



Advancing FAIR Agricultural Data: The AgReFed FAIR Assessment Tool

PRACTICE PAPER

CHRISTIANE BAHLO 

 ubiquity press

ABSTRACT

The FAIR principles provide guidance for improving the findability, accessibility, interoperability and reuse of research data and other digital objects. The Agricultural Research Federation (AgReFed) has developed a FAIR implementation consisting of policies, resources and tools to enable FAIR agricultural data in Australia. It prescribes minimum acceptable data standards and an associated set of FAIR metrics for agricultural research data. Existing FAIR assessment tools were examined and found to have limitations in serving the purposes of AgReFed.

The AgReFed FAIR assessment tool addresses these needs with novel features to help improve FAIRness of datasets and other digital resources. By providing the ability to assess and subsequently re-assess datasets while also enabling users to add comments, it facilitates building work lists that support and document the improvement in FAIRness. Other innovative features include customisable FAIR metrics, storage of assessment results, a versioning system and inbuilt help resources. In addition to user-friendly reporting of AgReFed standards compliance, the integrated automated F-UJI FAIR assessment API provides a supplementary set of machine-readable FAIR scores.

The AgReFed FAIR Assessment Tool has been deployed for public use and released as open source to help data custodians enact FAIR principles in domains and data communities within and beyond Australian agriculture.

CORRESPONDING AUTHOR:

Christiane Bahlo

Centre for eResearch and
Digital Innovation (CeRDI),
Federation University, Australia
c.bahlo@federation.edu.au

KEYWORDS:

AgReFed; FAIR data; FAIR
assessment; agricultural data

TO CITE THIS ARTICLE:

Bahlo, C. 2024. Advancing
FAIR Agricultural Data: The
AgReFed FAIR Assessment
Tool. *Data Science Journal*, 23:
18, pp. 1–11. DOI: [https://doi.
org/10.5334/dsj-2024-018](https://doi.org/10.5334/dsj-2024-018)

Agricultural data is generated from a variety of sources in increasing volumes (Osinga et al. 2022) and cost (Misra et al. 2022). There is a recognised need for data sharing between researchers and the broader community (Top et al. 2022), but data is often locked up in inaccessible formats and locations, lacking metadata, licence information, attributions and provenance information. This inhibits data discovery, reuse, and federation, which hinders potential new discoveries and innovations (Bahlo et al. 2019; Osinga et al. 2022; Devare et al. 2021).

These issues exist across multiple domains, and the FAIR Principles have emerged as an overarching guide to help make data more findable, accessible, interoperable and reusable (Wilkinson et al., 2016). The FAIR principles are intended as a broad guide that is domain-agnostic, but interpretation and implementation, including the development of specific metrics to measure FAIRness, require further effort (Jacobsen et al. 2020), which includes community involvement (Top et al. 2022; Mons et al. 2020) and may vary between domains (Gehlen et al. 2022; van Reisen et al. 2020).

FAIR implementations seek to leverage investments in digital objects and foster cross-domain data federation to facilitate new discoveries by, for example, illuminating some of the climatic, environmental, and societal linkages with agriculture (Sales et al. 2020). Digital objects include datasets and associated metadata (Caracciolo et al. 2020; Soares et al. 2022), as well as models and algorithms (Torres, Balbi and Villa 2023; Rosenzweig et al. 2013), vocabularies and ontologies (Subirats-Coll et al. 2022; Arnaud et al. 2020) and software (Lamprecht et al. 2020; Katz, Gruenpeter and Honeyman 2021; Hasselbring et al. 2020).

Specific FAIR implementations were first developed in the life sciences (van Reisen et al., 2020) and then in other domains (Betancort Cabrera et al. 2020). More recently, agriculture-specific implementations have begun emerging (Top et al. 2022). To support different domains and data communities, the WorldFAIR Project (WorldFAIR 2022) is developing FAIR Implementation Profiles (FIP) (Schultes et al. 2020), but to date no agriculture FIP exists. The Agricultural Research Federation (AgReFed) is an Australian cooperative of communities of data providers, whose technical council developed a set of FAIR thresholds for data and services (Wong et al. 2022). Existing FAIR assessment tools (both manual and automatic) use numeric scoring systems or star ratings, which are unable to express the minimum thresholds adopted by AgReFed.

BACKGROUND

AgReFed was initiated in 2018 with funding from the Australian Research Data Commons (ARDC) to address challenges with FAIR data in agriculture. During the enactment phase, a community of agricultural research data users and providers was formed, and the governance and stewardship model was developed (Box et al. 2020).

AgReFed's common vision is to 'enable Findable, Accessible, Interoperable and Reusable (FAIR) agricultural data to accelerate innovation and increase the profitability and sustainability of Australian agriculture' (Agricultural Research Federation 2023). The AgReFed community developed a domain-specific FAIR assessment that defines thresholds based on minimum standards, acceptable standards and stretch goals (Wong et al. 2022). AgReFed recognises that FAIR is a journey that is supported through education and tooling (Mons et al. 2020; Devaraju et al. 2021) and that FAIR assessments are among the tools and technologies that assist with the provision of FAIR data (Krans et al. 2022; Thompson et al. 2020).

