Computers and Electronics in Agriculture 226 (2024) 109381

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Making Australian Drought Monitor dataset findable, accessible, interoperable and reusable

Francis Gacenga[a,b,*], Duc-Anh An-Vo[c,d], Jillian McCulloch[d], Richard Young[a], David Cobon[d]

[a] Research Infrastructure, University of Southern Queensland (UniSQ), Australia
[b] Centre for Sustainable Agricultural Systems, UniSQ, Toowoomba, QLD 4350, Australia
[c] UniSQ College, UniSQ, Toowoomba, QLD 4350, Australia
[d] Centre for Applied Climate Sciences, UniSQ, Toowoomba, QLD 4350, Australia

ARTICLE INFO

Keywords:
FAIR data
Drought
Agricultural dataset
Research data management
Cloud computing for agriculture
High performance computing
Digital research infrastructure

ABSTRACT

Making agricultural research datasets Findable, Accessible, Interoperable, and Reusable (FAIR) is an evolving priority for research organisations in Australia. Indigenous data governance standards, described in the CARE (Collective benefit, Authority to control, Responsibility and Ethics) principles complement FAIR principles when managing research datasets. Agricultural research data have traditionally been difficult to publicly access and share due in part to conflicting interests in ownership, commerce, multiparty contracts, and diverse research practices.

As part of an agriculture digital research platform development project (AgReFed Platform project), we develop here a workflow that applies the FAIR data and CARE principles to the Australian Drought Monitor dataset, a product developed as part of the Northern Australia Climate Program (NACP), a joint project funded by Meat and Livestock Australia, the Queensland Drought and Climate Adaptation Program and the University of Southern Queensland (UniSQ). We present here a complete process on how to apply the FAIR principles to the Australian Drought Monitor dataset, including a digital infrastructure development to enable its re-use in the AgReFed Platform project.

1. Introduction

Data shareability, and findability in Agriculture is an ongoing challenge for data collectors and users alike. Agricultural research is experiencing a rapidly evolving and growing data environment, with the potential to facilitate a transformation towards resilient, sustainable, and environmentally sound food production systems globally (Coble et al., 2018; Hackett et al., 2019; Knapen et al., 2023; Petrosyan et al., 2023; Top et al., 2022). There are calls to share agricultural research data between researchers, the research community, data users and data providers (Top et al., 2022) as this is considered crucial both for the advancement of the discipline and ability to meet global challenges (Petrosyan et al., 2023). The FAIR data principles as outlined by Wilkinson et al. (2019) act as a guiding principle to ensure that data are Findable, Accessible, Interoperable and Reusable. Adherence to the FAIR principles furnish agricultural researchers with the concepts and framework to unlock multiple and varied data sources not otherwise accessible, and allows reuse, integration, collaboration, and investigation to answer broad-scope questions surrounding food production worldwide (Hackett et al., 2019; Knapen et al., 2023). As demand for data sharing increases, concerns have been raised about the need to protect the ownership and interests of indigenous owners when integrating Indigenous knowledges and approaches into data practices and policies. The CARE Principles outlined by the Indigenous Data Sovereignty Interest Group within the Research Data Alliance as outlined by Carroll et al. (2020) ensure that Indigenous Peoples retain control, protection, and rights of development of Indigenous cultural heritage embedded in their data. CARE principles embody Collective Benefit, Authority to Control, Responsibility and Ethics and are considered along with the FAIR (Findable, Accessible, Interoperable, Reusable) data principles of Wilkinson et al. (2019) for a 'best-practice' approach for data management. The CARE principles extend the FAIR principles to reverse historic power imbalances that resulted from colonisation, by creating value from Indigenous data that are grounded in Indigenous worldviews and by realizing opportunities for Indigenous peoples within the research sphere (Carroll et al., 2020).

* Corresponding author at: Research Infrastructure, University of Southern Queensland (UniSQ), Australia.
E-mail address: francis.gacenga@unisq.edu.au (F. Gacenga).

https://doi.org/10.1016/j.compag.2024.109381
Received 29 February 2024; Received in revised form 14 August 2024; Accepted 23 August 2024
Available online 6 September 2024

The sharing of research data that is FAIR provides multiple benefits to the research community, corporate and government partners, and end-users of the datasets. FAIR datasets can be validated, evaluated, verified, peer reviewed and facilitate progression for researchers with a lower level of technological readiness, and streamlines the collection and collation of multiple datasets while reducing labour and errors (Coble et al., 2018; Top et al., 2022). Major impediments to achieving these benefits include agricultural data lacking interoperability and standards (Top et al., 2022), agricultural research encompassing multiple data types across multiple sub-discipline and computational agents that undertake data retrieval, management and storage currently existing in a variety of formats that are increasingly decentralised (Wilkinson et al., 2019). There is also a lack of reliable or complete metadata of shared datasets, confidential proprietary data, limited documentation accompanying the dataset, lack of open accepted data formats and APIs (Application Programming Interfaces), undocumented data sharing processes, availability of automated tools for data providers and data users, lack of finance to build the required computational infrastructure and data ownership (Coble et al., 2018; Hackett et al., 2019; Top et al., 2022). As the awareness of FAIR Data Principles spreads in there is a gradual increase in the number of datasets available. Agricultural industry is still in the early stages of incorporating computational data intensive workflows and until a critical mass is achieved, the growth of data services is expected to be slow (Coble et al., 2018; Top et al., 2022).

Implementation of FAIR agricultural datasets is a developing area of interest and necessity, and because of this, knowledge and practice gaps exist in the achievement of FAIR data principles in agricultural research projects. In the present work, we propose to address a FAIR implementation of an agricultural dataset by developing our dataset to be stored on premise at UniSQ and accessible via the AgReFed Platform which aims to be a portal for FAIR agricultural research datasets and workflows. The formal publication of workflows in the agricultural data sphere are vitally important for transparency and reproducibility, and it has been identified that there is a need for greater guidance in data curation and archiving (Hackett et al., 2019; Wilkinson et al., 2019) with the goal to increase effectiveness and transparency of agricultural research and innovation (Top et al., 2022). We present here a case study of how the Australian Drought Monitor (ADM) dataset was made FAIR, shared, and reusable on the AgReFed Platform. The AgReFed Platform project aims to unlock the potential of agricultural research datasets and workflows from organisations in Australia by providing governance, data stewardship and a data sharing platform to increase knowledge, innovation and improve decision making in Australian agriculture (Agricultural Research Federation, 2023a).

There are many stages in making a research dataset FAIR. The process of making the ADM dataset FAIR as a part of the AgReFed Platform project is dependent on an output dataset originating from the ADM project under the Northern Australia Climate Program (NACP). The ADM project dataset was co-owned by the project collaborators who were part of the multi-party agreement that enabled the reuse of the datasets in other projects. The ADM project had many stakeholders, each with their own interest and unique perspectives on challenges and considerations. Stakeholders included researchers wishing to share their data, citation of each other's research publication; institutions ensuring compliance; professional data publishers; software engineers; DataOps (integration and automation of Software Development and IT Operations work); funding bodies (public and private); computational infrastructure providers, data scientists; and the end-users who access the data. We outline in our workflow how we addressed multiple stakeholder interests in ensuring that the dataset is FAIR without breaching proprietary or institutional obligations. Research projects often do not consider the long-term legacy of the dataset and its utility and often insufficient funding is supplied for the personnel and technological components of ongoing data storage, maintenance and management which challenge interoperability (Hackett et al., 2019). The data findability, accessibility, interoperability and reusability were not planned for in the ADM project under the NACP. The AgReFed Platform project and digital research support teams at UniSQ align the dataset with the FAIR and CARE data principles by applying a methodological approach described in the method section.

The present work is not only about defining the FAIR principles for agricultural data. FAIR data principles by their initial definition are almost exclusively about metadata (Wilkinson et al., 2019) and how it is created, shared, and made available to support data re-use. In the present work, we aim to propose and develop an implementation workflow method on how to achieve FAIR for a typical dataset in agriculture. The proposed implementation workflow is not only about providing accompanying metadata that accurately describes the dataset but a comprehensive development including an operable infrastructure for a FAIR agricultural dataset and its results. This includes assigning persistent identifiers to the dataset, making the dataset available through open access repositories, using standard format and protocols for data exchange, and ensuring that the dataset is properly licensed for reuse. Additionally, it is important to ensure that the dataset is properly curated and maintained over time, particularly for datasets with the regular update such as the ADM dataset to achieve its long-term usability. We should also note that there is not a unique implementation to achieve FAIR data in agriculture. We aim to propose and develop one of them.

