Computers and Electronics in Agriculture 230 (2025) 109855

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Implementing FAIR principles in data management systems: A multi-case study in precision farming

Ngakan Nyoman Kutha Krisnawijaya [a], Bedir Tekinerdogan [a,*], Cagatay Catal [b], Rik van der Tol [c], Yeni Herdiyeni [d]

[a] Information Technology Group, Wageningen University & Research, The Netherlands
[b] Department of Computer Science and Engineering, Qatar University, Qatar
[c] Agricultural Biosystems Engineering Group, Wageningen University & Research, The Netherlands
[d] Department of Computer Science, IPB University, Indonesia

ARTICLE INFO

Keywords:
FAIR principles
Data management system
Data ecosystem
Interdisciplinary research

ABSTRACT

Over the past few years, the agricultural sector has witnessed a significant transformation, increasingly adopting a data-driven approach. Adopting an advanced data management system becomes imperative to effectively manage and govern the vast amounts of data generated in this context. Within the realm of data management, the FAIR principles provide valuable guidance. These principles aim to make data Findable, Accessible, Interoperable, and Reusable (FAIR), ensuring that data can be effectively managed, shared, and reused across different domains and disciplines. However, implementing the FAIR principles is not a straightforward task, requiring careful consideration of various factors such as data organization, metadata standards, interoperability protocols, and accessibility mechanisms. To address these challenges, this paper focuses on presenting a systematic approach for implementing the FAIR data principles within the data management system of an interdisciplinary agricultural project. In this paper, each FAIR principle is analyzed in detail, delving into the specific requirements and considerations for achieving them in the context of agricultural data. The current implementation approaches for each principle are identified and synthesized, taking into account both common practices and variant approaches that may be applicable to different scenarios. To provide practical insights, a multi-case study approach is applied to an interdisciplinary project involving dairy and fish farming. This research underscores the importance of metadata, secure data access protocols, semantic interoperability, and comprehensive documentation for implementing FAIR principles in agricultural data management systems, offering valuable insights applicable to dairy and fish farming domains.

1. Introduction

Agricultural sectors have been immensely shifted to the data-driven domain in the last decade (Mehrabi et al., 2020; Tekinerdogan, 2022). The volume of digital agricultural data grows exponentially, dominantly generated by digital technologies used in farming activities, business and decision-making processes (Tantalaki et al., 2019). On top of that, the advancements in managing, processing, and analyzing the generated data to support agricultural decision-making transform agricultural activities into more data-driven (Saiz-Rubio and Rovira-Más, 2020). The agricultural data consists of information regarding the whole farm's conditions and actions, which come from various and multiple tools and sources (Roussaki et al., 2023; Tantalaki et al., 2019). Most advanced smart farming tools, software, and applications deployed to support farm activities use their proprietary formats and standards for the generated data (Bahlo et al., 2019). In addition, agricultural data can be performed in different forms, such as survey data, field records, lab reports, or weather data (Kharel et al., 2020). As a result, it complicates data exchange and integration among processing systems or stakeholders. Furthermore, stakeholders' different needs and interests resulted in group dynamic challenges regarding the use of agricultural data (Himanen et al., 2019).

To mitigate those issues, proper data management applications are needed, not only to provide features for storing and handling data but also to manage and govern the data properly (Rix et al., 2021; Oliveira et al., 2019), especially for managing and sharing heterogeneous and

* Corresponding author.
E-mail address: bedir.tekinerdogan@wur.nl (B. Tekinerdogan).

https://doi.org/10.1016/j.compag.2024.109855
Received 30 May 2024; Received in revised form 23 November 2024; Accepted 17 December 2024
Available online 26 December 2024

large-scale data is a massive challenge for effective use by users (Kamilaris et al., 2017). Specifically, a data management system with proper rules to manage and store data is needed to address the problem (Arefolov et al., 2021). FAIR data principles provide guidance for developing data management systems. FAIR principles help to manage and prepare the data for sharing and reusing among stakeholders across different domains, disciplines or organizations by making it Findable, Accessible, Interoperable, and Reusable (FAIR) (Top et al., 2022).

Even though FAIR principles provide valuable guidance, implementing them is not straightforward (Jacobsen et al., 2020a; Wilkinson et al., 2018). It requires careful consideration of various factors, including data organization, metadata standards, interoperability protocols, accessibility mechanisms and stakeholders. To address these challenges, this paper focuses on presenting a systematic approach for implementing the FAIR data principles within the data management system of an interdisciplinary agricultural project. Furthermore, to provide practical insights, a multi-case study approach is applied to a multidisciplinary project involving dairy and fish farming. Smart Indonesian Agriculture (smart-in-ag)[1] is used as a multi-case study, which is a multidisciplinary and collaborative project between Wageningen University and Research in the Netherlands, IPB University in Indonesia, and other stakeholders to introduce and establish a smart farming technology for Indonesian dairy and fish farming. In this project, as an interdisciplinary consortium, there are different members' backgrounds to improve the quality and production of Indonesian dairy and fish products.

In this paper, each FAIR principle is analyzed in detail, delving into the specific requirements and considerations for achieving them in the context of agricultural data. In specific, the objectives of this study are as follows:

1. Developing a systematic approach for implementing FAIR principles for data management systems in an interdisciplinary agricultural project.
2. Evaluating this systematic approach using an actual multi-case study, dairy and fish farming in Indonesia.

The structure of the remainder of this paper is as follows: Section 2 explains the FAIR principles, particularly in the context of data management systems. The next section details the systematic approach for implementing these principles. Section 4 then delves into our multi-case study, providing practical insights and applications, while Section 5 discusses the findings. Finally, the last section concludes the paper with key takeaways and implications of our study.

2. Related work

With the emergence of data-driven agricultural activities, creating a data management system and following proper procedures have become essential (Jacobsen et al., 2020a; Thompson et al., 2020). A data management system is a solution to manage, process, and analyze the generated data, which can guide farmers and other stakeholders to make the best decisions that lead to innovative agricultural actions (Andrade et al., 2021). However, the main challenge in creating a proper data management system is to ensure the efficiency of knowledge transfer regarding several aspects of data management (i.e., organizational, procedural and technical aspects) to the stakeholders from different backgrounds and domains to maximize opportunities for good data handling and data reuse (Thompson et al., 2020). Therefore, it is crucial to implement an effective data management strategy by following the FAIR data principles since these principles will amplify data utility, especially for legacy data sets (Ali and Dahlhaus, 2022). Furthermore, applying FAIR principles enhances the value of data assets by making the data more discoverable and accessible for machines and humans (Wise et al., 2019).

The first principle of FAIR is to ensure data is easily found, which is a fundamental aspect of the remaining principles. The diversity of quality and precision of data sources may lead to various data schemas (Arefolov et al., 2021; Wise et al., 2019). Consequently, establishing a uniform approach for cataloging data sources is challenging. To address this, data sources should be equipped with suitable metadata to enable the creation of self-reporting data catalogues, providing details such as access methods, content, and other relevant information (Koers et al., 2020; Wise et al., 2019). The second aspect of FAIR principles is making data more accessible with clearly and adequately defined access rules. Three main components of accessibility are access protocol, access authorization and metadata longevity (Wise et al., 2019). The next element of FAIR principles is interoperability, which addresses the need to express data in a formal, accessible, sharable, and broadly applicable language (Wilkinson et al., 2016). Finally, the last principle of FAIR is the reusability of data, which allows data to be recycled for new user communities, new purposes or needs and new applications (Celebi et al., 2020).

Hence, the FAIR data principles can be seen as a guideline to enable digital data infrastructure. A multi-stakeholder community can benefit from implementing the FAIR data principle (FAIRification) in their data management system since it will support the data to become easily reused by machines and people (Borycz and Carroll, 2020). Several communities, such as the GO Fair organization (GO FAIR, 2022), Research Data Alliance (FAIR Data Maturity Model Working Group, 2020), and Jacobsen et al. (Jacobsen et al., 2020b), have published several common routes to implement the FAIR data principles and make the FAIRification process more robust and consistent. The generic workflows proposed in (GO FAIR, 2022) are 1) The purpose of FAIRification identification; 2) Data analysis; 3) Metadata analysis; 4) Semantic model for data and metadata; 5) Linked data and metadata; 6) FAIR data hosting; 7) FAIR data assessment.

Ref. (Singh et al., 2021) followed the FAIRification process to set a standard for creating image-based datasets and advanced research resources to be more effectively used. They also added metadata, vocabularies, and unique identifiers to the data to simplify data sharing and retrieval. These strategies aimed to integrate the datasets among stakeholders in the plant science and phenotyping community since they will reduce data redundancies and lead to cost savings.

