorities of the FDMM could be better aligned with the communities.
14. Tailored and deepened assessments details are necessary
15. The FDMM requires additional and specific guidelines, i.e. more operationalisation of certain aspects of the FDMM.
16. The FDMM should be gradually improved through additional iterations and testing
17. It was suggested to create discipline-specific variants of the FDMM

All detailed answers can be found in the in report. The findings should be considered for the next revision of the FDMM.

The report also outlines the fact that this survey is not statistically representative to derive general conclusions nor trends as only ten people provided their views. A wider survey could be taken up to potentially explore the ideas, with regards to the topics identified. For future research in this topic, specific questions could be used to identify concrete action and solutions supporting the claim expressed in the survey. It would also be useful to reconduct the survey in a years' time to evaluate whether the opinions have changed.

1 Introduction

1.1 Background

During the development of the FAIR Data Maturity Model, participants in the RDA Working Group represented a wide range of backgrounds, including researchers and funding agencies. In some of the discussions in the Working Group and elsewhere in the period from early 2019 to late 2020, it appeared that those two stakeholder groups, research communities on one hand and funding agencies on the other hand, have different perspectives on the relevance and objectives of assessment of FAIRness.

In October 2020, the editorial team and the chairs of the FAIR data maturity model decided to address these different perspectives by designing a survey to ask for opinions from a limited number of representative of those two stakeholder groups to see whether conclusions and recommendations could be derived from the difference in perspective.

1.2 Approach

The following problem statement was formulated:

Improve the understanding of benefits and challenges of the FAIR assessments from the perspective of the funders and communities.

The objective of this survey was twofold:

1. Formulate conclusions and recommendations on the level of policy, i.e. better understanding of the views and challenges of both sides; and
2. Finding out how research communities and funding agencies might want to use the model and what changes they would want to see

In accordance to the two objectives of the survey, questions were divided in three parts. A first one related to the baseline views, a second one related to the policy level, and a third one related to the role and future development of the FAIR Data Maturity Model. Each part contained tailored questions for the two stakeholder groups.

In order to gather opinions, as briefly mentioned above, the editorial team selected a small number of active participants from the Working Group, who were invited to participate in the survey. Four representatives of funding agencies were invited of which three contributed their views, while seven of the twelve invited representatives of research communities provided their answers to the survey.

The answers to the survey were entirely anonymous.

It needs to be kept in mind that the number of participants contributing to the survey was relatively small and therefore did not represent a statistically representative group. Therefore, the results of the survey can only be interpreted as opinions of a small group of players involved in the work of the FAIR Data Maturity Model and not taken to be valid for a wider group of research communities and funding agencies.

2 Survey responses

2.1 Baseline views

The first section of the survey was concerned about establishing the baseline and understanding the views of the two stakeholder groups.

2.1.1 The most important objectives sought with FAIR initiatives

With this question we aimed to find out how funders would formulate their main objectives of assessment of FAIRness.

Funders

The objectives sought by FAIR initiatives are intrinsically linked to the FAIR principles, namely, findability, accessibility, interoperability and reusability. Enabling FAIR research data will make research reusable across disciplines allowing it to tackle major societal challenges and have a business impact.

The FAIR principles are an essential building block to create an ecosystem where research effort is minimised, and research reuse is maximised through good data management and scientific practices to enhance reproducibility and enable cross-disciplinary research.

2.1.2 How can the cooperation within communities be improved to allow data access and reuse (i.e. improved FAIRness)?

While a longer-term aim of the FAIR principle may be thought of as creating an environment that allows for cross-domain potential for reuse of research data, a first challenge is to improve the cooperation within research disciplines so that discipline-specific data can be readily reused, before sharing with other disciplines can start to take place. With this question, we aimed to get opinions from the research communities on the problems and potential solutions to this challenge.

Communities

Despite being widely mentioned in the RDA context, the term “community” remains unclear. The respondents are still questioning who these communities are and who are the stakeholders constituting them. Some respondents see communities made up of participants in the life cycle of data, e.g. creators, curators, repository providers, while others see them as collections of researchers in a specific discipline.

Once communities and stakeholders are defined, collaboration between them can be enhanced by sharing best practices, (open source) tools and expertise, e.g. indexing and rewarding FAIR initiatives and practices in a community could serve to enhance collaboration.

For that, communication is key. Therefore, having a strong leadership representing each community may be helpful. In that context, mutual benefit is an important factor and should be outlined. In this perspective, when established, data access and reuse become an essential component of cooperation within a community and allow for increased efficiency and democracy. Benefits of having access and the possibility to reuse quality data should be showcased to a wide audience through examples of initiatives that succeed in reusing data across communities.

Furthermore, establishing policies at the community level is helpful towards the common objective of improving FAIRness.