AgReFed provides infrastructure, resources, and tools as well as a framework for governance and data stewardship, which are accessible via AgReFed's website.¹ AgReFed, being a data federation, requires data and access controls to be maintained by data providers, ideally within trusted repositories (Wong et al. 2022). Data providers are also responsible for the creation and maintenance of adequate metadata and for attaining and sustaining an acceptable level of FAIRness (Box et al. 2020). FAIR assessments should be conducted on all agricultural research datasets, or at least on exemplars of similar collections. In reflecting on the development of AgReFed's FAIR thresholds, Wong et al. (2022) suggest an automated assessment tool as a next step.

¹ <https://agrefed.org.au>.

Education, minimum metadata standards and appropriate organisational policies will improve the level of FAIR data. Although this is expected to lessen the need for FAIR assessments over time, they will continue to be useful to data providers in meeting FAIR requirements in the agricultural science community.

MOTIVATION FOR SOFTWARE

Given that FAIR implementations are domain-specific (Jacobsen et al. 2020) and that assessment tools support FAIR implementations (Krans et al. 2022), it can be argued that the agricultural research domain has a need for a suitable FAIR assessment tool. Where existing tools don't provide suitable solutions, a community can overcome this implementation challenge by creating new tools, which in turn contribute to the FAIR ecosystem (Jacobsen et al. 2020). As none of the existing FAIR assessment tools were able to provide the functionality required to support AgReFed's goals, the creation of a new tool was indicated in conjunction with reusing the existing F-UJI Application Programming Interface (API) to test the machine readability of digital resources. The new FAIR assessment tool was released as open source, being a reusable contribution to the broader FAIR ecosystem, and an effort was made to make the assessment tool itself FAIR, all of which are in the spirit of the Open Science concept (Hasselbring et al. 2020).

METHODS

In this work, I introduce the tool by outlining previous research, addressing the limitations of existing FAIR assessment tools and demonstrating how I used the insights gained during community consultation within AgReFed to design software, which is a novel approach to FAIR assessments. I also describe the structure of the assessment tool and demonstrate its functionality.

PRIOR RESEARCH AND EXISTING FAIR ASSESSMENT TOOLS

At the time of writing, dedicated FAIR assessment tools for the agriculture domain could not be found, except for the initial spreadsheet-based AgReFed assessment. However, several assessment tools exist in the broader data community. The Research Data Alliance's FAIR Data Maturity Model Working Group analysed 12 different FAIR assessment tools with regard to alignment with the FAIR principles and their specific facets (Bahim, Dekkers and Wyns 2019); however, the tools were not evaluated with regard to features or usability. Krans et al. (2022) reviewed 10 tools for characteristics such as level of user expertise required, guidance provided, ease of use, types of input and output, tool maturity and recommendations for improvement for making a dataset more FAIR. The FAIR assessment tools listed in both aforementioned reviews, as well as five tools compared by Gehlen et al. (2022), were reviewed as potential candidates for use within AgReFed. The following section lists the outcome of this review.

LIMITATIONS OF EXISTING FAIR ASSESSMENT TOOLS

During the development of the minimum thresholds for AgReFed, exemplar datasets were initially found to vary in findability, that is, whether they had a permanent identifier, a metadata record and whether that metadata record was indexed in a searchable repository (Wong et al. 2022). Fully automated assessment tools are unable to capture any FAIR metrics in scenarios where a dataset doesn't at least have a Uniform Resource Locator (URL) and a machine-readable metadata record. Systematic analyses testing the FAIRness of agricultural datasets using the FAIR Evaluation Services (Bahlo and Dahlhaus 2021) and the F-UJI API (Petrosyan et al. 2023) illustrate both the potential and the limitations of automated assessment tools.

Generic FAIR assessment tools lack domain specificity. Every domain has its own standards, or de facto standards, conventions, accepted file types and licences, and therefore should have its own set of FAIR metrics (Genova et al. 2021). The FAIR Evaluation Services addresses this issue through an open authoring framework where communities can create and publish their own metrics (Wilkinson et al. 2019), whereas other assessment tools provide a fixed set of generic or domain-specific FAIR metrics. A further consideration is that metrics can become outdated as data communities evolve, making a case for updatable tests.

FAIR assessment results should be permanently recorded for purposes such as compliance checking and auditing when required, as well as for future analysis. Some online tools will save all results in a public listing, which can be accessed at any time (for example, in the FAIR Evaluation Services – <https://fairsharing.github.io/FAIR-Evaluator-FrontEnd/#!/collections/new/evaluate>), whereas others allow the export of the assessment result (e.g., JSON output from the online F-UJI Tool – <https://www.f-uji.net>). However, most assessment tools only display the results as a percentage or graphic representation without allowing them to be saved.

None of the tools investigated allow the reassessment of digital resources that have already been tested or provide for an easy comparison of multiple assessments of a resource. A reassessment of digital resources in existing assessment tools is possible only by re-entering all information as a new assessment. Even where it is possible to export the assessment results, comparisons of an initial FAIR assessment with subsequent assessments are not straightforward and require manual processing. AgReFed places a strong emphasis on the journey towards improvement in FAIRness. Therefore, an easy comparison between the initial and subsequent assessment results is considered essential.