Ref. (Dorich et al., 2020) established The Global N2O Database to host all datasets of Nitrous Oxide (N2O), one of the greenhouse gas emissions produced by agricultural activities. This database also became a public resource for researchers or others in the same domain of interest for analytical purposes. They applied the FAIR principles to their database to follow proper and consistent standards for data sharing since sampling N2O emissions is expensive and takes too much time to process.

Several recent studies have demonstrated the benefits of FAIRification on agricultural systems. In (Zhang et al., 2023), FAIR principles were implemented to help the dissemination of agro-geoinformation through a cyberinformatics web-based tool. Meanwhile, Ref. (Gacenga et al., 2024) developed a workflow based on FAIR principles along with CARE (Collective benefit, Authority to control, Responsibility and Ethics) principles for the Drought Monitor dataset in Australia for solving the difficulty of data sharing in agricultural research practices. According to (Kumar et al., 2024), it was found that to enhance the digitalization of agri-food and rural development sector practices, the FAIR principles adoption and data management practices needed to be standardized.

Despite the benefits offered by FAIR principles, a few published papers discussing FAIR implementation were found in the literature, specifically in the agricultural domain (Ali and Dahlhaus, 2022). Therefore, this study presents a systematic approach to implementing FAIR principles within the data management system of an interdisciplinary

[1] https://smart-in-ag.com/.

agricultural project. The current study proposes the common practices and variant approaches of each FAIR principle that may be applicable in agricultural data management. Furthermore, by studying real-world examples, the paper examines how to apply the FAIR principles effectively and investigates the potential benefits and challenges that arise in an agricultural context.

3. Research method

To achieve the goal of this study, the design science research (DSR) method was deployed. DSR is a problem-solving approach aimed at producing artifacts that enable the effectiveness and efficiency of information systems (Hevner et al., 2004). This method consists of several phases: problem definition, solution design, and validation (Giray and Catal, 2021). The first step is identifying several practical problem instances, which were analyzed, leading to the development of a problem statement based on theoretical concepts from existing literature. The next step is to design a conceptual solution by following a systematic approach, which involves domain analysis to derive and represent domain knowledge for the purpose of solution design (Giray and Catal, 2021). Domain analysis consists of domain scoping and domain modeling activities. Domain scoping aims to identify relevant knowledge sources to derive the key concepts of solution (van Geest et al., 2021), while domain modeling seeks to unify and represent the domain knowledge gathered from relevant sources through a feature model (Tekinerdogan and Öztürk, 2013). The last step of the method is validation, and this study utilizes two real case studies to evaluate the feature model that demonstrates the solution. The DSR method used in this study is illustrated in Fig. 1, and a detailed explanation is provided in the following subsections.

3.1. Problem definition

This study was motivated by various cases related to implementing FAIR principles in the agricultural domain, as seen below.

Problem Case 1: Limited framework regarding FAIR implementation in the agricultural domain.

Regarding implementing the FAIR principles in the agricultural data management system, a few publishing papers discussed FAIR implementation were found in the literature (Ali and Dahlhaus, 2022). Ref. (Top et al., 2022) presented the FAIRification of data management systems in agriculture and food science and explored every FAIR principle step using three cases. They found that three dominant factors should be considered before FAIRification and data can be effectively reused: 1) Advanced tools for data providers and users; 2) Involvement of community in developing vocabularies and tools; 3) An open-by-default policy is not the only rule for data sharing. Although it was found that the FAIR principles could raise the awareness of data sharing between researchers and professionals in the domain, it still needs a concrete investigation of several principles that are hard to implement.

Problem Case 2: Complexity of FAIRification process.

According to (Jacobsen et al., 2020a), all community or organization can define their FAIRification solutions based on their digital resources and purposes as long as they align with the points of the FAIR principles. However, the choice of freedom of FAIRification can also lead to a risk of conflicting solutions among stakeholders in the domain (Jacobsen et al., 2020a). The literature shows that FAIRification is a complex step in harmonizing the data for publishing and reusing by other people and machines. Some steps need to be carefully considered in FAIRification. It was also found that the purposes of FAIRification of data management, proper data infrastructure, and community are crucial in implementing the FAIR principles.

3.2. The implementation of FAIR principles

3.2.1. Domain scoping

This section discusses our systematic approach to implementing the FAIR principles to address the problem cases provided in the previous section. At this stage, we investigated the domain of interest and applied approaches to achieve the standards of each principle. The first step of the domain scoping stage is defining the domain of interest, which, for the current study, is the implementation of FAIR principles in agricultural systems. As a result, the phrase "FAIR data principles" or "FAIR principles" will be combined with "data management", "data ecosystem", and "farming", "agriculture" as a search strategy to obtain relevant papers in the Scopus and ScienceDirect databases. After that, exclusion criteria were applied to the candidate papers to retain only those relevant to the study. The papers without full text available, those not written in English, duplicates, those not involving the agricultural domain, and those not discussing the implementation of FAIR principles were excluded from the list of candidate papers. The information was extracted from the papers, including the title, publication year, and information regarding FAIRification, which was recorded in an Excel sheet for further analysis. As the automation search with the search query resulted in a very limited number of papers, we also conducted a manual search, including grey literature, such as websites, reports, and books, to ensure sufficient coverage. The websites included in the study are GO-FAIR,[2] Australian Research Data Commons (ARDC),[3] General Data Protection Regulation (GDPR)[4] and Research Data Alliance (RDA).[5] The relevant literature was used to develop a domain model that represents the knowledge of the domain. Table 1 shows the list of literature, including papers, reports, and books.

3.2.2. Domain modeling

A feature model is employed to visualize the common and familiar approaches of FAIRification identified through domain scoping analysis. Feature modeling is a method for capturing domain knowledge in a structured way that allows for reuse (Tekinerdogan and Öztürk, 2013). The model is depicted as a tree structure, where the root represents the concept of FAIRification, and its nodes represent the common approaches to applying FAIR principles. Three types of dependencies, mandatory, optional, and alternative, illustrate the relationships among the nodes in the tree (Krisnawijaya et al., 2024).

Fig. 2 illustrates the feature diagram of the FAIR principles. As mentioned in the previous section, the FAIRification process involves complexities in ensuring that data is findable, accessible, interoperable, and reusable (Garcia et al., 2019; Lin et al., 2022). The FAIRification process includes several aspects that must be implemented in the dataset. It is essential to understand each aspect to develop concrete steps for effective implementation. In this study, we identify and synthesize the current implementation approaches from the literature for each principle, considering both common practices and alternative methods that may apply to various scenarios, which will be discussed in the following subsections.

3.2.2.1. Findable. In the agricultural domain, various data resources exist to provide data to data consumers. Several options are available to make data findable, from old-fashioned approaches to the most advanced ones. Traditionally, data consumers use their networks to discover (un)published datasets from colleagues, farmers, agricultural organizations, and even librarians or other data producers or managers (Gregory et al., 2018). Another way is by searching the literature, such as conference papers, scientific articles, or work reports, which are

[2] go-fair.org.

[3] ardc.edu.au/resource/fair-data-self-assessment-tool.

[4] https://gdpr.eu/.

[5] rd-alliance.org.