2. Methods

In the present agricultural research, we aim to unlock the ADM dataset developed in a climate science project for the reuse in agricultural research applications. The data curation, archiving, and formats in climate science can impose challenges to agricultural applications. In the present work, we aim to propose and develop a FAIR implementation workflow for a typical dataset in agriculture which was initially developed in another scientific discipline.

This method section is organised into four sub-sections that define and explain the approach and steps undertaken in this study following a generalised workflow outlined in Hackett et al. (2019) describing the people, places, and processes involved in the transfer of biodiversity data from the field to users. In the first subsection a background on the generation of the ADM is provided. In the second sub-section, the FAIR alignment workflow, we provide an overview of the motivation for applying the FAIR data principles to the ADM then describe the people, places, processes, activities and outputs in the ADM FAIR alignment workflow. The third subsection details the data collector activities undertaken to make the ADM dataset FAIR including FAIR and CARE assessment, contracts negotiation, licensing the ADM, registering and publishing the metadata record, adding the accompanying metadata and linking research outputs to the metadata record. In the fourth subsection we describe the data curator and DataOps activities undertaken including the design and development of a platform to extend and enable the reuse of the ADM dataset and outline the underpinning computing infrastructure for processing, storage and publication of the ADM.

2.1. The Australian Drought Monitor dataset

The ADM (Cobon et al., 2022) was developed by the UniSQ's Centre for Applied Climate Sciences under the Northern Australia Climate Program (NACP) as detailed in Guillory et al. (2023) specialised for Australian conditions (Fig. 1). The ADM comprises a software package developed in Python that calculates Combined Drought Indicators (CDI) at a range of time scales e.g., last month, last three months, last six months and last twelve months, based on the United States Drought Monitor (Svoboda et al., 2002) (USDM). The USDM was developed at the National Drought Mitigation Centre at the University of Nebraska-Lincoln in the late 1990 s. The CDIs are the combination of four different input indices/indicators: 3-month Standardized Precipitation

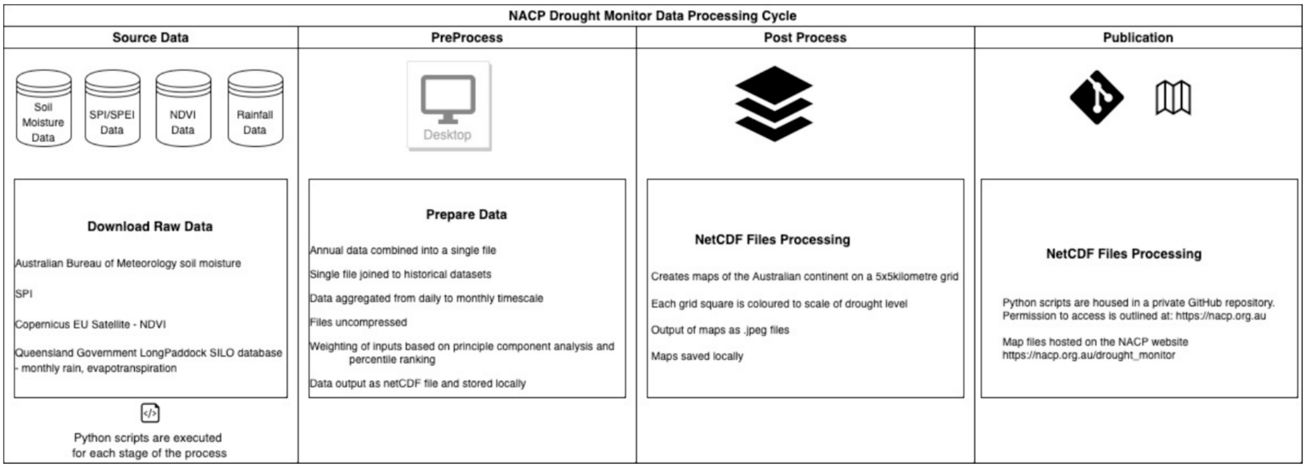

Fig. 1. Process for developing the ADM dataset in the UniSQ NACP-ADM project.

Index (SPI), Soil Moisture (SM), Evapotranspiration (ET) and Normalised Difference Vegetation Index (NDVI). Each input dataset is percentile ranked over a baseline period and the results combined using a weighted average. Principal Component Analysis (PCA) was used to determine the optimal weighting of each input dataset for generating the CDIs at each grid cell (5 km x 5 km) and each time scale over Australia.

The outputs are Australian-continent-coloured maps of the CDI values for all grid cells, representing the local drought condition which is categorised in Table 1. The detailed methodology for generating the ADM dataset can be found in Cobon et al. (2022); Guillory et al. (2023) and the maps are hosted on the NACP website.[1] The dataset and its maps are updated monthly and made available from April 1998. Initially, the dataset and source codes were stored inaccessibly, lacking metadata and without an explicit usage license. This meant that this key dataset and its source codes were not findable, accessible or reusable per the FAIR data principles.

2.2. The FAIR alignment workflow

A plan was developed for the computational and data processing steps (workflow). We aimed to create a FAIR dataset that would be made accessible via the AgReFed platform project, utilising federated cloud architecture interconnecting cloud-based systems for data storage and data backup, a custom-built data server, as well as a local data repository as recommended by Ali and Dahlhaus (2022). We firstly created a metadata record that utilised the DublinCore metadata schema (Dublin Core Metadata Initiative, 2002) and the RIF-CS metadata standard as recommended by Research Data Australia (Australian Research Data Commons, 2013). This has created fully contextualised and descriptive machine-readable metadata, a robust database and record that is easily utilised by humans and machines. This process allows new consumers to reuse data for new needs and applications (Ali and Dahlhaus, 2022; Bahlo and Dahlhaus, 2021; Knapen et al., 2023).

Table 1
Drought categories given by the CDI values.

CDI value	Category
<0.02	Exceptional Drought
0.02–0.05	Extreme Drought
0.05–0.1	Severe Drought
0.1–0.2	Moderate Drought
0.2–0.3	Slightly Dry
0.3–0.7	Near Normal
0.7–0.8	Slightly Wet
0.8–0.9	Moderate Wet
0.9–0.95	Severe Wet
0.95–0.98	Extreme Wet
>0.98	Exceptional Wet

In this project the alignment of the ADM dataset with the FAIR principles (Wilkinson et al., 2019) was achieved by organising the team roles and workflow groupings derived from a "generalized workflow describing the people, places, and processes involved in the transfer of data from the field to users" (Hackett et al., 2019). The implementation of the principles followed a non-linear, iterative path where several processes occurred simultaneously and in parallel as progress was made and goals achieved. The implementation progressed organically and with processes undertaken iteratively to complete tasks. Fig. 2 is presented on a two-dimensional plane showing the FAIR principles alongside the activities undertaken by the team roles of data collector, data curator, and DataOps. This depiction is intended to make it easier to show how the processes, activities and outputs undertaken by the data collector, data curator and DataOps roles aligned with the FAIR principles as originally outlined and referenced to their numbers by Wilkinson et al. (2019).

2.2.1. The FAIR guiding principles for the Australian Drought Monitor dataset

Findability of a dataset refers to the ease with which the dataset can be located, being crucial for promoting data sharing and reuse. It is an important aspect of data management as it ensures that the dataset is discoverable by potential users. To improve the findability of a dataset, it is important to provide descriptive and accurate metadata that includes information such as the title, author, date, and keywords (Australian Research Data Commons, 2020). Additionally, assigning a persistent identifier such as a Digital Object Identifier (DOI) or Unique Resource Locator (URL) can make it easier to locate the dataset. Making the dataset available through a searchable repository or catalogue can also improve its findability.