In most cases, assessment tools provide no support to users moving towards improving the FAIRness of digital resources beyond providing a score (Krans et al. 2022). AgReFed intends to facilitate self-assessment. This requires that users are provided with contextual help and linked resources, particularly in the early stages. Help resources within the user interface are limited in most of the assessment tools. Therefore, prior knowledge of FAIR is required; otherwise, a user must search for information online while conducting the assessment.

BRIEF OVERVIEW OF FUNCTIONALITY AND RUNNING ENVIRONMENT

The AgReFed FAIR Assessment Tool is a web-based application that runs in a browser and measures a digital resource (at this stage, a dataset, whether tabular, a map-based data product or data delivered by a service) against the FAIR thresholds set out in the AgReFed Technical and Information Policy Suite (MacLeod et al. 2020). The tool is a full-stack PHP application (using the Laravel framework) with a PostgreSQL database server-side and JavaScript (using the Vue framework) client-side. The FAIR assessment questions, selectable answers, scores and ancillary information are loaded dynamically from the database as a JSON object. All user responses and submitted FAIR assessments are saved in the database. Users are required to sign up to use the tool so that usage metrics can be collected and assessments can be attributed to users, as described in the tool's disclaimer. This data collection was approved by Federation University's Human Research Ethics Committee (HREC approval no: A21-046). This approval principally relates to interviews that were conducted in the early stages of AgReFed but also covers the usage metrics for the assessment tool.

STRUCTURE

The AgReFed FAIR Assessment Tool consists of the assessment form, an assessment listing and an assessment results page. Users can access only their own assessments. Additional information on the home page and the help page is accessible to all users.

The assessment form (shown in Figure 1) allows the user to enter details about the dataset to be assessed, including the name, description and reason for the assessment. This section is followed by 14 questions relating to the FAIR principles that are split into the sections Findable, Accessible, Interoperable and Reusable. Every section heading and every individual question have clickable help links, which open pop-ups containing external supplementary information.

All questions are multiple-choice. The selected answer will give immediate visual feedback in the form of a bar graph and pop-up that indicates whether AgReFed thresholds are met. Additional information can be entered for each question. Depending on the selected answer for each question, evidence is requested (for example, URLs) in support. A user can also select implementation status and enter other comments, for example, additional evidence or to-do notes.

Form contents are automatically saved after every change so that a user may reopen an unsubmitted assessment for editing at any time. This also prevents data loss if the form is accidentally closed. When complete, the assessment can be submitted, and upon submission, the results page will be displayed. When a user has chosen to reassess a dataset, the same assessment form will be opened, pre-filled with the answers from the most recent assessment.

AgReFed FAIR Assessment Form

Version 1.1

Assessment Details

Name of digital resource: Corangamite Soil Health Monitoring Program Data

Description of digital resource: A set of 100 soil health monitoring sites from across the Corangamite CMA region of Victoria.

Reason for assessment: Initial assessment

? FINDABLE

Q1 The data product has been assigned (an) identifier(s)

Globally unique, citable and persistent identifier (e.g. DOI, PURL, or Handle) ?

Please provide: Globally unique, citable and persistent identifier (e.g. DOI, PURL, or Handle)

Status: fully implemented

Notes:

Agrefed compliance level:
Acceptable (ideal)
 Score: 8

Q2 The data product identifier is included in all metadata records/files describing the data

Yes ?

Please provide link to the metadata record that contains the data identifier. Where possible provide the link to your most comprehensive 'point of truth' metadata record here - for example, your

Figure 1 FAIR assessment form showing first two questions and visual user feedback.

The assessment listing displays a list of assessments that the logged-in user has created previously. The entries in the list indicate the status of each assessment, which may be editable (if an assessment is a draft that has not yet been submitted) or not editable (if the assessment has been submitted). Each assessment in the list has a clickable link to view the complete assessment.

The assessment results page displays results for one or more assessments of a single digital resource. It consists of three sections: the first describes the resource, the second shows scores, and the third (and main) section contains tabs for each assessment and reassessment (where applicable). [Figure 2](#) shows an example of a result with multiple assessments.

In the scores section, bar graphs indicate assessment scores for each of F, A, I, R categories, as well as the overall FAIRness of the assessment (green) and the machine readability scores obtained from the embedded F-UJI API (blue). If a resource was reassessed, like in the example shown, these bar graphs are stacked, showing the assessment and reassessment scores below to clearly indicate any improvement in FAIRness.

The main part of the results page shows the AgReFed compliance level for each of the 14 metrics, together with additional information that the user entered for each question. Where the user has added notes such as actions to take, this section of the results page will serve as a to-do list for improvements to undertake before attempting a re-assessment.

Before an assessment is submitted, it can be resumed and edited, as indicated by the presence of an edit button. If an assessment has already been submitted, it cannot be edited again, but the user can choose to re-assess the digital resource. For each assessment, results can be printed in Portable Document Format (PDF) or in Comma Separated Values (CSV) format.

The home page contains a brief statement about the purpose of the tool and links to the help page, the AgReFed website and the Plain Language Information Statement (PLIS) that informs and relates to the ethics approval. For signed-in users, the home page also displays links to start a new assessment and to list existing ones.

The help page provides information about the tool and the related AgReFed technical policy. It also contains help and advice on the features and how to use the assessment tool.