Also, in relation with improving cooperation within communities, they should be supported to develop – or adopt – their own standards and develop crosswalks across communities. For that, communities will require infrastructure, allowing open data and standards development, which will require consistent funding sources.

2.1.3 Summary

Funders agreed on the necessity of FAIRness to enable reuse of research data across communities. Communities on their end, reported on the need to collaborate and cooperate further by sharing best practices, tools and expertise. Additionally, communication and leadership are two factors that can contribute to a successful collaboration between communities.

2.1.4 What are the challenges related to different communities having different targets for the level of FAIRness?

A factor that could complicate the assessment of FAIRness across a wide range of research disciplines is that communities in different disciplines may have different levels of maturity; for example, some communities are already very advanced in documenting research results and providing machine-

processable data based on formal standards, while other communities are less advanced. Especially for funders, this makes it hard to determine what a realistic and reasonable target level of FAIRness is. With this question we aimed to get a view of what funders saw as the main challenges on this point.

Funders

According to the funders, assessments across communities are not very helpful. They acknowledged that there is a need for community-specific approaches. In the same way, for the communities that are less advanced in terms of FAIRness, having further discussion on the standards and interoperability of data is an imperative need that should be given priority.

To improve FAIRness of less advanced communities, funders advised on tailored guidance. Field-specific practices should be designed on top of core universal aspects.

Nevertheless, prior to assessing FAIRness across communities, funders shared the need to continuously engage and raise awareness across communities on the value proposition of data being FAIR.

2.1.5 Which organisations should take the lead in creating standards and tools to assess FAIRness? How could they be supported by funders and communities?

There are many types of participants in the work on the FAIR principles and the FAIR Data Maturity Model, from individual researchers to data curators, representatives of research institutes and universities, professional societies, standards organisations and funding agencies. The question is which organisations would be best placed to drive the movement to higher FAIRness and the resulting improvement of the reuse of research results.

Communities

Standard setting should ideally be done at global level (e.g. RDA, W3C), but adoption should be done at EU level (or equivalent in other regions). For that, it is essential that the organisations involved in the process of setting standards and tools are neutral and independent. Leading policy-setting organizations within each community, across disciplines, and across borders should work together in creating standards, including those for tools to assess FAIRness. As the assessment and interpretation FAIRness is often related to domain-specific knowledge, domains should take the lead in creating standards and tools. Finally, the bottom line is having community review.

When developing tools, FAIR assessment should never be used to “qualify” a resource at any point, but only to support the choice of what can be done to further enhance FAIRness. Comparison of FAIRness will not be needed so assessment frameworks do not need to give comparable results.

Additionally, funders can help speeding up the creation and applications of standard tools by providing funding. Besides, being beneficiaries of FAIR data, universities and publishers should also partake in funding activities.

Beyond the funding aspects, research institutes are also essential in developing, testing and applying the standards and tools to evaluate data FAIRness.

Funders

According to the opinions of the funders, research community organisations are ideally placed to develop standards as they understand the current practice in a discipline and have knowledge of the infrastructure as well as the repositories.

Bottom-up initiatives like RDA are best suited for these exercises. Funders could play a useful leadership role in bringing communities together, investing in piloting and scaling up approaches and accelerating the adoption of the tools by the research

community. Conversely, evaluation and assessment of FAIRness can be more top-down, where funders contribute to set the requirements that need to be evaluated.

It was mentioned that the RDA could play a similar role for the FAIR assessment of software and tools.

Funders and communities can raise the value of FAIR standards and tools with learned societies and organisations that provide infrastructure for research disciplines.

2.1.6 How could funding and/or other measures help to improve cross-community cooperation on improving FAIRness and the reuse of the existing infrastructure?

Improving cross-community communication and reuse potential will require support measures, in the form of funding or other ways to create understanding, incentives and capabilities. With this question, we aimed to collect suggestions on which measures would help to overcome this challenge.

Communities

Communities recommended an increased across-domain collaboration and cross-community cooperation to improve FAIRness. Early adopters could showcase cross-community cooperation. Support should be given to communities to clarify their own FAIR requirements and support to researchers by data management professionals embedded in institutions. Related to that, respondents believe organisations should be assessed based on how well they support their researchers in becoming FAIR advocates.

Next to increased collaboration and cooperation, communities identified the need of funding for supporting cross-community case studies as well as for developing standards and crosswalks between standards.

Use cases should be undertaken within the disciplinary communities in order to guarantee standardisation in the fields as well as between related fields.

Funders

Funders agreed on the need to provide dedicated funding to develop underlying FAIR standards and agreements. Specifically, they would coordinate investment in the further development of tools that build on emerging consensus.

It was also recommended to have university curricula and syllabi supporting training and skills development related to FAIR for future researchers.

2.1.7 Summary

With respect to policy and adoption, both the funders and communities see a clear role for neutral and independent platforms, such as the Research Data Alliance, allowing to bring stakeholders from different disciplines together and create more cross-community understanding.