Accessibility of a dataset refers to the ability of users to obtain and use the dataset for research and other purposes. The accessibility ensures that the dataset is available to all potential users, regardless of their location or technical expertise. To improve the accessibility of a dataset, it is important to provide clear instructions on how to access and use the data, including any necessary software or tools. Additionally, ensuring that the dataset is available in multiple formats, such as Comma Separated Values (CSV), JavaScript Object Notation (JSON), or network Common Data Form (netCDF) can make it easier for users to access and use the data. Providing open access to the dataset, either through a

[1] https://www.nacp.org.au.

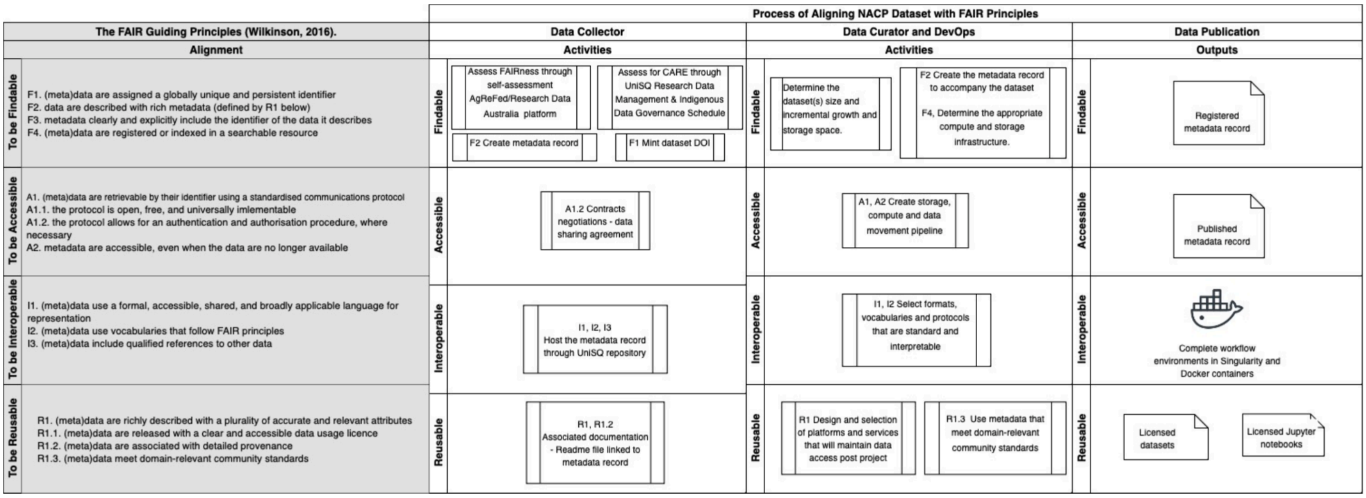

Fig. 2. Outline of the process, activities and outputs in making the Australian Drought Monitor dataset FAIR.

public repository or by releasing it under an open license, can also improve its accessibility.

Interoperability refers to an approach whereby individual data components are coupled in a flexible way that permits data exchange across platforms, allowing to re-use and substitution of data components within an overarching framework. Interoperability also refers to the ability of different systems and applications to communicate and exchange data seamlessly. It is proposed that the availability of better, more trusted, and targeted datasets within an interoperable data framework will result in more effective use of datasets for improved production and environmental management. In the context of agriculture, interoperability enables the integration of different agricultural data sources, such as soil data, weather data, and crop data, to provide a comprehensive view of agricultural systems. This can lead to improved decision-making and increased efficiency in agricultural practices. To achieve interoperability, various instruments such as metadata, schema languages, and ontologies have been developed to support data description and sharing (Cerba et al., 2011). Additionally, initiatives such as AgBioData and API-AGRO have been established to improve data exchange and interoperability between agricultural stakeholders (Sine et al., 2015).

The reusability of agricultural data by humans or machine agents is the main goal of making it FAIR. Reusability refers to the ability to use data for different purposes beyond its original intended use. Agricultural data can be reused for various purposes, including research, policy-making, and informed decision-making. By making agricultural data FAIR, it becomes easier to reuse the data for different purposes, which can lead to increased collaboration and innovation in agricultural research. Additionally, the reusability of agricultural data allows to develop applications, contributing to sustainable agricultural performance and bringing in economic and social benefits.

2.2.2. *FAIR alignment workflow – people, places, processes, activities and outputs*

This section describes the roles undertaken by the authors of this article in making the Australia Drought Monitor dataset FAIR and highlights the contribution to the project by other subject matter experts in UniSQ research support roles. The core project team and support roles are presented in summary in Table 2 and Table 3.

2.3. *A FAIR Drought Monitor metadata record – data collector Activity*

The definition of "data collector" in this workflow is a trained individual gathering, recording, analysing or transforming datasets for transfer to the data curator, and preparing the metadata, license, contract, FAIR and CARE assessments of the dataset (Hackett et al. (2019)).

2.3.1. *FAIR assessment*

The dataset was iteratively tested and progressed for alignment with FAIR principles using the AgReFed platform FAIR self-assessment tool (Agricultural Research Federation, 2023b) which extends the Australian Research Data Commons (ARDC) FAIR assessment tool (Schweitzer et al., 2021) by applying it to the Agricultural domain. The assessment tool is designed to be able to assess the dataset's FAIRness to identify gaps where information is lacking. The AgReFed FAIR self-assessment tool is designed to give the user immediate feedback through a progress-bar that indicates how the database matches FAIR principles. A benefit of using a standardised tool is in overcoming subjectivity in FAIR assessments done manually (Kumar et al., 2024). A challenge experienced when using this tool was that the researcher is unable to get an interim result of the FAIRness of the dataset, especially during the planning stage, because every field must be populated for the tool to give an output. Unless researchers have already planned how they are going to make their data FAIR prior to the project commencement, the tool by design requires multiple iterations to get a valid assessment result while the project is still 'in-progress'. This can be challenging for researchers as they are confronted with trying to work through FAIR alignment post project completion. This was the case for our research and testing with the tool was done repeatedly through the FAIR updates to the dataset until it reached completeness. Once the dataset was deemed FAIR aligned, and the metadata record publishable engagement with UniSQ Library Support and Research Contracts teams was completed.

2.3.2. *CARE assessment*

An associated guide to the FAIR data principles are the CARE principles (Carroll et al., 2020) that acknowledge the right of indigenous peoples to govern the collection, ownership and application of their own data including traditional knowledge (Australian Research Data Commons, 2022). The CARE principles complement the FAIR principles by ensuring that appropriate data sharing and reuse is done with awareness of indigenous governance and stewardship for indigenous data. The Data Collector consulted authors of the UniSQ's Research Data Management and Indigenous Data Governance Schedule (RDM&IDGS) to investigate whether the raw data sources used to generate the Australian Drought Monitor dataset accessed any indigenous data or knowledge. UniSQ's RDM&IDGS are based on the code of ethics from the Australian

Table 2
Core project team roles and activities.

Person	Organisation Role	Project Role	Research Activity	FAIR Principles Alignment Activity
Dr Francis Gacenga	Senior Digital Research Advisor	Funding acquisition	Project Management	Guide FAIR and CARE assessment
		Data Collector	Literature Review	Contracts negotiation and data sharing agreement
		Data Curator	Research Data Management	
		Data Publication	Information Systems Design and Architecture	
Dr Duc-Anh An-Vo	Agricultural Research Data Scientist	Data Curator	Data Interpretation	Authoring Jupyter notebooks and publishing workflows, interoperability testing and documentation for data accessibility and reuse.
		Data Publication	Data Analysis	
			Data Visualisation	
			Platform testing	
Jillian McCulloch	Research Assistant	Data Collector	Literature Review	Assess FAIRness through AgReFed RDA tool
			Research Data Management	CARE assessment through UniSQ guidelines
				Metadata record creation
Richard Young	HPC Systems Engineer	DataOps	Design implementation, coding, scripting and containerisation	Implement storage and compute infrastructure
Prof David Cobon	Associate Professor (Climate and Agricultural Systems Modelling)	Data Collector	Original data contributor	Contracts negotiation and data sharing agreement

Table 3
Project support roles and activities.