AgReFed FAIR Assessment Result

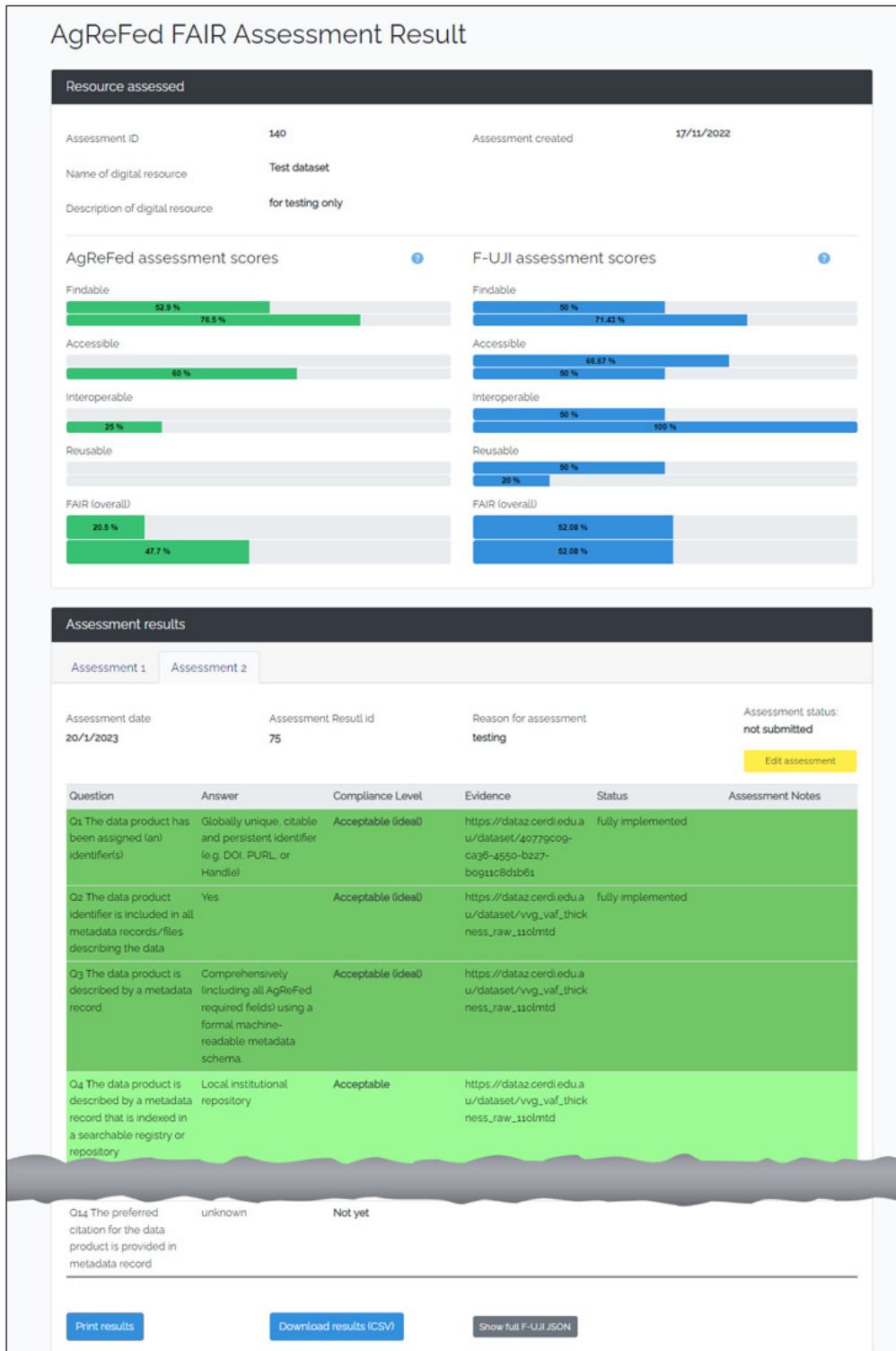


Figure 2 FAIR assessment results page example (abbreviated).

DISCUSSION

This FAIR assessment tool has been developed specifically to address the needs of a data federation in the agricultural sector and is the result of a community approach to greater FAIRness of data in this domain.

The Agricultural Research Federation (AgReFed) Technical and Information Policy Suite (MacLeod et al., 2020) describes a set of 14 FAIR questions with corresponding minimum, ideal, and stretch goal requirements for agricultural datasets. These formed the basis of a spreadsheet containing the questions and associated possible answers, FAIR thresholds and required evidence that make up the AgReFed FAIR assessment. Based on this initial spreadsheet version, the AgReFed FAIR Assessment Tool was developed as a web-based application that presents these questions in a form with multiple-choice answers. This tool also implements the learnings and recommendations of Wong et al. (2022) in relation to FAIR assessments that followed the completion of the initial stages of AgReFed.

During the development and early testing of the alpha version of the tool, additional user interface features were identified, subsequently implemented and further refined during beta testing. These additional features aimed to improve the user experience as well as the usefulness of the tool in a data custodian's journey to improve the FAIRness of a digital resource.