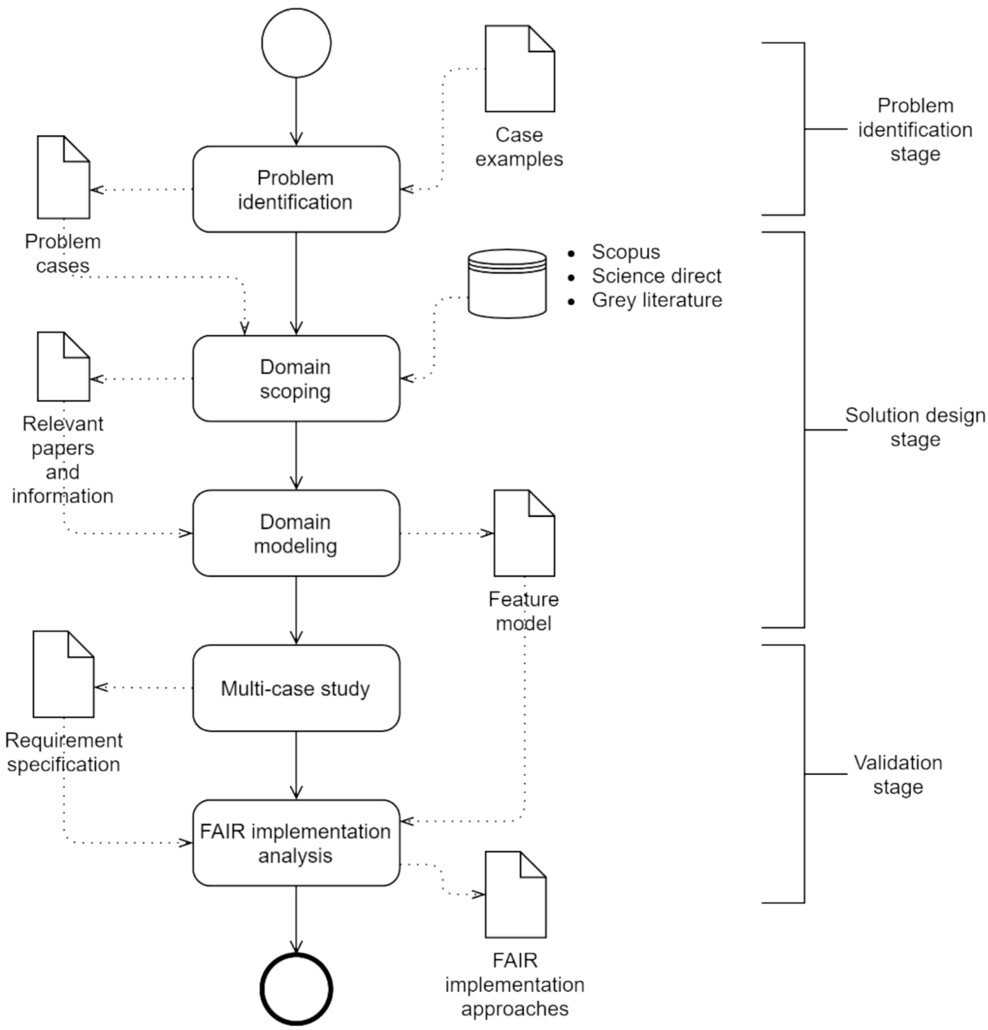

Fig. 1. The DSR method used in this study.
Adapted from Giray and Catal (Giray and Catal, 2021).

Table 1
List of the literature to obtain the FAIR principles adoption (FAIRification).

Samourkasidis, A., & Athanasiadis, I. N. (2020). A semantic approach for timeseries data fusion. Computers and Electronics in Agriculture, 169. https://doi.org/10.1016/j.compag.2019.105171 (journal article)
Dey and Shekhawat (2021). Blockchain for sustainable e-agriculture: Literature review, architecture for data management, and implications. Journal of Cleaner Production, 316. https://doi.org/10.1016/j.jclepro.2021.128254 (journal article)
Top et al. (2022). Cultivating FAIR principles for agri-food data. Computers and Electronics in Agriculture, 196. https://doi.org/10.1016/j.compag.2022.106909 (journal article)
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. Ecological Informatics, 76. https://doi.org/10.1016/j.ecoinf.2023.102126 (journal article)
Specka, X., Gärtner, P., Hoffmann, C., Svoboda, N., Stecker, M., Einspanier, U., Senkler, K., Zoarder, M. A. M., & Heinrich, U. (2019). The BonaRes metadata schema for geospatial soil-agricultural research data – Merging INSPIRE and DataCite metadata schemes. Computers & Geosciences, 132, 33–41. https://doi.org/10.1016/j.cageo.2019.07.005 (journal article)
Ali, B., & Dahlhaus, P. (2022). The Role of FAIR Data towards Sustainable Agricultural Performance: A Systematic Literature Review. Agriculture, 2(12). https://doi.org/10.3390/agriculture12020309 (journal article)
Bahim et al. (2019). Results of an Analysis of Existing FAIR Assessment Tools. DOI https://doi.org/10.15497/rda00035 (report)
FAIR Data Maturity Model Working Group. (2020). FAIR Data Maturity Model. Specification and Guidelines. DOI https://doi.org/10.15497/rda00050 (report)
MacLeod et al. (2020). The Agricultural Research Federation (AgReFed) Technical and Information Policy Suite. Version 1.0. Endorsed by the AgReFed Council on 13 May 2020. DOI https://doi.org/10.5281/zenodo.3993784 (report)
Science Europe. (2021). Practical Guide to the International Alignment of Research Data Management. https://www.scienceeurope.org/media/4brkxxe5/se_rdm_practical_guide_extended_final.pdf (report)
Food and Agriculture Organization of the United Nations (FAO). (2021). AGROVOC – Semantic data interoperability on food and agriculture. https://doi.org/10.4060/cb2838en (Book)

usually based on data and see if they contain the supplementary dataset. Besides searching it manually through a network or literature, data could also be found by using a web search engine. Finding data through a web search engine could be done either by searching it in the domain-specific portal or repository (Gregory et al., 2018). There are several types of repositories, such as local servers, private cloud, public cloud, and hybrid cloud. A local server repository is data storage that is developed by the data owner to store the data. A private cloud is a repository dedicated exclusively to one organization, while a public cloud is a repository made available to the public and shared between

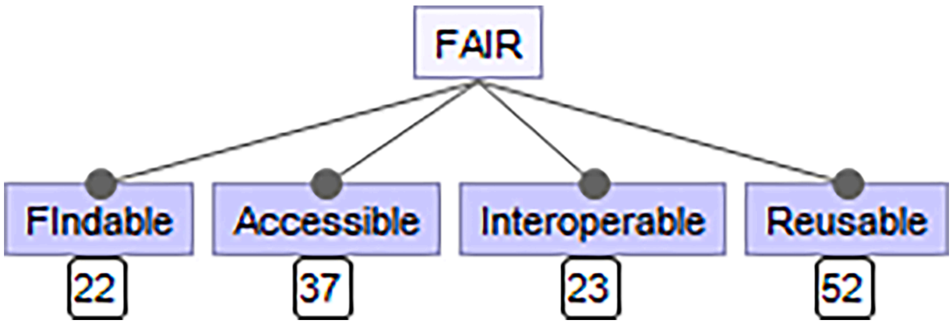

Fig. 2. Feature diagram of FAIR data principles.

organizations hosted by cloud computing services vendors. The last repository type is the hybrid cloud, which combines private and public clouds (Goyal, 2014).

The more advanced approach is finding data using a persistent identifier. The persistent identifier (PID) is critical in data management systems to solve the problem of long-lasting finding of cited resources. PID can be divided into two categories: people and objects. Open Researcher and Contributor ID (ORCID) is the PID for people and Digital Object Identifier (DOI) for objects (Bahim et al., 2019; Hauschke et al., 2021). Thus, PID is strongly recommended as an identifier for data or metadata since the PIDs can permanently connect with a set of (meta) data, which contains information about the digital objects, to allow people and machines to track and find a reliable way to the digital objects.

The last method to find the data is using metadata. Metadata is a set of data that provides information about other data (Allemang and Teegarden, 2017). Metadata can vary in richness, which is how much of the data is described and captured in the metadata (MacLeod et al., 2020). There are three metadata types: basic, specialized, and rich. Basic metadata contains basic information about the data, such as authorship, year, title, and short description of data. Specialized metadata contains more complex information than the basic one, such as a persistent identifier, license, and access rights. Meanwhile, rich metadata contains complete and concrete information and follows the standardized format and vocabulary to allow the machine to find and read the included information within the metadata (Australian Research Data Commons (ARDC) (2022); Thompson et al., 2020). Fig. 3 depicts the family features of findability.