Organisation Role	Support Role	FAIR Principles Alignment Activity
Research Contracts Team	Contracting	Guide FAIR and CARE assessment
		Contracts negotiation and data sharing agreement
Research Operations Team	Contracting	
Senior Research Librarian	Curation	Metadata curation,
		DOI minting,
		Licensing
ICT infrastructure team	DataOps	Virtual server provision

Institute of Aboriginal and Torres Strait Islander Studies,[2] the Global Indigenous Data Alliance statement on CARE Principles,[3] the National Health and Medical Research Council Australia's ethical guidelines for research with Aboriginal and Torres Strait Islander People,[4] the Australian Research Data Alliance Practice paper on the CARE Principles for Indigenous Data Governance[5] and the United Nations Declaration on The Rights of indigenous Peoples.[6] Based on this, we found that none of the raw or processed data came under any Indigenous Data Governance Schedules. The CARE principles in this project were applied to ensure the collective benefit of the dataset to the broader community, maintenance of authority to control the dataset via licencing, responsible operation of the dataset through clear role assignment and documented practice and adherence to ethical standards throughout all processes.

2.3.3. *Contracts renegotiation*

It is well documented that proprietary, government and non-government funding bodies, partners or contractual agreements can hinder the sharing of research data (Bahlo and Dahlhaus, 2021; Coble et al., 2018; Hackett et al., 2019; Top et al., 2022; Wilkinson et al., 2019). Before deciding on a creative commons license for the dataset, contracts had to be reviewed between the NACP and AgReFed projects and their respective funding bodies. The NACP-ADM project and subsequent development of the ADM dataset was jointly funded through Meat and Livestock Australia (MLA), the Queensland Department of Agriculture and Fisheries' Drought and Climate Adaptation Program and UniSQ. The ADM dataset is a product subject to several contractual arrangements for use on the AgReFed project.

To ensure that contractual obligations were met and to prevent contract breaches, we consulted UniSQ research contract and legal teams, who provided expert guidance on gaining permission from MLA and the Drought and Climate Adaptation Program (DCAP) in amending the project contracts to include the sharing of the dataset on the AgReFed platform. The consultations were initiated through meetings with the Research Contracts Manager, outlining the aim of the project and rationale for contracts review and amendment. Following the consultations, a written request was made to funding bodies to approve in-principal support for the sharing of the netCDF dataset on the AgReFed platform. Once this was granted, the contracts team amended the AgReFed and NACP contracts to enable sharing of the dataset that incorporates and protects all the partner's interest. Agricultural researchers must take contracts into consideration when planning the sharing of their intellectual property or data, as some contracts stipulate ownership of the intellectual property and the rules surrounding the sharing of datasets. This is often set at the start of a project and should be undertaken as good practice that facilitates responsible sharing and reuse of datasets.

2.3.4. *Licensing the Australian Drought Monitor dataset*

The ADM dataset required licence assignment to protect the research partners right, complete the metadata record and to ensure that end-user's awareness of their obligations when using the dataset (Top et al., 2022). The UniSQ Library Research Support team (Research Librarians that are experts on Licensing at our Institution) was consulted for guidance on the appropriate licensing of the dataset which was shared with the UniSQ Contracts team for alignment with the contractual obligations among projects and any needed amendments. The resultant information was communicated through the DataOps team and a licensing arrangement agreed upon. Accounting for the contractual arrangements surrounding the CDIs, it was decided that a Creative Commons Attribution NonCommercial License ShareAlike 4.0 (Creative Commons, 2023) was the most appropriate choice because it allows widespread and open use of the dataset for all research purposes, with caveats for commercialisation. This aligns with the project goals of AgReFed, in making sure that the data is as FAIR as possible. Three layers of licensing information (Fig. 3a) were then added to the metadata record, the dataset's readme file and its current derivative works

[2] https://aiatsis.gov.au.
[3] https://www.gida-global.org/care.
[4] https://www.nhmrc.gov.au/research-policy/ethics-and-integrity.
[5] https://datascience.codata.org/articles/https://doi.org/10.5334/dsj-2020–043.
[6] https://social.desa.un.org/issues/indigenous-peoples.

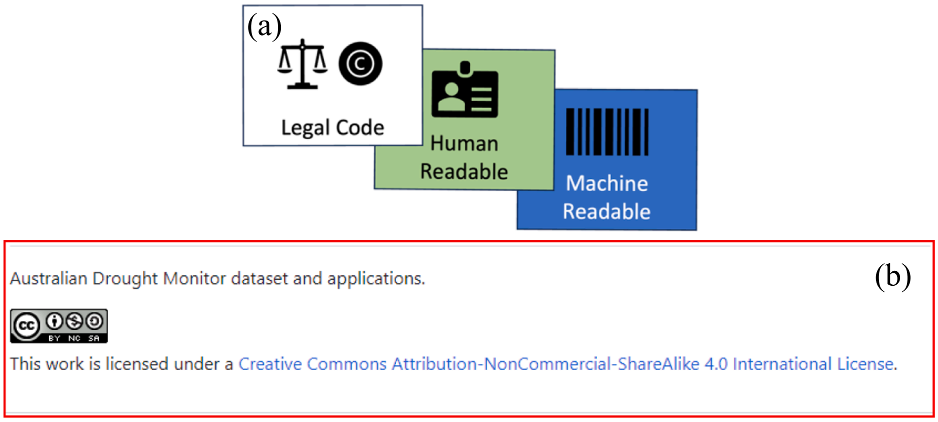

Fig. 3. Three Layers of a Creative Commons Licence (a), and its license badge (b) was imposed to the Drought Status work package of the AgReFed Platform.

such as the NACP website[7] that hosts the post-processed maps and the Drought Status work package of the AgReFed Platform[8] (Fig. 3b).

Legal code (Fig. 3a) is a text document being familiar to legal professionals. For non-legal readers, the license is also available in a format that is understandable – the Common Deed (also known as the "human readable" version of the license). To make it easy for the machine agents and web engines to discover when a dataset is available under a Creative Common license, a "machine readable" version of the license is also provided.

2.3.5. Metadata record registration and publication

A comprehensive metadata record is one of the most important factors in ensuring a dataset is aligned with FAIR and CARE principles. Metadata schemas exist to ensure that a uniform system is applied to datasets and that metadata adhere to recognised standards for datasets and the indigenous governance structures that surround them (Carroll et al., 2020; Wilkinson et al., 2019).

Once a digital object identifier (DOI) was minted by UniSQ library, required fields were populated for the metadata record listing which include an introduction, the spatial region, research areas (fields of research and social-economic objectives), a citation, the DOI, related publications, related information, related organisations, links, contact details, researchers and their ORCiD, temporal data coverage, link to the dataset, licensing and rights.

The metadata record was created, registered and published at UniSQ research data and publication repository (Cobon et al., 2022) and also via the Research Data Australia repository.[9] The metadata follows the Dublin Core schema and adheres to the RIF-CS metadata standard used by Research Data Australia (RDA). The metadata record aims to closely follow the Climate and Forecast Metadata Conventions, with spatial and temporal properties of the data listed (CF Community, 2023). RDA metadata standards were adhered to when creating the metadata record. For model-generated climate data, RDA metadata standards recommend for data to be written in the netCDF format.

2.3.6. Accompanying metadata – ReadME file

The development of the ReadME file is done to ensure that other parties accessing the dataset can understand the context and its original intended usage to make the data interoperable and reusable for their own research purposes. ReadME files are also recommended for use if a metadata record is unavailable (Research Data Management Service Group, 2023), but in this case we used it to enhance the metadata record.

The ReadME file was developed to describe the entire dataset from the Australian Drought Monitor including inputs and outputs, with details for each individual file listed at the bottom of the document. The ReadME file accompanies the metadata record and the dataset to enhance the findability of the dataset and its accessibility. The ReadME file is written in plain text document to enable all formats of computers to be able to read the document. The ReadME file format was based on those of the Cornell University, Research Data Management Service Group (2023).

2.4. A FAIR Australian Drought Monitor dataset – data curator and DataOps Activity

The data curator is supplied the data from the data collector (Hackett et al., 2019). In our case the supplied dataset is the ADM project output of CDIs at a range of time scales stored in netCDF files. The data curator works on developing and augmenting local databases, uploading data to a server, testing and quality controlling the dataset for querying, labelling, transcribing, and managing the data. They oversee the data portal and the dissemination of data. This allows data users to focus on data analysis and ensures that someone is responsible for the quality of available data (Top et al., 2022). In our research project the data curator role was undertaken by the UniSQ AgReFed Project Research Data Scientist.