Unlike other FAIR assessment tools, its primary purpose is not to provide a FAIRness percentage or other arbitrary score but rather to determine whether minimum or ideal thresholds (as set out in the AgReFed Tech Policy document) are met by a digital resource. The tool provides clear and unambiguous feedback to the user both visually, through colour coding, and in written form.

A scoring system was added to show improvements in FAIRness (if any) when a digital resource is reassessed. For example, the findability of a dataset is improved by moving it from local storage to a place on the Internet and making it accessible via a URL. As the AgReFed minimum threshold requires a URL with a unique and persistent identifier (e.g., a Document Object Identifier (DOI)), another URL will not suffice to achieve compliance. However, any URL is an improvement over none, and therefore the data custodian going through the process of making the dataset FAIRer obtains improved scores to indicate progress, even if the AgReFed minimum requirements are not yet met. The scores for each assessment are shown as bar graphs, giving an indication of the increasing levels of F, A, I, R, and overall FAIR score from the first to subsequent assessments. This feature was added to encourage data custodians in their efforts to make resources more FAIR.

To further encourage and support the progression of digital resources from less to more FAIR, the tool includes entry fields to indicate the implementation status and freeform comments or notes. This seemingly simple additional functionality elevates the tool from a mere assessment tool to a more integrated solution that supports improvement towards AgReFed minimum standards (or higher). A data custodian can use the notes field so that the assessment output can serve as a worklist for future improvements and, if required, the list can be printed. Alternatively, the notes field can be used to provide additional evidence or other relevant comments.

The AgReFed FAIR Assessment Tool uses a hybrid approach; it is manual in the sense that the data custodian or other person running the assessment must enter data into a form and select answers for multiple-choice questions. However, it is integrated with the F-UJI tool's API, which performs an automated assessment and provides feedback on the machine readability of the digital resource to be assessed ([Devaraju and Huber 2021](#)). As previously stated, an automated tool will not be able to provide any scores for datasets that are not online or not listed in a catalogue; therefore, it will not be able to support or give good feedback to data custodians whose resources are initially at a very low level on the FAIR scale. AgReFed's goals include providing guidance and support, so a manual approach incorporating plentiful help and feedback was chosen. Once a digital resource has been improved to where it is findable and has machine-readable metadata, integration with the F-UJI tool will yield additional insights regarding the machine-readability of metadata related to interoperability and reusability. This approach combining the advantages of manual and automated approaches is supported by the findings from assessment tool evaluations ([Gehlen et al. 2022](#); [Krans et al. 2022](#)) and has been adopted in other domains ([Bonello, Cachia and Alfino 2022](#)).

Access to saved assessment results is available in the list of assessments previously run within the tool's interface. They can also be printed and exported as CSV, and the raw F-UJI score can be viewed as JSON. This allows data custodians to inspect, save and query records of FAIR assessments for purposes such as reporting, quality control or other needs.

When an assessment is run in the web interface, every change is saved so that the user can return at any time to continue adding and editing until the assessment is completed and submitted. An assessment can be edited as many times as necessary without a time limit until it is completed and the user submits the form.

Once submitted, any digital resource may be reassessed. A reassessment will be pre-populated with the most recent assessment results to speed up data entry. Where a digital resource has been assessed more than once, all sets of assessment results are displayed on the results page. The AgReFed compliance levels are shown in a separate tab for each assessment on that resource, and scores are displayed as stacked bar graphs for a visual representation of the level of FAIRness and the improvement(s) between assessment iterations.

The FAIR assessment questions, metrics and other supporting information are dynamically generated from JSON fields in the database, so that it is possible to:

- upgrade the metrics in subsequent versions,
- substitute the metrics for another set that may suit another organisation or domain, and
- use different metrics for different digital resources (currently supports datasets)

This means the tool can be adopted for other purposes without substantial changes to the source code.

SOFTWARE ACCESSIBILITY

The source code is available at <https://doi.org/10.25955/22105712> and on GitHub at <https://github.com/AgReFed/agrefed-fair-assessment>.

A working version of the AgReFed FAIR assessment tool is available for public use at <https://assessment.agrefed.org.au>.

CONCLUSIONS AND FUTURE WORK

Prior to tool development, community consultation and the efforts of the technical committee resulted in the AgReFed minimum standards and a set of 14 questions (Wong et al. 2022). This recognises the importance of the community-driven aspect of FAIR (Wilkinson et al. 2019; Gehlen et al. 2022; Sansone et al. 2019).

No comparable alternative domain-specific set of FAIR questions exists for comparison with the AgReFed standards. To ensure they are fit for the purpose of the agricultural sector in Australia, these questions were used to test seven datasets through a manual process enacted by a data steward during the first stage of AgReFed. These datasets included plant, soil and climate variables and consisted of a variety of data formats, including point observations, sensor, spatial and temporal data from different providers (Wong et al. 2022). These datasets were assessed at the outset and after efforts to improve FAIRness with the help of the data steward. Applying the 14 FAIR questions to actual use cases showed they provided a workable solution. The questions, possible answers, supporting documentation and minimum standards in the FAIR assessment tool are based on this work.