From the community standpoint, it is important for policies to be coordinated on a global scale; implementations may be more realistically organised on a regional basis, for example through the European Open Science Cloud (EOSC) in Europe and similar initiatives in other regions. Additionally, communities should share best practices and develop community policies, to raise the overall awareness of the approach to and the benefits of FAIRness.

Funders believe that, within communities, professional societies and infrastructure providers are seen to be in the best position to drive the adoption of FAIR assessment methodologies. Funding can help in the development of best practices and standards, and in encouraging the adoption of tools.

Lastly, the value of FAIR data will increase when benefits can be shown from cross-discipline good practices and cross-disciplinary reuse. This can be achieved through awareness raising across communities with community-specific approaches and FAIR-related training for future researchers

2.2 Future work

2.2.1 Is the current approach in FDMM to create levels of FAIRness per indicator useful? What can be the role of the 'score' per FAIR area? How could FDMM be improved on these aspects?

An aspect of the FAIR Data Maturity Model that led to much discussion in the work of the RDA WG was the question in what way assessment approaches should try to deliver a result in the form of a score for the individual FAIR principles or even for overall FAIRness of a data resource being evaluated. With this question, we aimed to get the views of the research communities how scores and levels could be used.

Communities

Views on the scoring were quite varied. Respondents first observed that scoring is needed to evaluate the FAIRness of digital resources and can be used to establish a level of trust. Secondly, indicator levels, providing that they are clearly defined, can help data stewards to potentially improve their data management and stewardship practices. Levels are useful as communities are at very different stages of FAIRness, i.e. FAIR practices are different from one community to another and expectations on FAIR depend therefore on the use case.

Furthermore, scoring should avoid any kind of metrics that doesn't give room for improvement or makes data labelled as not FAIR.

Further on the levels and the indication of improvement needed it gives, it would be useful to identify and prioritise individual steps that could be undertaken on the way towards more FAIRness. The assessment results should be clear, actionable and benefiting the researcher. In short, scoring should be helpful as a development tool, to map improvements over time.

When scoring, it is crucial to communicate the context (i.e. input, output, assessment methods) so that users can interpret the results in a transparent and meaningful way.

Practically, the FDMM should provide additional user guidance to use the scoring mechanism. Training and examples from domains and disciplines could be used. For that, RDA adoption cases could help.

Funders can potentially define FAIR requirements, using the FDMM indicators and levels, for data produced by funded projects. In that context, scoring is one of the many ways for funding programmes to encourage or assess FAIRness.

Forward-looking, it could be useful to provide FAIRness assessment also for technologies tools.

Funders

The FDMM levels are not deemed useful for comparison across disciplines. But it is helpful as a development tool, to map improvements over time. Next to that, some funders expressed reservation about scores for FAIRness and most importantly how they'll be used.

On the other hand, others expressed the opinion that concrete scoring is necessary to evaluate the FAIRness of different digital resources. FDMM indicators and levels can be used by funders and organisations to build their own evaluation methodologies.

Assessments need to result in clear and actionable guidance.

2.2.2 What are the main ways that funding programmes could use FDMM to encourage or to assess FAIRness?

Funding programmes could use different approaches in using the model to deliver the policy objectives of improving the FAIRness of data resulting from funded activities. With this question, we aimed to get suggestions from the funders on which measures they were contemplating.

Funders

Funders have identified ways the FDMM to encourage or assess FAIRness.

- By starting to discuss with a research community on an acceptable base score of FAIRness and a trajectory to gradually improve the score.
- By evaluating whether resources comply with desired levels of FAIRness during project reporting.
- By educating and assisting researchers in their self-evaluation of the FAIRness of their resources.
- By assessing outputs reported by researchers and target support for areas where the principles are not yet being followed.

2.2.3 In what way could FDMM help in providing insight into the benefits of increased FAIRness?

One of the ways that the model could be used is to make the benefits of improving FAIRness quantifiable or at least visible to researchers and data curators. With this question, we aimed to gather ideas from research communities on how the model could support the creation of such insights.

Communities

The respondents have identified different ways for the FDMM to provide insight into the benefits of increased FAIRness.

- By gathering feedback from the implementers on the practical tests designed against the indicators of the FDMM model.
- By funding case studies. Identified best practices will help other data stewards to improve their FAIRness.
- By adding success stories from case studies to the FDMM to make the benefits of the model visible. Furthermore, FDMM should explain for every one of the indicators what the specific benefits are to projects and to science as a whole.
- By showing greater reuse or other impacts of data as it becomes increasingly FAIR.
- By regularly updating the FDMM.
- By using the FDMM as a common starting point to build an international FAIR community.
- By training the researchers and offering them a basic version of the FDMM that can be customised into learning material. Feedback from these new users would be valuable inputs for the FDMM.