The role of the DataOps in this workflow was to oversee and manage the process and ensure that the appropriate hardware, software and personnel for the project were obtained at the right time during the project to achieve project goals, timelines and reporting objectives. The DataOps role in this project was undertaken by the UniSQ eResearch Services team comprised of a High-Performance Computing (HPC) Systems Engineer and Senior Digital Research Advisor.

2.4.1. Extending and reusing the Australian Drought Monitor dataset

The data curator developed scientific applications that extended and repurposed the ADM dataset and was also responsible for testing the virtual machines and automation deployed by the DataOps team. The ADM dataset is updated monthly by pulling input data from external databases. The latest ADM dataset is then automatically uploaded into a secure UniSQ on-premises server that pushes the dataset to an UniSQ ARDC Nectar tenanted virtual server and finally to an AgReFed ARDC tenanted Nectar virtual server. The data curator developed agricultural science applications scripted in Python and deployed in Jupyter notebooks that enable end users to access the latest dataset and documentation.[10] To demonstrate and support further reuse of the dataset, the

[7] https://www.nacp.org.au/.

[8] https://github.com/AgReFed/drought-status-notebooks/wiki/Drought-status-workshop.

[9] https://researchdata.edu.au/australian-drought-monitor/2027489.

[10] https://github.com/AgReFed/drought-status-notebooks/wiki/Drought-status-workshop#drought-monitor-dataset.

data curator developed two user cases in separate notebook 1[11] and notebook 2[12] demonstrating the dataset applications. The first user case is a basic application giving step-by-step instruction on how to access and visualise the dataset spatially (maps) and temporally (time series). The second user case is at the research level investigating how to use the CDI dataset for better drought assistance allocation across the Australian wheatbelt. The user cases and their supporting files are published as part of the AgReFed platform tools and datasets (Agricultural Research Federation, 2023a). The tools developed for extending and reusing the ADM dataset meet key effectiveness criteria proposed by (Top et al., 2022) including making automated tools available for data providers and users, applying a community-based approach in developing tools and vocabularies and securing the data owners' rights through appropriate licensing.

2.4.2. Drought Monitor compute virtual Machines for Processing, storage and publication

The DataOps team was responsible for the design, build and deployment of the systems required to reuse the drought monitor dataset and for the interoperation with external systems used to publish the dataset. The DataOps team developed a virtual research environment as outlined in Knapen et al. (2023) that allows researchers to access datasets, storage and computational resources on a secure platform for data exchange, and complete metadata.

An outline of the processes, activities and outputs of the DataOps team and Data Curator are shown in Fig. 4.

The DataOps team worked through several iterations of requirements gathering, building, testing, and piloting candidate solutions. The initial concept design entailed the implementation of Thematic Real-time Environmental Distributed Data Services (THREDDS) data server. This design on testing and minimal viability pilot implementation proved to be resource intensive, onerous and not easy to interoperate with other platforms. A second build and test phase involved the use of virtual machine (VM) infrastructure deploying the CERN (European Centre for Nuclear Research) VM-filesystem (CERN VM-FS) on a Microsoft Azure Proof of Concept pilot. This deployment was an improvement and supported interoperability but was not cost efficient. A third iteration incorporated the learnings and evolved user requirements from the first two iterations and implemented a hybrid pilot solution that included a physical on-premises VM server and ARDC Nectar Cloud Service VMs still using CERN VM-FS this solution worked well and was cost efficient. A final adjustment for further improvement was implemented in the final iteration which entailed simplifying the data flow and processing and integration with the AgReFed platform by using the native operating system, Linux filesystem and NetCDF without the added layer of the CERN VM-FS. Interoperability was further improved by deploying process over open-source services including Jupyter Hub and GitHub services (Knapen et al., 2023). Cost efficiency in the final iteration was achieved by eliminating a hybrid architecture that relied on a low-cost scalable on-premise virtual machine on shared server infrastructure and the use of free to researchers Australian government funded ARDC cloud services VM and no cost open source Linux operating system. A summary of the final iteration deployment process, outputs and alignment with FAIR principles is shown in Fig. 4. Detailed explanations and technical specification DataOps design, build and deployment is currently in preparation for publication in systems journal and will also have its DOI linked to the metadata record.

[11] https://github.com/AgReFed/drought-status-notebooks/blob/main/cdi.ipynb.

[12] https://github.com/AgReFed/drought-status-notebooks/blob/main/info_value.ipynb.

3. Results

The following results were achieved in making the ADM dataset FAIR including:

- Data server with management and access capability to deliver monthly update drought monitor datasets for the AgReFed Platform and the NACP.
- Licensed CDI (Combined Drought Indicator) FAIR dataset published accessibly.
- Drought Monitor metadata published on the RDA (Research Data Australia) registry.
- Documentation of the development of Findable, Accessible, Interoperable, Reusable (FAIR) process
- Drought Monitor Use Cases – for application and training via the AgReFed Platform
- User Training Material

3.1. FAIR ADM dataset

The ADM dataset was made FAIR. The findability was achieved by published metadata record and readme file. Accessibility was achieved by a developed infrastructure, including the on-premises server supplying the local monthly CDI generation to the UniSQ ARDC Nectar tenanted virtual server and finally to the AgReFed ARDC tenanted Nectar virtual server which is a part of the AgReFed platform where credential end users can log in. The AgReFed platform also provides an AgReFed Python environment and allocate one virtual machine to each user when they select the environment enabling access to and work with the dataset by provided documentation on a step-by-step guide.[13] Interoperability was achieved by the published readme file, netCDF format, the developed interconnected servers, and the AgReFed platform proving the computing environment and virtual machines to credential end users. Reusability was enabled by amending the contracts, identifying a suitable license and imposing that license to the dataset and its derivatives.

3.2. GitHub documentation

We used GitHub wiki to provide document on steps to access the dataset in the AgReFed platform.[14] This is a part of the Drought Status Workshop[15] documentation which also include the pre-workshop setup accessing the AgReFed Platform, Python environment and virtual machines, and applications represented in the two use cases. This workshop documentation is useful for training and extension activities and constitutes the legacy of the related projects including the NACP-ADM and the AgReFed Platform.

3.3. Demonstration of reusability by use cases

End users might be interested in how the CDIs can represent realistic drought conditions at locations, supporting farm managers in their management decisions and policy makers in drought relief policies. Two use cases were developed using Jupyter Notebooks and AgReFed Python environment to help users work with the ADM dataset fostering its reusability via such potential applications. The first use case in a developed notebook is about common tasks in using the dataset such as

[13] https://github.com/AgReFed/drought-status-notebooks/wiki/Drought-status-workshop#drought-monitor-dataset.

[14] https://github.com/AgReFed/drought-status-notebooks/wiki/Drought-status-workshop#drought-monitor-dataset.

[15] https://github.com/AgReFed/drought-status-notebooks/wiki/Drought-status-workshop.