Through an investigation of the literature and existing FAIR tools, shortcomings of existing tools were identified, and novel solutions were implemented while addressing the specific needs of the AgReFed community. The approach taken for AgReFed FAIR Assessment Tool is supported by the recommendation by Gehlen et al. (2022), which is a 'mature hybrid approach', where hybrid relates to a mix of automatic and manual assessment methods and maturity relates to community involvement. Krans et al. (2022) recommended that tool developers provide built-in guidance for users and provide a fully developed tool published as licenced open source. These recommendations were implemented here. However, the AgReFed FAIR assessment tool goes one step further by using an architecture that facilitates the use of other (or multiple) sets of FAIR metrics.

Learnings from the analysis of other FAIR tools guided the development of the AgReFed FAIR assessment tool so that it will better suit the needs of the agricultural community. It is hoped that it can also provide ideas and incentives for the development and provision of better tools in other domains.

The AgReFed FAIR assessment tool has undergone preliminary user acceptance testing. Testing was conducted in-house, and a link to a test version of the tool was circulated to AgReFed partners. The tool was also demonstrated to ARDC and other partners. Feedback received during the early stages (mainly user interface design improvements) was incorporated into the tool, and later feedback has been added to future work. At the time of writing, 27 users had signed up for the assessment tool.

It was released as open source code and as an online tool as part of the AgReFed website in December 2022. This release addresses one of the milestones of the AgReFed platform project,

which is an ARDC-supported platform transforming agricultural research. The assessment tool webpages and database of tested resources (and FAIR scores gained) will be monitored to collect statistics as part of informing the evolution and monitoring the impact of AgReFed on enabling FAIR agricultural data over time. The public Github repository will also be monitored. These statistics will provide feedback regarding the acceptance and usefulness of the assessment tool and guide future development.

The current suite of FAIR assessment tools (including the AgReFed FAIR Assessment Tool) relate to datasets (or collections of datasets) more than to other digital resources within the FAIR ecosystem. There is a significant overlap of metadata requirements between these, but it has been recognised that research software, models, algorithms, services, ontologies and vocabularies should be assessed differently for datasets (Katz, Gruenpeter and Honeyman 2021; Hasselbring et al. 2020; Koers et al. 2020; Torres, Balbi and Villa 2023; Amdouni and Jonquet 2022). Therefore, ideally, an assessment tool should use specific questions, supporting assessment help, evidence and possible answers for each of those resource types. As the AgReFed FAIR Assessment Tool has been designed to dynamically load a set of questions (and related information), the functionality to provide different assessment types is already accommodated within the tool, so it would require only minor additional changes for full implementation. More work is required to develop sets of questions for different digital resources, which could promote recognition of research software and vocabularies as first-class research outputs. Such recognition is identified as a vision both nationally by the Australian Research Data Commons (Australian Research Data Commons 2022) and internationally by the Research Data Alliance (Gruenpeter 2021).

It is expected that FAIR implementations will continue to evolve (Sustkova et al. 2020). This may mean the FAIR assessment metrics need to be updated. The updating process can be accommodated in the tool through the same dynamic metric loading system mentioned above.

For example, it has been recommended to add data quality to FAIR metrics (Stausberg et al. 2022), as FAIR does not address the quality of data. In line with most FAIR implementations, the AgReFed FAIR metrics currently make no mention of data quality beyond requirements for open and interoperable file formats. Accordingly, data quality tests were not considered in the design of this assessment tool but could be added by updating the dynamically loaded metrics, if required by the AgReFed Technical Committee.

Another possible further development of the assessment tool is an improvement in the integration with the F-UJI API. Because F-UJI is open source, its repository could be forked and the code base edited to bring the assessment metrics more in line with AgReFed minimum standards. That way, the F-UJI score would more closely resemble the AgReFed threshold in the future for cases where the resource is machine-readable. Additionally, the results from the automated assessment could be used to pre-populate some of the answer fields for the AgReFed FAIR questions.

Other possible features for future versions include a bulk export of assessment scores to facilitate reporting for entities that host many digital resources, user summaries showing the number of assessments, an average FAIR score and the ability to identify low-scoring assessments for reporting and follow-up activities. The addition of an option to delete a user account has been recommended and is planned for a future software release.

Documented benefits of data sharing and reuse in the broad research community (Data sharing and the future of science 2018) led to the FAIR principles, which provide guidance on data management and stewardship (Wilkinson et al. 2016). Agricultural research and industry, which produce a broad and diverse range of datasets, can benefit from applying FAIR principles to data management (Top et al. 2022), particularly in view of their overlap with the environmental sciences, animal health, soil sciences, weather and climate science and other domains. AgReFed's FAIR assessment tool is intended to help the agricultural research community create and maintain FAIR data and metadata by providing valuable feedback on whether agricultural datasets meet community-developed criteria for findability, accessibility, interoperability and reusability. The application of objective FAIR measurements of datasets can also help inform the agricultural research community on whether the collective goal of increasing FAIRness is being achieved.

ACKNOWLEDGEMENTS

I acknowledge the work of the developers of the F-UJI FAIR assessment tool, whose open source API was integrated into the AgReFed FAIR assessment tool.

I would like to thank Helen Thompson, Megan Wong and Judi Walters for feedback on this manuscript, and Richard Archer and Scott Limmer for code reviews and software deployment.

FUNDING INFORMATION

The AgReFed Platform project received investment (<https://doi.org/10.47486/PL005>) from the Australian Research Data Commons (ARDC). The ARDC is funded by the National Collaborative Research Infrastructure Strategy (NCRIS).