3.2.2.2. Accessible. Transforming raw data into FAIR data is not about

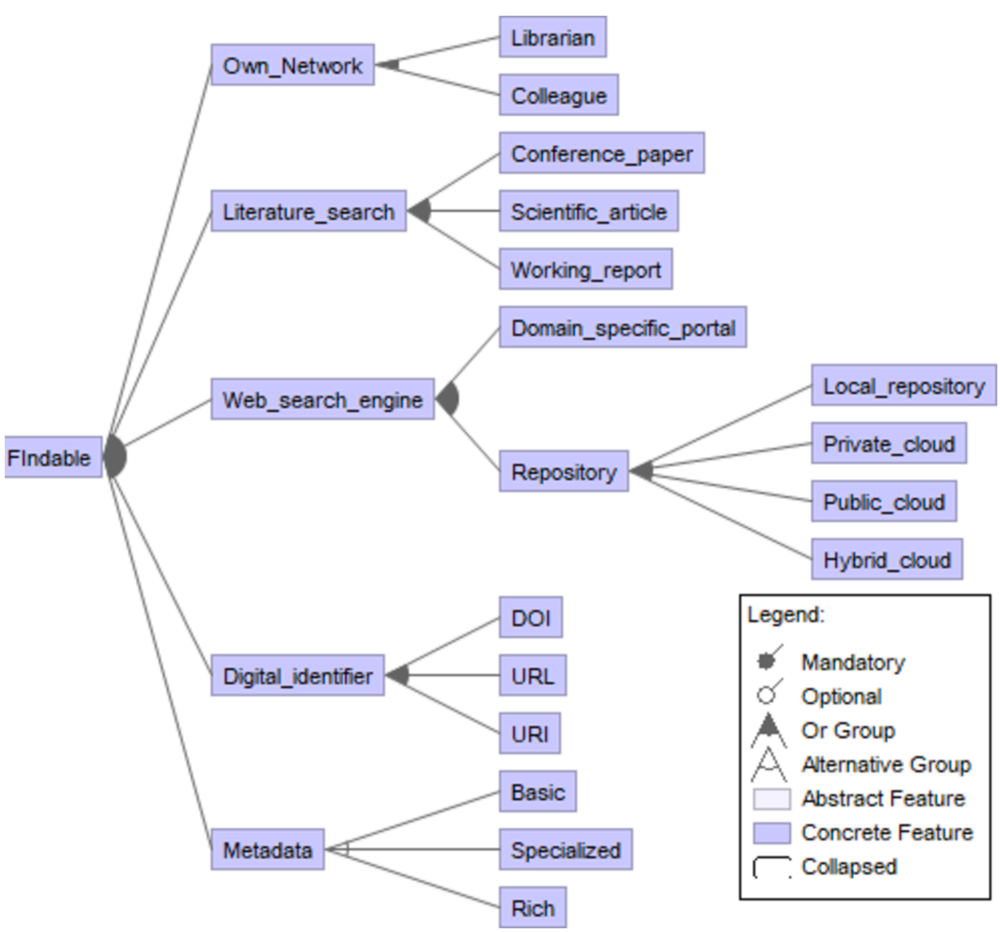

Fig. 3. Feature diagram of findability.

making data fully open (Wong et al., 2022). Accessibility, as the second principle of FAIR, guides people or machines to access the data with correct rights and proper ways. In this principle, the data owners can state their data status and choose the protocol or policy to access their data. Accessibility can be divided into two components: types of access and access properties (Wise et al., 2019). There are several accessibility features, which can be seen in Fig. 4.

It was found that the dataset can be accessed publicly from repositories, and data consumers can freely download the dataset. For instance, the weather data provided by government websites can be accessed by the public, or datasets can also be downloaded from public resources, such as GIT or GitHub (Bahim et al., 2019). Furthermore, the data owners or managers have to consider their dataset's sensitivity level before publishing publicly. The dataset has to align with the ethics or policies related to sensitive data. For more sensitive data, registered-based access (Gregory et al., 2018) could be opted to ensure proper user authorization, authentication, and control access. For everyone who needs the datasets, registration to the systems is required. By doing so, the data managers can control user authentication and authorization and set every system member an appropriate right to access the dataset. This approach is more secure in protecting sensitive data within datasets.

Unspecified conditional access can also be an option for accessing data, and the data customers need to ask data custodians to retrieve the preferred dataset directly (Australian Research Data Commons (ARDC) (2022)). Meanwhile, embargoed access could be chosen for the data owners who want to make the data available at a predetermined time (Australian Research Data Commons (ARDC) (2022)). The following

Fig. 4. Feature diagram of accessibility.

approach to accessing data is using a query. Data users can filter the dataset with a query to retrieve only the needed information. Data users can apply queries in several systems, such as local, domain-based, or general indexing systems (e.g., Google, Bing, Yahoo) (Gregory et al., 2018; Hauschke et al., 2021).

Other than types of access, data owners or customers have to consider access properties when accessing the preferred data. The first access property is access protocols, which are essential for securing communication between two or more machines to access the data. Open and secure protocols such as API, HTTPS, and FTP can be chosen (Hauschke et al., 2021). Secondly, the access devices that are used to access the data also need to be determined. Sometimes, data can be only accessed on personal computers and laptops, or it is also accessible through server systems or mobile devices, such as tablets or smartphones. Another critical issue that needs to be addressed is how long data will be kept and accessible. It could be long-term (e.g., more than ten years) and/or still available in the organization, although the person who collected the data has left it. Thus, data preservation is essential to prevent the potential risk of data loss (Science Europe, 2021).

3.2.2.3. Interoperable. Data exchange is crucial for research and innovation in agriculture (Food and Agriculture Organization of the United Nations (FAO). (2021)). For instance, it could be a competitive factor since data sharing enables a more knowledge-based production and demand-driven to meet different needs of the consumers or stakeholders (Food and Agriculture Organization of the United Nations (FAO). (2021)). Thus, enhancing the data exchange practices in the agricultural sector is necessary. In order to improve data sharing, current practices need to be analyzed to harmonize it subsequently. Fig. 5 depicts the feature diagram of interoperability. To achieve organizational interoperability, it is necessary to have effective interoperability from the lowest to highest layer, which consists of technical, syntactic, and semantic interoperability (Khatoon and Ahmed, 2022).

The first level is technical interoperability, which focuses on basic exchange capabilities between systems (Lehne et al., 2019). This interoperability is related to the protocol between hardware, software, or networks that allows communication among the machines (Khatoon and Ahmed, 2022). The second layer is syntactic interoperability, which specifies the format and structure of data (Lehne et al., 2019). As regards data format, it consists of non-machine-readable and machine-readable data. For non-machine-readable data, interoperability can be implemented manually by integrating data in the proprietary format or consulting it with the data owners. However, recent technologies, such as big data systems, cloud computing, and machine learning, can change and improve data exchange in the agricultural domain. In recent years, these technologies have been the primary data consumers to increase the quality of the decision-making process; therefore, the digital data must be in machine-understandable format (Food and Agriculture Organization of the United Nations (FAO). (2021)). A standard language such as RDF, XML or JSON can be used to express the metadata and dataset. The data structure could be divided into structured, semi-structured, and unstructured data (Nyoman Kutha Krisnawijaya et al., 2022).

The next dimension of interoperability is semantic interoperability, which focuses on the mutual apprehension between parties that exchange data (Khatoon and Ahmed, 2022). Finding relevant data and information can be challenging as the data is fragmented or scattered in digital storage. Therefore, information from data or metadata should be delivered in a formal, standardized, sharable and applicable language (Wise et al., 2019). Metadata should contain elements that give clear information regarding the dataset and follow community guidelines