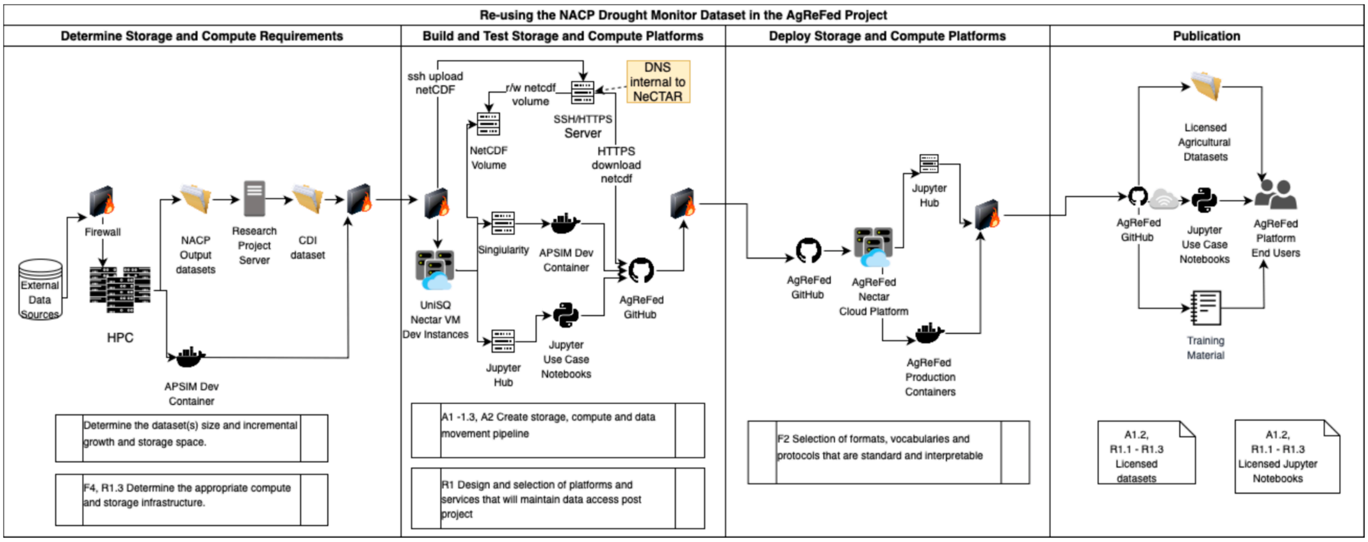

Fig. 4. Reusing the NACP-ADM Dataset in the AgReFed project.

reading and analysing the drought indices[16] including data plotting and mapping. For instances, the users can create an Australia map of spatial CDI values in a current month to understand the drought status (Fig. 5a) or plotting time series of CDI values at a local site to analyse the drought dynamics there (Fig. 5b).

The second use case in another developed notebook[17] is an advanced application at the research level using the generated CDIs to provide objective assistance to minimise the economic and environmental impacts of droughts. We found that using the ADM dataset to inform drought assistance results in objective assistance allocations across 59 wheat production sites of the Australian wheatbelt which are different from those without drought monitor information (Fig. 6).

4. Discussion

Making agricultural research data FAIR is a complex process involving various challenges. A key challenge is the lack of community-based standards and implementation plans for FAIR data in agriculture (Ali and Dahlhaus, 2022; Kumar et al., 2024; Petrosyan et al., 2023). This can lead to inconsistencies in data management practices and hinder data sharing and reuse. Additionally, there is a lack of awareness and training among researchers on FAIR data sharing, which can result in poor data quality and limited data sharing (Devare et al., 2023). Moreover, the diversity in the research data landscape of agricultural sciences can make it difficult to establish minimum thresholds for FAIR data (Wong et al., 2022). This can lead to difficulties in aligning expectations for data delivered from providers' distributed data stores through a community-governed federation. Furthermore, there are challenges, deficits and uncertainties in handling research data in agricultural sciences, which can hinder access and efficient reuse of valuable research data (Petrosyan et al., 2023; Senft et al., 2022). A further challenge is limitations to data governance and knowledge sharing for informed decision-making (Kumar et al., 2024). It thus does not exist a unique practice in achieving FAIR for all agricultural data sources.

In the present work, we proposed a method including detailed development steps in achieving FAIR ADM dataset whose data management was not planned under the NACP-ADM project generating the dataset. These development steps include a standardised metadata record and readme file and their publications in data catalogue services such as the Research Data Australia. Moreover, the dataset is updated monthly and stored in a common netCDF file format fostering interoperability. Such transient dataset is transferred seamlessly among data servers in our developed hybrid cloud infrastructure minimising cost while ensuring a stable service for accessibility. We negotiated a revision of the research contract allowing to access and reuse the dataset in the AgReFed platforms. This is then reflected and guaranteed by identifying and imposing a best suitable license to the dataset and its derivatives. The license and its imposition allow data sharing and reuse without worrying about research contracts and intellectual property commitments. To demonstrate reusability, we developed two use cases in two Jupyter Notebooks using the developed AgReFed Python environment.

The sustainability of FAIR data and service management in establishing open access to research data is also a significant challenge. This includes financial sustainability and funding, as well as legal, governance, and technical issues that concern the scientific communities (Cocco et al., 2020). Making agricultural research data FAIR comes with other costs such as developing community-based standards and implementation plans for FAIR data in agriculture, the establishment of minimum thresholds for FAIR data, and the need for research data policy and data repositories to support data sharing practices (Ali and Dahlhaus, 2022; Devare et al., 2023; Wong et al., 2022). The cost also includes the need for training and education of researchers on FAIR data sharing, which require funding and resources. In the present work, we developed a hybrid cloud infrastructure seamlessly transferring the transient ADM dataset among data servers minimising cost while ensuring a stable service for accessibility and reusability. Furthermore, we provided detailed workshop documentation of dataset access and the two use cases in the AgReFed platform by using opensource GitHub wiki. The documentation can be used for further training and user support.

Agricultural research data can be made reusable by following FAIR and CARE principles and practices. The establishment of community-based standards is considered an essential ingredient in the implementation plans for FAIR data in agriculture. Our project benefited from leveraging the AgReFed network which provided a community base with a guiding structure and standards. The AgReFed network will benefit as use of the AgReFed platform increases and the network expands to have more stakeholders engaging with the research outputs. It is also recommended that researchers be trained on FAIR data sharing and provided with educational material to ensure that they understand the importance and have the capability and skills on incorporating FAIR principles is sharing research data. Our project developed training

[16] https://github.com/AgReFed/drought-status-notebooks/blob/main/cdi.ipynb.

[17] https://github.com/AgReFed/drought-status-notebooks/blob/main/info_value.ipynb.