COMPETING INTERESTS

The author has no competing interests to declare.

AUTHOR AFFILIATIONS

Christiane Bahlo  orcid.org/0000-0003-4185-5542

Centre for eResearch and Digital Innovation (CeRDI), Federation University, Australia

REFERENCES

- Agricultural Research Federation.** 2023. *About us*. Available at <https://www.agrefed.org.au/WhoAreWe> [Last accessed 5th February 2023].
- Amdouni, E and Jonquet, C.** 2022. FAIR or FAIRer? An integrated quantitative FAIRness assessment grid for semantic resources and ontologies. *Metadata and Semantic Research*. DOI: https://doi.org/10.1007/978-3-030-98876-0_6
- Arnaud, E, et al.** 2020. The ontologies community of practice: A CGIAR initiative for big data in agrifood systems. *Patterns*, 1(7): 100105. DOI: <https://doi.org/10.1016/j.patter.2020.100105>
- Australian Research Data Commons.** 2022. *A national agenda for research software*, 28 March 2022. DOI: <https://doi.org/10.5281/zenodo.6378082>
- Bahim, C, Dekkers, M and Wyns, B.** 2019. *Results of an analysis of existing FAIR assessment tools*. Research Data Alliance. DOI: <https://doi.org/10.15497/RDA00035>
- Bahlo, C, et al.** 2019. The role of interoperable data standards in precision livestock farming in extensive livestock systems: A review. *Computers and Electronics in Agriculture*, 156: 459–466. DOI: <https://doi.org/10.1016/j.compag.2018.12.007>
- Bahlo, C and Dahlhaus, P.** 2021. Livestock data – Is it there and is it FAIR? A systematic review of livestock farming datasets in Australia. *Computers and Electronics in Agriculture*, 188. DOI: <https://doi.org/10.1016/j.compag.2021.106365>
- Betancort Cabrera, N, et al.** 2020. *White paper on implementing the FAIR principles for data in the social, behavioural, and economic sciences*. Berlin, Germany: RatSWD. DOI: <https://doi.org/10.17620/02671.60>
- Bonello, J, Cachia, E and Alfino, N.** 2022. AutoFAIR-A portal for automating FAIR assessments for bioinformatics resources. *Biochimica et Biophysica Acta (BBA) – Gene Regulatory Mechanisms*, 1865(1): 194767. DOI: <https://doi.org/10.1016/j.bbagr.2021.194767>
- Box, P, et al.** 2020. White paper for the enactment phase of the Agricultural Research Federation (AgReFed). DOI: <https://doi.org/10.5281/zenodo.3707057>
- Caracciolo, C, et al.** 2020. 39 hints to facilitate the use of semantics for data on agriculture and nutrition. *Data Science Journal*, 19(47): 1–12. DOI: <https://doi.org/10.5334/dsj-2020-047>
- Data sharing and the future of science.** 2018. *Nature Communications*, 9(1): 2817. DOI: <https://doi.org/10.1038/s41467-018-05227-z>
- Devaraju, A, et al.** 2021. From conceptualization to implementation: FAIR assessment of research data objects. *Data Science Journal*, 20(4): 1–14. DOI: <https://doi.org/10.5334/dsj-2021-004>
- Devaraju, A and Huber, R.** 2021. An automated solution for measuring the progress toward FAIR research data. *Patterns*, 2(11): 100370. DOI: <https://doi.org/10.1016/j.patter.2021.100370>
- Devare, M, et al.** 2021. AgroFIMS: A tool to enable digital collection of standards-compliant FAIR data. *Frontiers in Sustainable Food Systems*, 5. DOI: <https://doi.org/10.3389/fsufs.2021.726646>
- Gehlen, K, et al.** 2022. Recommendations for discipline-specific FAIRness evaluation derived from applying an ensemble of evaluation tools. *Data Science Journal*, 21(1). DOI: <https://doi.org/10.5334/dsj-2022-007>
- Genova, F, et al.** 2021. *Recommendations on FAIR metrics for EOSC*. Luxembourg: Publications Office of the European Union. DOI: <https://doi.org/10.2777/70791>