2.2.4 What are the challenges of applying the FDMM for your own purpose?

During the development of the model, a testing phase took place where participants applied the model to existing assessment approaches and existing data collections. It turned out that in some cases the application of the model was not immediately obvious to the testers. In this question, we aimed to get some views from the community representatives as well as from the funders what the main challenges of the application of the model could be.

Communities

While the FDMM has proven to be of value to the respondents, some expressed challenges of applying the FDMM. Indicators are expressed following the FAIR

principles, hence overlaps exist and priorities are not always aligned with the communities. Indeed, the fact that community practices are not convergent makes it very difficult to apply FDMM.

Additionally, the assessment details in the FDMM could be improved, the interpretation of maturity indicators is still arbitrary to a certain extent, and they should not be technology/service specific. More importantly, any practical tests against the indicators should be built on existing community practices.

Using the FDMM to evaluate FAIRness assumes that an existing dataset in its final location is evaluated; it cannot deal with future choice nor evaluation in earlier stage. However, evaluating the consequences on FAIRness of choices made in a project are essential.

Funders

According to the funders the FDMM lacks discipline-specific implementation guidelines to support researchers. It is important to operationalise the FDMM. Funders also observed that in its current version, the FDMM contains too many indicators to assess.

Additionally, it is challenging, in an evaluation methodology built on top of the FDMM, to identify the mark above which a resource is or isn't "FAIR enough".

2.2.5 What are your suggestions for further work on FDMM?

The model as published in June 2020 can be seen as a first result based on discussions and consensus on a basic set of indicators, priorities and maturity levels for the assessment of FAIRness. It is not intended as the final answer to the assessment of FAIRness, and the WG will continue to work on it and possibly publish newer versions in the future. With this question, we aimed to get suggestions from the two stakeholder communities for further activities of the RDA WG.

Communities

With respect to the future of the FDMM, communities provided various ways for further improving the model. In the first instance, additional iteration and testing of the model, while gradually improving its guidance is needed. To facilitate this process, a feedback loop should be initiated, collating information on the practical tests. The reach of the model should be extended as there is interest beyond the research community (e.g. libraries).

Once the process is in place, a way to advertise success stories based on adoption use cases could be to publish a paper with all active participants as co-authors.

Concretely, additional feedback on the indicators and especially on their priority levels should be gathered from a wide variety of use cases and implementations. Additionally, actionable best practices – tailored for the communities – helping managing data and improve FAIRness should be added to the model.

Funders

Funders suggested to create discipline specific variants of the FDMM informed by research communities. Additionally, the baseline methodology, on top of which evaluation mechanisms can be further developed, should be extended to concrete evaluation and assessment methodologies that can be of use for funders that need to evaluate the FAIRness of the data that is created by funded projects.

2.2.6 Summary

With respect to the future of the FDMM, both the funders and communities agreed to the need of creating actionable guidelines and making the assessment approach more practical and tailored, i.e. more operational details and discipline-specific guidance. In that context, maturity levels serve primarily as an

educational tool for researchers and stewards to understand what the important aspects are and where improvements can be made. To facilitate that, it could be useful to have more information on the priorities of the steps to be taken towards improvement.

It was mentioned that it is important to be realistic and flexible, not taking the FAIR principles as articles of faith and acknowledge that the journey towards FAIRness is dependent on community practices and targets. This would allow the widest range of communities to be involved and create a global, cross-discipline network of research data that can help to tackle major societal challenges.

From the standpoint of the communities, scoring is considered to be useful for evaluation of existing data and for identifying areas where improvement can be made, but misuse should be avoided, and, in particular, approaches where the result of an assessment declare data to be unFAIR.

It was also highlighted that further work on the FDMM is needed, especially on gathering feedback and gaining visibility through impactful adoption use cases and success stories.

According to the funders, the FDMM has too many indicators and remains on a high level. However, selecting the most relevant indicators in specific contexts could be relevant. Besides, funders expressed their willingness to set the bar 'FAIR enough', while acknowledging the divergences in community practices.

3 Conclusion

This survey aimed at improving the understanding of benefits and challenges of the FAIR assessments from the perspective of the funders and communities.

Given the fact that the data reported in the document was collected from a small set of ten respondents, the views might not be exhaustive and therefore a wider survey could be taken up to explore the ideas further.

For future research in this topic, specific questions could be used to identify concrete actions and solutions supporting the opinions expressed in the survey. It would also be useful to reconduct the survey in a years' time to evaluate whether the opinions have changed.

In that regards, two strands of work were identified to help defining concrete actions and solutions:

- **Deepen** – Work on different elements to understand them better (e.g. metadata, granularity etc.)
- **Widen** – Follow up on the FDMM indicators and the FAIR assessment methods in different domains, sectors (e.g. private and public research) and regions.