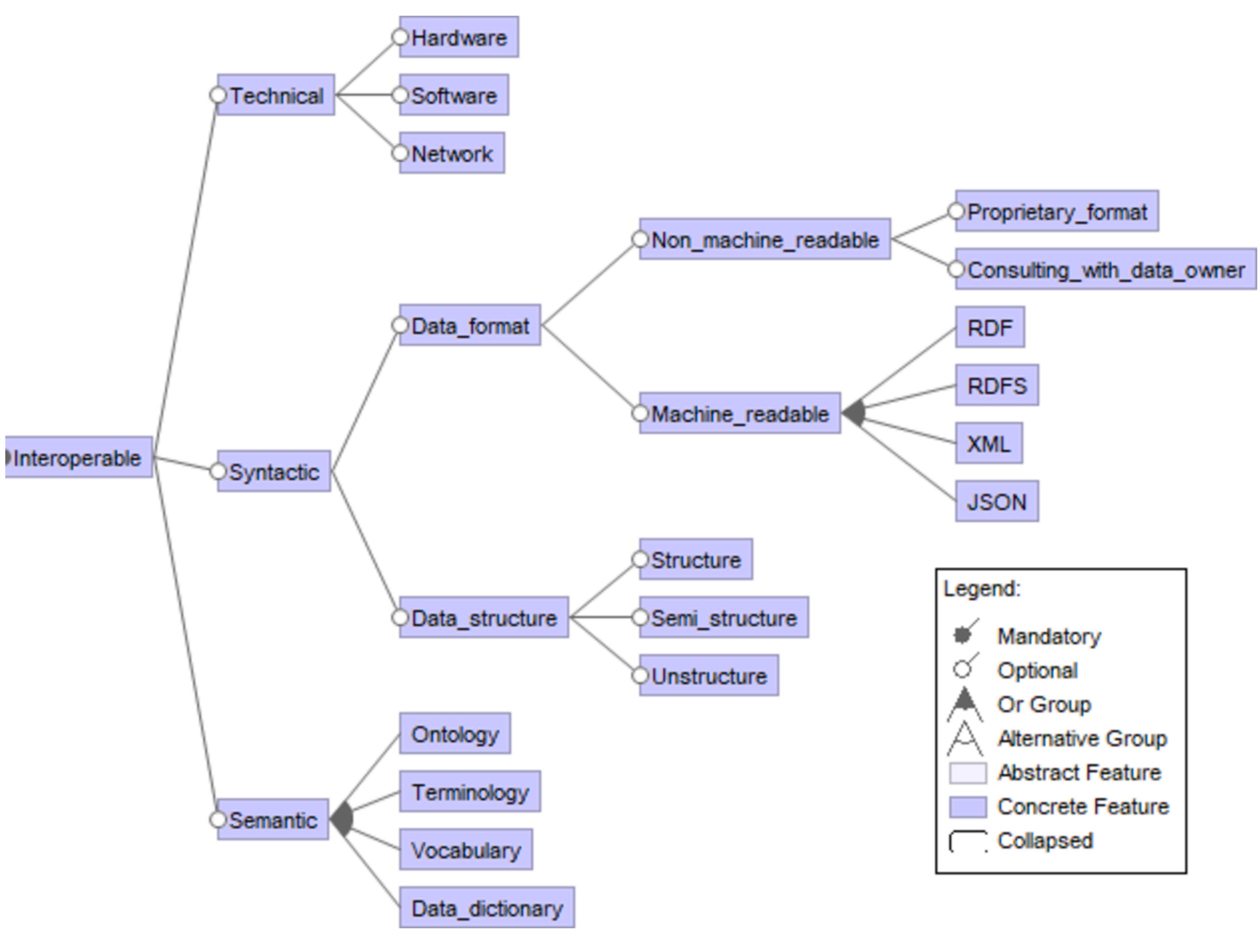

Fig. 5. Feature diagram of interoperability.

with well-defined vocabularies, such as AGROVOC (Food and Agriculture Organization of the United Nations (FAO). (2021)). In addition, the metadata elements must not be ambiguous to support the interpretation of the dataset. Thus, metadata should have an explicit ontology, terminology, vocabulary or data dictionary (Allemang and Teegarden, 2017; Hauschke et al., 2021).

3.2.2.4. Reusable. In this principle, other parties can reuse the dataset by following data usage conditions and restrictions defined by data owners. The data should be easy to understand and citable to maintain the value of data. Fig. 6 shows the reusability feature diagram. This research divides reusability into two mandatory sub-features: data ownership and privacy and security. Data ownership is necessary to make data reusable since data consumers can understand the data usage and how to cite the dataset. The data ownership properties, such as license (e.g., open or commercial), open digital rights language, and intellectual property rights, should be determined (Allemang and Teegarden, 2017). It is necessary to understand data ownership before reusing others' datasets or making the data reusable since it explains the data's owner(s), policies and procedures, and regulations that influence dataset usage.

Regarding privacy and security, data managers/producers have to check datasets before sharing or storing them with other parties to determine whether they contain sensitive and personal data, ethical considerations, data privacy, and security laws. The data often becomes sensitive in agriculture since it is acquired from the field, animal, and/or the farmer. Such data must be treated special before being shared with other parties (Allemang and Teegarden, 2017). For instance, the data owner can restrict their dataset, determine what kind of variables others can access, and determine which ones cannot or who can have full or partial access to the dataset. These data-sharing policies should be clearly mentioned in the metadata so that if other systems or humans

Fig. 6. Feature diagram of reusability.

want to reuse the data, they can easily understand the procedures. In addition, it is necessary to ensure that the data anonymity and pseudonymity are in the right place (Allemang and Teegarden, 2017). Anonymous data are no longer regarded as personal data, while pseudonymous data has been de-identified from the data's subject but can be re-identified as needed (Science Europe, 2021).

Besides data ownership, privacy, and security, it is also essential to consider the preferred data format to ensure the reusability of the dataset. For instance, in image data, jpeg, png, and tiff are the preferred formats since all machines can read this format. Another necessary reusability feature is data documentation to accompany existing datasets and enable reusable data. In the data documentation, several pieces of information, such as policy, constraint, and rules of reusing the dataset, must be stated clearly and explicitly. For instance, providing support for data versioning is crucial in reusability. Data versioning properties, such as file naming, version number, and date, must be determined to distinguish between versions of the file. Additionally, the details of data volume are also essential to estimate how much data storage is required when reusing the data. They can be expressed in numbers of objects (files, rows, and columns) and/or in storage space required (bytes).

3.3. Validation using A multi-case study approach

To address the second question, the multi-case study was designed to assess the FAIRification in the smart-in-ag project. The main purpose of this section was to evaluate and validate the systematic approach presented in Section 3.2.2. To achieve this, the case study approach was applied, following the protocol proposed by (Runeson and Höst, 2009). The case study protocol includes five key procedures that need to be followed (Runeson and Höst, 2009):

1. Case study design.

The goal of the case study research was initially defined. As mentioned earlier, the case study was employed to evaluate and validate the proposed systematic approach for FAIRification.

2. Data collection preparation.

Two different case studies were chosen in the current research: Indonesian dairy and fish farming. Both studies involved researchers, experts, and data analysts with sufficient experience managing and using datasets for their business or research. Since the smart-in-ag project is collaborative research between Wageningen University and Research in the Netherlands and several universities in Indonesia, the selected respondents were from the Netherlands and Indonesia and were directly involved in the project as they represent the most relevant stakeholders. Furthermore, this study also involved several industrial experts in Indonesia who were involved in the smart-in-ag project to provide a more comprehensive insight into precision farming.

3. Evidence collection.

In this step, the survey instruments were initially designed to accommodate all approaches in Section 3.2.2 through an iterative discussion and process among the authors. The current study used an online questionnaire through the online tool Google Forms to obtain the opinions of dairy and fish farming experts regarding FAIRification approaches. The online questionnaire method was chosen since the respondents were in the Netherlands and Indonesia. Thus, the survey was distributed through participants' email addresses. Furthermore, the questionnaire introduced the research objective, accompanied by an explanation of FAIRification procedures and how the data obtained from the questionnaire would be used. The questionnaire consists of five sections. The first section includes a set of introductory questions to ask about respondents' background and experience in smart farming systems. The following sections asked and evaluated the main questions regarding FAIR principles and each approach to achieving FAIRification.

The questionnaire consists of close-ended questions, such as multiple-choice questions with one or multiple answers allowed, and rating scales. Furthermore, in some questions, this study used a combination of close-ended and open-ended questions by providing predefined options and an "Other" option, allowing respondents to add their own choices. This will enable participants to provide additional responses that may not be covered by the given options. Before distributing the questionnaire to the targeted participants, a pilot test was conducted to gain the preferred level of the questionnaire's content.

As regards ethical considerations, personal data was not requested to maintain the confidentiality and anonymity of the respondents. It was explained in the general description section of the survey that participants would not be asked to provide any information related to their name, company name, or address anywhere in the questionnaire. The questionnaire only asked about participants' backgrounds (i.e., dairy or fish farming) and their experience with smart farming systems, as well as their opinions regarding the implementation approaches of FAIR principles in their domains.

4. Data analysis.

After collecting data, the feedback from the survey was organized, analyzed, and interpreted by looking at the frequency of the answers to determine mandatory and optional approaches based on experts' opinions. The approaches with high-frequency responses indicate that many experts suggest they are important approaches; hence, they were selected as mandatory approaches to achieve FAIRification. Meanwhile, the approaches that the experts did not choose were chosen as optional.

5. Reporting.

The final step is reporting the data analysis results. After identifying important approaches based on survey results, the next step is presenting the selected approaches in the table. Based on experts' opinions, the table summarizes the mandatory approaches to achieve FAIRification in dairy and fish farming.

The following sections discuss the survey results using feature diagrams to depict the mandatory approaches from experts' opinions for the FAIRification process in a certain domain.