- Gruenpeter, M.** 2021. Software as a first class output in a FAIR ecosystem. In: *Workshop on Sustainable Software Sustainability 2021*. (WOSSS21), 6–8 October 2021. DOI: <https://doi.org/10.5281/zenodo.5563028>
- Hasselbring, W,** et al. 2020. From FAIR research data toward FAIR and open research software. *It — Information Technology*, 62: 39–47. DOI: <https://doi.org/10.1515/itit-2019-0040>
- Jacobsen, A,** et al. 2020. FAIR principles: Interpretations and implementation considerations. *Data Intelligence*, 2: 10–29. DOI: https://doi.org/10.1162/dint_r_00024
- Katz, DS, Gruenpeter, M, Honeyman, T, Katz, DS, Gruenpeter, M and Honeyman, T.** 2021. Taking a fresh look at FAIR for research software. *Patterns*, 2: 100222. DOI: <https://doi.org/10.1016/j.patter.2021.100222>
- Koers, H,** et al. 2020. Assessment report on ‘FAIRness of services’. *Zenodo*. DOI: <https://doi.org/10.5281/zenodo.3688762>
- Krans, NA,** et al. 2022. FAIR assessment tools: evaluating use and performance. *NanoImpact*, 27: 100402. DOI: <https://doi.org/10.1016/j.impact.2022.100402>
- Lamprecht, AL,** et al. 2020. Towards FAIR principles for research software. *Data Science*, 3: 37–59. DOI: <https://doi.org/10.3233/DS-190026>
- MacLeod, A,** et al. 2020. The Agricultural Research Federation (AgReFed) technical and information policy suite. DOI: <https://doi.org/10.5281/zenodo.3993784>
- Misra, NN,** et al. 2022. IoT, big data, and artificial intelligence in agriculture and food industry. *IEEE Internet of Things Journal*, 9: 6305–6324. DOI: <https://doi.org/10.1109/JIOT.2020.2998584>
- Mons, B,** et al. 2020. The FAIR principles: First generation implementation choices and challenges. *Data Intelligence*, 2: iii–vi. DOI: https://doi.org/10.1162/dint_e_00022
- Osinga, SA,** et al. 2022. Big data in agriculture: Between opportunity and solution. *Agricultural Systems*, 195. DOI: <https://doi.org/10.1016/j.agsy.2021.103298>
- Petrosyan, L,** et al. 2023. FAIR degree assessment in agriculture datasets using the F-UJI tool. *Ecological Informatics*, 76. DOI: <https://doi.org/10.1016/j.ecoinf.2023.102126>
- Rosenzweig, C,** et al. 2013. The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies. *Agricultural and Forest Meteorology*, 170: 166–182. DOI: <https://doi.org/10.1016/j.agrformet.2012.09.011>
- Sales, L,** et al. 2020. GO FAIR Brazil: A Challenge for Brazilian data science. *Data Intelligence*, 2: 238–245. DOI: https://doi.org/10.1162/dint_a_00046
- Sansone, SA,** et al. 2019. FAIRsharing as a community approach to standards, repositories and policies. *Nature Biotechnology*, 37: 358–367. DOI: <https://doi.org/10.1038/s41587-019-0080-8>
- Schultes, E,** et al. 2020. Reusable FAIR Implementation Profiles as Accelerators of FAIR Convergence. In: Grossmann, G, Ram, S, Grossman, G and Ram, S (eds.), *Advances in Conceptual Modeling*. Cham: Springer International Publishing pp. 138–147. DOI: https://doi.org/10.1007/978-3-030-65847-2_13
- Soares, FM,** et al. 2022. Building a community-based FAIR metadata schema for Brazilian agriculture and livestock trading data. In: *18th International Conference on Semantic Systems*. Vienna, Austria on 12–15 September 2022.
- Stausberg, J,** et al. 2022. FAIR and quality assured data — the use case of trueness. *Studies in Health Technology and Informatics*, 289: 25–28. DOI: <https://doi.org/10.3233/SHTI210850>
- Subirats-Coll, I,** et al. 2022. AGROVOC: The linked data concept hub for food and agriculture. *Computers and Electronics in Agriculture*, 196. DOI: <https://doi.org/10.1016/j.compag.2020.105965>
- Sustkova, HP,** et al. 2020. FAIR convergence matrix: Optimizing the reuse of existing FAIR-related resources. *Data Intelligence*, 2: 158–170. DOI: https://doi.org/10.1162/dint_a_00038
- Thompson, M,** et al. 2020. Making FAIR easy with FAIR tools: From creolization to convergence. *Data Intelligence*, 2: 87–95. DOI: https://doi.org/10.1162/dint_a_00031
- Top, J,** et al. 2022. Cultivating FAIR principles for agri-food data. *Computers and Electronics in Agriculture*, 196. DOI: <https://doi.org/10.1016/j.compag.2022.106909>
- Torres, AM, Balbi, S and Villa, F** 2023. Scientific modelling can be accessible, interoperable and user friendly: An example for pasture and livestock modelling. *PloS One*, 18(2): e0281348. DOI: <https://doi.org/10.1371/journal.pone.0281348>
- Van Reisen, M,** et al. 2020. Towards the tipping point for FAIR implementation. *Data Intelligence*, 2: 264–275. DOI: https://doi.org/10.1162/dint_a_00049
- Wilkinson, MD,** 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,** et al. 2019. Evaluating FAIR maturity through a scalable, automated, community-governed framework. *Scientific Data*, 6: 174. DOI: <https://doi.org/10.1038/s41597-019-0184-5>
- Wong, M,** et al. 2022. Development and governance of FAIR thresholds for a data federation. *Data Science Journal*, 21. DOI: <https://doi.org/10.5334/dsj-2022-013>
- WorldFAIR.** 2022. *The WorldFAIR Project*. Available at <https://worldfair-project.eu/> [Last accessed 23 March 2023].

TO CITE THIS ARTICLE:

Bahlo, C. 2024. Advancing FAIR Agricultural Data: The AgReFed FAIR Assessment Tool. *Data Science Journal*, 23: 18, pp. 1–11. DOI: <https://doi.org/10.5334/dsj-2024-018>

Submitted: 04 September 2023

Accepted: 12 January 2024

Published: 12 April 2024

COPYRIGHT:

© 2024 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/>.

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