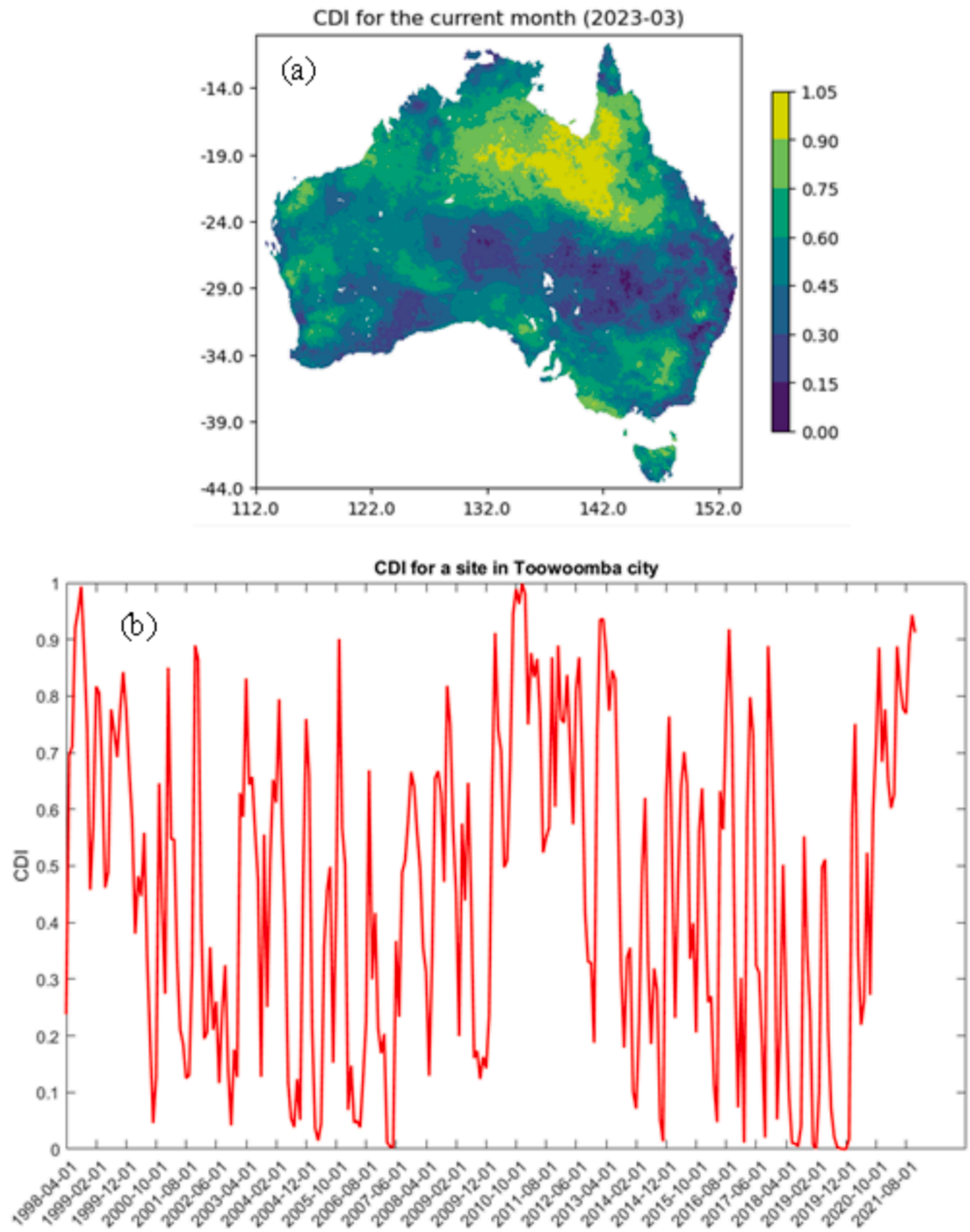

Fig. 5. CDI mapping of the Australian continent for the current month in March 2023 (a) and CDI time series for a site in Toowoomba city (latitude = −27.54, longitude = 151.93), Queensland, Australia (b). The colour schemes in (a) reflect the CDI values and their associated drought categories given in Table 1 (see the method).

materials and workshops that were deployed as extensions of the NACP dataset, aimed at increasing the capability and skills of researchers in achieving FAIR practices. As recommended in Wong et al. (2022) minimum thresholds for FAIR data in agriculture should be established to align expectations for data delivered from providers' distributed data stores through a community-governed federation. Agricultural practitioners should also be trained and educated on big data principles, database management, improved data visualisation, as well as incentives for data sharing for optimizing big data in systems agriculture (Kharel et al., 2020). By making the NACP dataset FAIR and augmenting use cases and training materials an exemplar is provided to aid researchers accept that their responsibilities include the stewardship of data assets to assure long-term preservation, access, and reuse (Devare et al., 2023).

We found as prescribed in Katabalwa et al. (2021); (Kumar et al., 2024) that research data policy and data repositories should be established to support data sharing practices. It has been noted that the benefits of open access to research data go beyond academia and can contribute to socio-economic outcomes such as reducing hunger and poverty and improving human nutrition in the tropics through research aimed at increasing the eco-efficiency of agriculture (Ramirez-Villegas et al., 2012).

It should be noted that the cost of making agricultural research data FAIR can be offset by the benefits of FAIR data sharing, which include increased collaboration, improved data quality, and increased research impact (Ali and Dahlhaus, 2022). FAIR data sharing can also contribute to sustainable agricultural performance, which can have economic and social benefits (Ali and Dahlhaus, 2022).

To make an agricultural dataset FAIR, it is important to ensure that the data is of high quality. Quality data is essential for making the data findable, accessible, interoperable, and reusable. There are several factors that contribute to the quality of an agricultural dataset, including

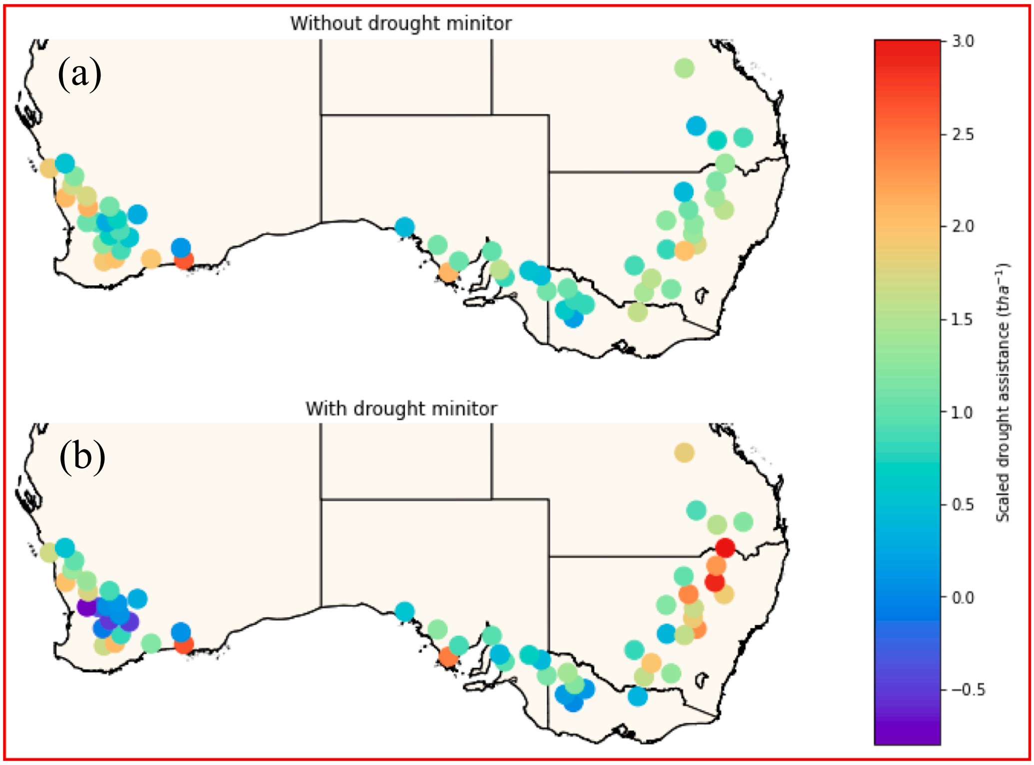

Fig. 6. Mapping the drought assistance across 59 production sites of the Australian wheatbelt without (a) and with the ADM dataset (b). The September CDI was used to inform drought assistance in (b). Details of the 59 production sites can be found in (An-Vo et al., 2018; Mushtaq et al., 2017).

accuracy, completeness, consistency, and timeliness (Ali and Dahlhaus, 2022). One way to ensure quality data is to harness relevant expertise in ontology development and identify innovative solutions that support quality data annotation (Arnaud et al., 2020). Furthermore, it is important to assess the quality of the data information and provide some inspiration for the subsequent research on data mining, as well as for the dataset optimisation for practical applications (Li and Chao, 2022).

Making agricultural research data FAIR requires a concerted effort from all stakeholders involved in agricultural research. Our interdisciplinary and systems approach in hosting FAIR data in a remote cloud computing environment enabled stakeholders to work together as observed by Knapen et al. (2023). By following the FAIR guidelines, we developed and implemented a practical method to make the ADM dataset FAIR, demonstrating that agricultural research data can be made FAIR, which will contribute to the advancement of agricultural research and ultimately benefit society. We acknowledge that there are other infrastructure implementation approaches that can be taken to make Agricultural datasets FAIR, and these will vary based on access to resources, supporting infrastructure and expertise. It may be possible to achieve FAIR with alternative and rudimentary client–server architecture on a smaller scale. It is important that researchers consider appropriate entry level options to achieve FAIR and approach the undertaking as a journey that scales over time.

We developed a standardised metadata record and readme file and their publications in the Research Data Australia. We developed the necessary resources allowing the transient ADM dataset transferred seamlessly among data servers in our developed hybrid cloud infrastructure minimising cost while ensuring a stable service for accessibility. This provides a sustainable solution in making agricultural research data FAIR. Though making agricultural research data FAIR comes at a cost, the benefits of FAIR data sharing does offset the cost. By harnessing relevant expertise, using appropriate agricultural practices, and utilizing big data analysis, the quality of agricultural datasets can be improved, being more convenient in the FAIR process.

Data availability

The Australian Drought Monitor dataset is updated monthly and available in the AgReFed platform which can be accessed through the ARDC Nectar Research Cloud Jupyter Notebook Service[18] using the Australian Access Federation (AAF) credentials and selecting the AgReFed Python environment. We can then download a copy of the dataset by following steps here https://github.com/AgReFed/drought-status-notebooks/wiki/Drought-status-workshop#drought-monitor-dataset.