4. Results

4.1. Participants characteristics

The participants in this study are the stakeholders of the smart-in-ag project, consisting of project members and industrial experts involved in this project. The survey was conducted over one month, from September to October 2023, with a total of 18 people participating in the survey. However, three of the participants had limited experience in the dairy or fish farming domains, and one was unsure about their expertise. Therefore, 14 questionnaires obtained from the survey were used for further analysis. Out of the 14 respondents, five are dairy farming experts, eight are fish farming experts, and one is an expert in both fields. In dairy farming, most of the experts have up to two years of experience, while in fisheries, the respondents mostly have from three to ten years of experience, with expertise in both theoretical and practical aspects. The details of the respondents' background and experience are shown in Table 2.

Based on Table 2, it can be seen that this study included respondents who represent both domains quite equally. They are the key stakeholders involved in the smart-in-ag project, so they have sufficient knowledge of the project and its requirements. Furthermore, by excluding those with limited experience in dairy or fish farming, this

Table 2
Background and experience of respondents.

Domain	Experience		Time of Experience		
			Up to 2 Years	3–10 Years	Total
Dairy	Experience	Theoretical Experience	3	0	3
		Practical Experience	0	0	0
		Both Experience	2	0	2
	Total		5	0	5
Fishery	Experience	Theoretical Experience	1	2	3
		Practical Experience	1	0	1
		Both Experience	1	3	4
	Total		3	5	8
Both (Dairy and Fishery)	Experience	Theoretical Experience	0	0	0
		Practical Experience	0	1	1
		Both Experience	0	0	0
	Total		0	1	1
Total	Experience	Theoretical Experience	4	2	6
		Practical Experience	1	1	2
		Both Experience	3	3	6
	Total		8	6	14

study ensures that only respondents with substantial experience in both domains provide their opinions regarding the implementation approaches of FAIR principles in their respective fields. Their feedback or preferences are used to evaluate and validate the proposed systematic approach for FAIRification.

4.2. Case study 1- Indonesian dairy farming

The first study was conducted with experts from the dairy farming domain who are involved in the smart-in-ag project. About 84 % of the respondents agreed that data exchange amongst data producers/owners is vital for their research or business. The summary of mandatory approaches to achieve FAIRification in dairy farming can be seen in Table 3.

Findability. Findable is the first principle asked to the respondents in this study. Fig. 3 shows several approaches that can be used to find the preferred data for the data seekers. This study found that about 50 % of the respondents usually use a literature search, their network (e.g., colleagues or librarians), or metadata to find a relevant dataset for their research. Specifically, all participants stated that metadata is mandatory and crucial to get and explain the data, and they agreed that metadata should be included in data sources. The appropriate metadata, such as specialized metadata, is a minimum requirement to establish the FAIR dataset, and half of the respondents agreed that rich metadata, which can be self-reported to the machine by using appropriate metadata schema, is essential. As regards data storage, participants mentioned private cloud repositories as the preferred systems.

Accessibility. The survey results showed that the respondents had adequate knowledge related to data security to access the data. More than half (83 %) of participants chose registered-based access, with user authorization being preferred the most, with a 67 % response rate. Another way to access the data is using a query, and it was identified that the domain-based system (e.g., specific domain repositories) and general indexing system (e.g., Google, Yahoo) are the most popular query systems. Regarding access devices, most participants tend to access the data through their mobile devices (e.g., tablets and smartphones) or personal computers. Furthermore, they opted for HTTPS and FTP as the access protocol. By doing all the data protection scenarios, the respondents would like to have their dataset always available, although they have left the organization.

Table 3
FAIR implementation approaches in dairy farming.

FAIR Principles	Implementation Approaches	Mandatory Approaches
Findability	Own network	Colleagues Librarians
	Literature search	Scientific article Conference paper Working report
	Metadata	Specialized metadata Rich Metadata
Accessibility	Registered-based access	User authorization
	Query	Domain-based system General indexing system (e.g., Google, Yahoo)
	Access protocol	HTTPS FTP
	Access devices	Personal computer Mobile devices
	Data preservation	Always available
Interoperability	Technical interoperability	Software (e.g., API) Network (e.g., HTTPS)
	Syntactic interoperability	Structured data
	Semantic interoperability	Ontology Terminology Vocabulary
Reusability	Data documentation	File naming Version date Methods and tools used for collecting data Data object Recorded date Location of data collection Number of objects (e.g., files, rows, columns)
	Data ownership	Data citation Intellectual property rights
	Data privacy and security	Data backup scenario Ethics Sensitive data security Data anonymity Data pseudonymity
	Data provenance	Fully recorded in a text format Fully recorded in a machine-readable format

Interoperability. Data interoperability can allow data providers and seekers to exchange the data's information properly. The process is not easy to do as several factors have to be considered, such as data structure, data type, etc. In the current study, for the data exchange approach, about 67 % of the respondents selected to use software (e.g., API) and through the network (e.g., HTTPS, FTP) instead of using hardware devices such as external hard disk or USB to exchange the data. Almost all respondents preferred data elements with standardized descriptions, such as ontology, terminology, and vocabulary, and only one person opted for a data dictionary to make data interoperable. Furthermore, the survey participants mostly worked with structured data in proprietary or non-machine-readable data format.

Reusability. Data information is crucial to reusability in order to allow other people, groups, organizations, or machines to understand the data. Hence, in the data documentation, the detailed information such as file naming, version date of the file, methods and tools used for collecting data, data object, recorded date, location of data collection, and number of objects (e.g., files, rows, columns) are mandatory items to be included. Furthermore, data ownership properties are also essential to get an insight into data reusability. Data citation and intellectual property rights are mandatory data ownership properties based on respondents' opinions. Furthermore, to make data ready for reuse, data privacy and security rules or scenarios have to be prepared, such as data backup, the protection of sensitive data, data anonymity or pseudonymity, and defined data ethics from the community and government. The respondents also considered fully recorded data provenance information in human and machine-readable formats.

4.3. Case study 2- Indonesian fish farming

The second case study involved fishery experts to gather their opinions regarding FAIRification in the fish farming domain within the project. Survey results indicate the participants have substantial experience working with data for their research or business. Similar to the first case, almost all respondents agreed that data exchange is genuinely needed to enrich the information of the data for the research. For findability, metadata is still essential to help data seekers find relevant information about the data. Thus, the experts opted for specialized and rich metadata to accompany the dataset so that the machine could understand it. In addition, web search engines are also a popular approach for finding data. As regards accessibility, the experts preferred to query domain-based and local-based systems to access the preferred dataset, while the general index, such as Google, is optional to get relevant datasets using their personal computers or mobile devices. Regarding data access security, user authorization and access control are mandatory approaches, as almost all participants (89 %) chose to use the registered-based system to give other people access to the data. Furthermore, they used internet protocols such as HTTPS and FTP to access the dataset instead of using old-fashioned practices such as email or telephone.

Regarding interoperability, 89 % of respondents considered the data exchange could be done using software (e.g., API) and network protocols (e.g., HTTPS) instead of hardware, such as an external hard disk or USB. In terms of syntactic processes, most of them worked with non-machine-readable data formats and with all types of data structures. They considered that the datasets should be able to explain using standardized descriptions, such as ontology, vocabulary and terminology. For reusability, data documentation is an important part of the FAIRification process, which contains file naming, recorded date, methods and tools for collecting data, data collection location, etc. The other things that need to be considered to make the data reusable are data citation and intellectual property. Data security and privacy scenarios, such as data backup, sensitive data protection, ethics, and constraint rules of data access protocol, should be carefully and properly governed. Finally, fully recorded provenance information in machine-readable format must be provided. Table 4 shows the summary of mandatory approaches for FAIRification in fish farming.

Table 4
FAIR implementation approaches in fish farming.

FAIR Principles	Implementation Approaches	Mandatory Approaches
Findability	Metadata	Rich Metadata
	Web search engine	Domain-specific portal Repository
	Repository	Private cloud
Accessibility	Registered-based access	User authorization Access control
	Query	Domain-based system Local-based systems
	Access protocol	HTTPS FTP
	Access devices	Personal computer Mobile devices
Interoperability	Technical interoperability	Software (e.g., API) Network (e.g., HTTPS)
	Syntactic interoperability	Structured data Semi-structured Unstructured
	Semantic interoperability	Ontology Terminology Vocabulary
Reusability	Data documentation	File naming Version date Data unit Methods and tools used for collecting data Recorded date Location of data collection
	Data ownership	Data citation Intellectual property rights
	Data privacy and security	Data backup scenario Ethics Sensitive data security Constraint policies
	Data provenance	Fully recorded in a machine-readable format

4.4. Comparing FAIR approaches in dairy and fish farming

Based on the previous section, the preferences for potential approaches to implementing FAIR principles were discussed. To further understand the process of FAIRification, a comparative analysis between the two domains is essential (Lupu et al., 2023). In this section, expert opinions from both domains were analyzed and compared to identify similarities and differences in their preferred approaches. In the context of findability, participants from both domains agreed on the importance of metadata in finding data, as shown in Table 5.

Table 5
Experts opinions on the preferred approaches in finding datasets.

Which approach(es) would you prefer to find other people's datasets relevant to your business/research	Dairy		Fish	
	Frequency of responses	Percentage	Frequency of responses	Percentage
Network	3	50 %	4	44 %
Literature search	3	50 %	4	44 %
Web search engine	2	33 %	7	78 %
Using digital identifiers	2	33 %	3	33 %
Metadata	3	50 %	7	78 %
Other: Combine	1	17 %	1	11 %
Other: Field database	0	0 %	1	11 %

Furthermore, Fig. 7 illustrates that participants consider metadata important for explaining datasets. As a result, the need for either specialized or rich metadata is evident, as depicted in Fig. 8. Regarding alternative methods for discovering data, network and literature searches are popular in dairy farming, while experts in fish farming tend to use web search engines.

According to Table 6, user authorization was the most popular method for protecting access to the data. Respondents from both domains agreed that it is crucial to secure data access. As shown in Fig. 9,

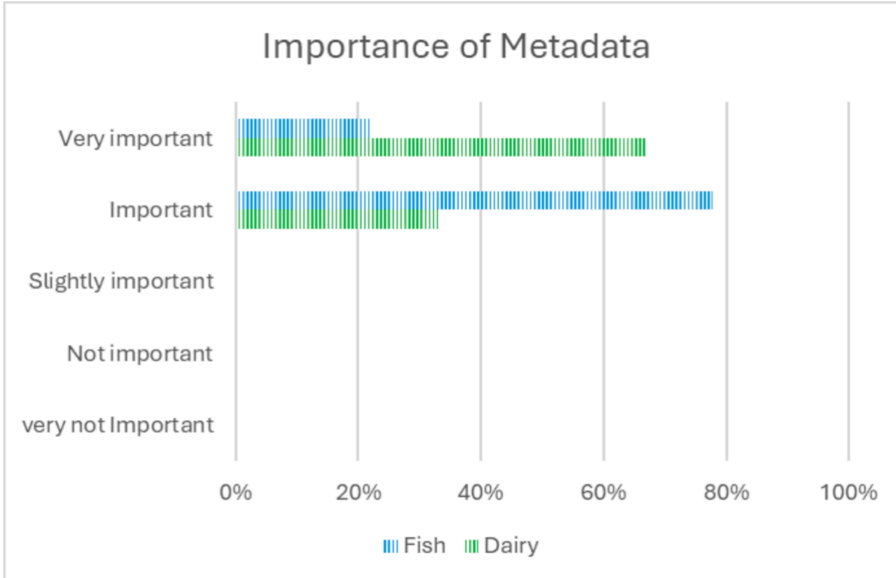

Fig. 7. The respondents' opinion on the importance of metadata.

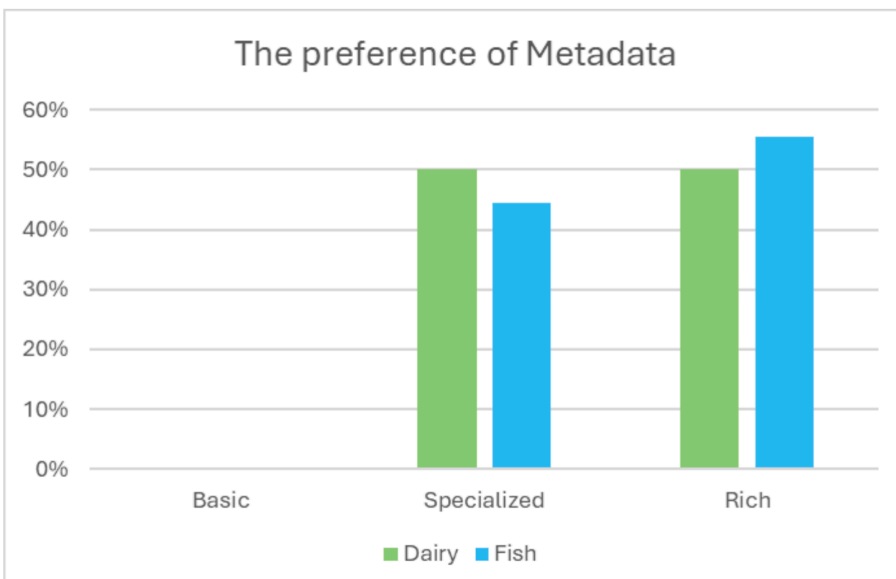

Fig. 8. The preference of Metadata types.

Table 6
The preferred data security approaches.

Which one of the data security approaches do you prefer to protect access to your data	Dairy		Fish	
	Frequency of responses	Percentage	Frequency of responses	Percentage
User authentication	2	33 %	2	22 %
User authorization	4	67 %	8	89 %
Access control	2	33 %	5	56 %

the registered-based access method was the most commonly chosen approach for granting data seekers access to the data. It indicates that stakeholders considered it essential to have a safe and proper way to manage access to or exchange their datasets. As a result, HTTPS and FTP were the preferred internet protocols for data exchange. Accessibility is not about opening access to the data; instead, it is the principle that governs how people can access or grant access to the dataset.

No respondent wants to exchange their data through hardware tools (e.g., USB drives or external hard disks). Instead, they tend to choose safer options, such as using software (e.g., APIs) or networks (e.g., HTTPS). Table 7 shows the respondents' opinions on data exchange approaches.

Furthermore, Table 8 shows that providing a detailed explanation of data elements using a standardized format, including vocabulary, terminology, and ontology, was the most preferred approach for integrating the data. None of the respondents wanted to choose data elements without any description. In fact, data elements with simple descriptions, like a data dictionary, were more commonly chosen by fisheries experts than those from dairy.

Regarding the reusable principle, dairy and fish farming experts use various file formats, making data documentation essential. Both domains agreed that file naming and the methods and tools used for data

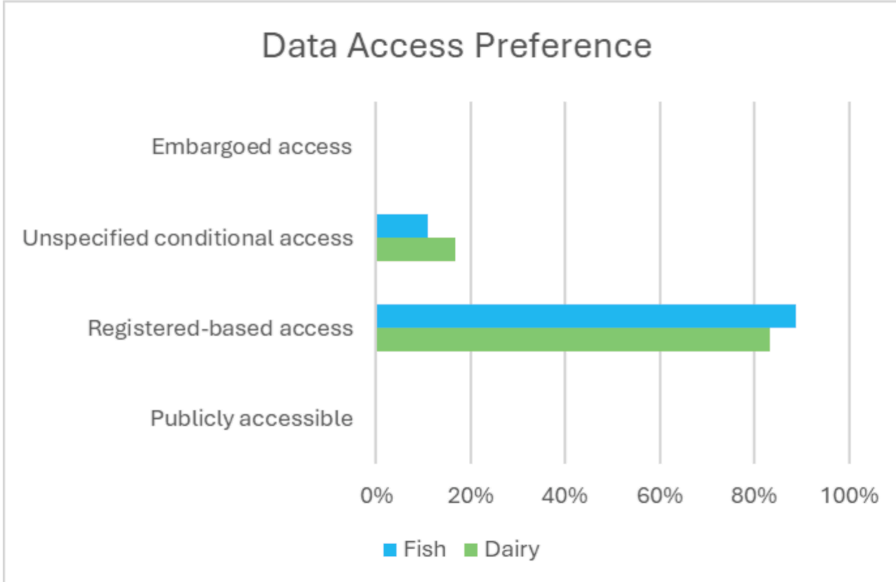

Fig. 9. The preferences of respondents on data access approaches.

Table 7
The preferred data exchange approaches.

Which one(s) of the following approaches do you prefer to exchange your data	Dairy		Fish	
	Frequency of responses	Percentage	Frequency of responses	Percentage
Hardware	0	0 %	0	0 %
Software	4	67 %	8	89 %
Network	4	67 %	8	89 %

Table 8
The responses of the experts on the data elements information.