Code availability

Codes and data to build applications in the Drought Monitor Work package can be obtained through a GitHub repository https://github.com/AgReFed/drought-status-notebooks.

CRediT authorship contribution statement

Francis Gacenga: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Duc-Anh An-Vo:** Writing – review & editing, Visualization, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Jillian McCulloch:** Writing – review & editing, Writing – original draft, Validation, Investigation, Data curation. **Richard Young:** Visualization, Software, Investigation, Data curation. **David Cobon:** Writing – review & editing, Project administration, Funding acquisition, Data curation.

[18] https://jupyterhub.rc.nectar.org.au/hub/spawn.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was funded by Queensland Department of Agriculture and Fisheries – Drought and Climate Adaptation Program (DCAP), Northern Australian Climate Program (NACP) [1] and the Australian Research Data Commons (ARDC) – AgReFed Platform Project. This research benefits from and acknowledges the use of the University of Southern Queensland *Fawkes* HPC cluster.

References

Agricultural Research Federation, A., 2023a. AgReFed Platform.
Agricultural Research Federation, A., 2023b. FAIR Assessment Tool.
Ali, B., Dahlhaus, P., 2022. The Role of FAIR Data towards Sustainable Agricultural Performance: A Systematic Literature Review. Agriculture (basel) 12, 309.
An-Vo, D.-A., Mushtaq, S., Zheng, B., Christopher, J.T., Chapman, S.C., Chenu, K., 2018. Direct and indirect costs of frost in the Australian Wheatbelt. Ecol. Econ. 150, 122–136.
Arnaud, E., Laporte, M.-A., Kim, S., Aubert, C., Leonelli, S., Miro, B., Cooper, L., Jaiswal, P., Kruseman, G., Shrestha, R., Buttigieg, P.L., Mungall, C.J., Pietragalla, J., Agbona, A., Muliro, J., Detras, J., Hualla, V., Rathore, A., Das, R.R., Dieng, I., Bauchet, G., Menda, N., Pommier, C., Shaw, F., Lyon, D., Mwanzia, L., Juarez, H., Bonaiuti, E., Chiputwa, B., Obileye, O., Auzoux, S., Yeumo, E.D., Mueller, L.A., Silverstein, K., Lafargue, A., Antezana, E., Devare, M., King, B., 2020. The ontologies community of practice: a CGIAR initiative for big data in agrifood systems. Gene Expr. Patterns 1, 100105.
Australian Research Data Commons, A., 2013. Registry Interchange Format - Collections and Services RIF-CS v1.6.5.
Australian Research Data Commons, A., 2020. ARDC Metadata Guide. Zenodo.
Australian Research Data Commons, A., 2022. CARE Principles.
Bahlo, C., Dahlhaus, P., 2021. Livestock data – Is it there and is it FAIR? A systematic review of livestock farming datasets in Australia. Comput. Electron. Agric. 188, 106365.
Carroll, S.R., Garba, I., Figueroa-Rodríguez, O.L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., 2020. The CARE principles for indigenous data governance. Data Sci. J. 19.
Cerba, O., Charvat, K., Ježek, J., Kafka, Š., 2011. Enhancing the Efficiency of ICT by Spatial Data Interoperability, pp. 1182-1197.
CF Community, 2023. CF Metadata Conventions.
Coble, K.H., Mishra, A.K., Ferrell, S., Griffin, T., 2018. Big Data in Agriculture: A Challenge for the Future. Appl. Econ. Perspect. Policy 40, 79–96.
Cobon, D., Gacenga, F., An-Vo, D.-A., Pudmenzky, C., Nguyen-Huy, T., Stone, R., Guillory, L., Jackson, J., Svoboda, M., Swigart, J., Meat, Australia, L., 2022. Australian Drought Monitor, Toowoomba.
Cocco, M., Bailo, D., Jeffery, K.G., Paciello, R., Vinciarelli, V., Freda, C., 2020. Sustainable FAIR Data management is challenging for RIs and it is challenging to solid Earth scientists, p. 18570.
Creative Commons, 2023. Attribution-ShareAlike 4.0 International (CC BY-SA 4.0).
Devare, M., Arnaud, E., Antezana, E., King, B., 2023. Governing Agricultural Data: Challenges and Recommendations. In: Williamson, H.F., Leonelli, S. (Eds.), Towards Responsible Plant Data Linkage: Data Challenges for Agricultural Research and Development. Springer International Publishing, Cham, pp. 201–222.
Guillory, L., Pudmenzky, C., Nguyen-Huy, T., Cobon, D., Stone, R., 2023. A drought monitor for Australia. Environ. Model. Softw. 170, 105852.
Hackett, R.A., Belitz, M.W., Gilbert, E.E., Monfils, A.K., 2019. A data management workflow of biodiversity data from the field to data users. Applications in plant sciences 7, e11310-n/a.
Dublin Core Metadata Initiative, D., 2002. The Dublin Core Metadata Element Set. Ringgold, Inc, p. 238.
Katabalwa, A.S., Bates, J., Abbott, P., 2021. Potential opportunities and risks of sharing agricultural research data in Tanzania. IASSIST quarterly 45.
Kharel, T.P., Ashworth, A.J., Owens, P.R., Buser, M., 2020. Spatially and temporally disparate data in systems agriculture: Issues and prospective solutions. Agron. J. 112, 4498–4510.
Knapen, R., Lokers, R., Janssen, S., 2023. Evaluating the D4Science virtual research environment platform for agro-climatic research. Agr. Syst. 210, 103706.
Kumar, P., Hendriks, T., Panoutsopoulos, H., Brewster, C., 2024. Investigating FAIR data principles compliance in horizon 2020 funded Agri-food and rural development multi-actor projects. Agr. Syst. 214, 103822.
Li, Y., Chao, X., 2022. Distance-entropy: an effective indicator for selecting informative data. Front. Plant Sci. 12, 818895.
Mushtaq, S., An-Vo, D.-A., Christopher, M., Zheng, B., Chenu, K., Chapman, S.C., Christopher, J.T., Stone, R.C., Frederiks, T.M., Alam, G.M.M., 2017. Economic assessment of wheat breeding options for potential improved levels of post head-emergence frost tolerance. Field Crop Res 213, 75–88.
Petrosyan, L., Aleixandre-Benavent, R., Peset, F., Valderrama-Zurián, J.C., Ferrer-Sapena, A., Sixto-Costoya, A., 2023. FAIR degree assessment in agriculture datasets using the F-UJI tool. Eco. Inform. 76, 102126.
Ramirez-Villegas, J., Salazar, M., Jarvis, A., Navarro-Racines, C.E., 2012. A way forward on adaptation to climate change in Colombian agriculture: perspectives towards 2050. Clim. Change 115, 611–628.
Research Data Management Service Group, C.U., 2023. Guide to writing “readme” style metadata.
Schweitzer, M., Levett, K., Russell, K., White, A., Unsworth, K., 2021. FAIR-Data-Assessment-Tool: Release v1.0 (v1.0). Zenodo.
Senft, M., Stahl, U., Svoboda, N., 2022. Research data management in agricultural sciences in Germany: We are not yet where we want to be. PLoS One 17, e0274677–e.
Sine, M., Theo-Paul, H., Emeric, E., 2015. API - AGRO: An Open Data and Open API platform to promote interoperability standards for Farm Services and Ag Web Applications. J. Agric. Informat. 6.
Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., Rippey, B., Tinker, R., Palecki, M., Stooksbury, D., Miskus, D., Stephens, S., 2002. THE DROUGHT MONITOR. Bull. Am. Meteorol. Soc. 83, 1181–1190.
Top, J., Janssen, S., Boogaard, H., Knapen, R., Şimşek-Şenel, G., 2022. Cultivating FAIR principles for agri-food data. Comput. Electron. Agric. 196, 106909.
Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Santos, L.B.d.S., Bourne, P.E., 2019. The FAIR Guiding Principles for scientific data management and stewardship (vol 15, 160018, 2016). Scientific data 6.
Wong, M., Levett, K., Lee, A., Box, P., Simons, B., David, R., MacLeod, A., Taylor, N., Schneider, D., Thompson, H., 2022. Development and Governance of FAIR Thresholds for a Data Federation. Data science journal 21.