Which one(s) of the following information do you prefer when integrating your dataset with others	Dairy		Fish	
	Frequency of responses	Percentage	Frequency of responses	Percentage
Data element without description	0	0 %	0	0 %
Data element with simple description	1	17 %	4	44 %
Data element with standardized description	5	83 %	6	67 %

collection are important documentation elements. In dairy farming, experts emphasized that documenting the number of objects (e.g., files, rows, columns) is particularly significant, as they primarily work with numerical data. Meanwhile, experts in fish farming highlighted the importance of recording the date and location of data collection, as they also frequently work with image and textual data in addition to numerical data. For data ownership, both domains favored data citation and intellectual property rights as the key ownership attributes. Regarding privacy and security, both domains agreed on the importance of data backup and complying with ethical guidelines set by the community or government. Furthermore, in fish farming, constraint policies and rules defining data access protocols were identified as one of the most popular methods for managing privacy and security, although this was the least preferred method in dairy farming. The details of the survey results, including tables and figures comparing FAIR approaches in dairy and fish farming, are provided in Appendix.

5. Discussion

5.1. General discussion

FAIRification is not a straightforward process. It needs to consider several detailed and highly specific approaches for each principle. The stakeholders' opinions and considerations are also crucial to the FAIRification task. In this study, each principle of FAIR was analyzed systematically. Thus, the current study presents a novel systematic approach for implementing FAIR principles (FAIRification) in data management systems of two different Indonesian agricultural domains: dairy and fish farming.

The domain analysis was used to systematically derive information from the literature regarding the adoption of FAIR principles in agricultural systems. Several research papers from the domain analysis discussed the FAIRification process in agricultural systems were obtained. To enhance our understanding of the FAIRification process, various grey literature, including websites, books, and reports, were also incorporated into the analysis. Thus, the domain modeling resulted in a feature diagram to depict the various possible approaches for meeting each point in the FAIR principles. All potential FAIRification approaches found in the literature were included in the feature diagrams to ensure a comprehensive analysis.

For both case studies, all variant approaches and practices for implementing FAIR principles from the feature diagrams were presented to the dairy and fish farming experts. Experts from universities and industries involved in the smart-in-ag project participated in the survey to provide their opinions on the FAIR approaches in the feature diagrams. The online survey was conducted from September to October 2023 using Google Forms. There are 14 experts who participated in the study and shared their views. The approaches chosen by at least 50 % of the experts in each domain were considered mandatory for the FAIRification process in the project.

The survey results revealed that experts from both domains unanimously agreed that metadata is essential for ensuring findability, particularly metadata containing rich dataset information. At this stage, the respondents demonstrated sufficient awareness of FAIR principles by acknowledging the critical role of metadata. It is a positive sign from the survey results as the awareness of the stakeholders is one of the challenges in FAIRification process (Gacenga et al., 2024). In addition to metadata, web search engines were among the most popular tools for finding desired data through repositories or domain-specific portals. The respondents preferred using private repositories to store their datasets, as these repositories are provided by domain organizations, and they are governed by strict access protocols. Accessing trustworthy datasets is critical for researchers, as they rely on such data to conduct their research and draw reliable conclusions (Allemang and Teegarden, 2017).

In agriculture, the sensitivity of data has become a critical issue that must be addressed when sharing and accessing agricultural datasets (Allemang and Teegarden, 2017; Dey and Shekhawat, 2021). Platforms designed to handle, manage, and secure data are essential to ensure the security of data sharing among users. For security reasons, most respondents preferred secure methods for accessing datasets. This preference is reflected in the survey results, where the majority of respondents favored systems requiring user permission to access datasets. Consequently, user authorization received the highest number of responses. Experts also prefer safe resources to obtain datasets, such as domain-based systems provided by domain-specific organizations (e.g., dairy organizations or aquaculture communities). Furthermore, as web search engines have become one of the most popular approaches for finding datasets, internet protocols such as HTTPS or FTP were chosen as mandatory methods for accessing the dataset, rather than contact protocols that involve directly contacting data owners via email, telephone, or Skype.

Furthermore, as the respondents agreed that the data generated by the project should remain accessible even after the researchers have left, data elements with standardized descriptions are especially important for facilitating data integration (Bahlo et al., 2019) and reducing dependence on the researchers (Ali and Dahlhaus, 2022). Therefore, participants agreed that rich metadata is required to provide information that humans and machines can read. To achieve this, data documentation plays a significant role by providing descriptions of data versioning, data collection, and data volume, which aligns with the survey results.

In addition to data documentation, aspects such as data ownership, data privacy and security, and data provenance are also important considerations when reusing data (Allemang and Teegarden, 2017), which aligns with the participants' view. According to the survey, data citation and intellectual property rights are mandatory approaches to ensure data ownership. This is especially relevant in the case of dairy farming, as some participants still use traditional methods to find data, such as through colleagues or literature. Therefore, data citation should be ensured. Furthermore, data backup scenarios and adherence to ethics defined by the government and community are crucial to ensuring privacy and security. Regarding data provenance, all participants require fully recorded provenance information, especially in machine-readable format. However, as most participants still use non-machine-readable data, they also prefer fully recorded provenance information in text format.

5.2. Benefits and challenges

Several potential benefits and challenges were identified through a systematic approach to investigating FAIR principles applications. The first benefit is that the study carefully analyzed each aspect of the FAIR principles, providing valuable insights into their practical implementation. By doing so, the study highlights various strategies for making a FAIR data management system. Second, stakeholders play a pivotal role in data management systems, so their opinions and perspectives on each proposed approach are invaluable. The study shows that by involving stakeholders in choosing the appropriate approach, the mandatory practices for FAIRification can be derived based on their requirements. This process is highly beneficial for future data management development.

To the best of our knowledge, research articles in the literature addressing the adoption of FAIR principles in agricultural systems are very limited. Therefore, this study was conducted to explore various potential FAIRification approaches that can inform the development of FAIR data management strategies. The investigation applies a systematic approach to identify potential FAIRification strategies using the DSR method, which can be easily adapted and reused in other domains.

Since the DSR method provides systematic steps, it allows researchers or practitioners from other domains to follow the same process when investigating FAIRification in their fields. First, they need to conduct a domain analysis, which consists of domain scoping and domain modeling. Domain scoping is used to identify relevant knowledge sources to derive the features, while domain modeling is used to visualize common and familiar features through a feature model. After that, validation is conducted by asking stakeholders involved in the domain to identify the necessary approaches to achieve FAIRification in their fields.

Aside from the benefits, several issues were identified in the study. First, nearly all stakeholders from the dairy or fish domain still used a proprietary data format. This poses a significant challenge for FAIRification, as additional efforts are required to address this issue, such as transforming these data into more interoperable formats (Allemang and Teegarden, 2017). Moreover, transforming datasets is a complex process that requires close collaboration between data owners and data engineers. However, the transformation process from proprietary to non-proprietary or machine-readable data formats offers the advantage of enhancing data maturity, as it enables data longevity and allows broad communities to access the data (Allemang and Teegarden, 2017; FAIR Data Maturity Model Working Group, 2020).

Although metadata is the preferred method for finding data, experts still consider conventional methods, such as contacting colleagues or librarians. From this finding, several aspects need to be considered for FAIRification purposes, especially when the dataset is created without metadata or proper documentation (Jacobsen et al., 2020b; Top et al., 2022). The first aspect that needs to be considered is the privacy and security of the data, including ethics and sensitive data protection. After that, to integrate the data, the dataset should be accompanied, at a minimum, by a data dictionary or, even better, by vocabulary, terminology, and ontology. To make the data more findable, the use of open data platforms is encouraged, such as institution repository, along with proper data citation. All these efforts can significantly enhance trust in data sharing between data seekers and providers (Food and Agriculture Organization of the United Nations (FAO). (2021)).

Another issue is that the respondents' experience in the domain must be carefully considered in order to implement FAIR principles (Gacenga et al., 2024). To address this, the study excluded participants without sufficient experience in smart farming systems from the survey. By doing so, the study ensured that the expected level of FAIRness could be reliably achieved and trusted.

5.3. Potential improvements

Although the process of selecting and defining potential FAIRification approaches was thorough and carefully executed, we acknowledge that many FAIRification approaches were not covered in the current investigation due to the scope of our research. Additionally, while the involvement of domain experts from the project provided sufficient input to derive the mandatory approaches, including experts from outside the smart-in-ag project could offer different perspectives that might improve the results of this study. Furthermore, incorporating interviews with experts to verify the FAIRification process could complement the survey data and enhance the overall understanding of the process.

6. Conclusion

This research presented and examined how the FAIR principles can be effectively applied by studying real-world examples using an interdisciplinary project involving dairy and fish farming. Each FAIR principle is analyzed in detail, delving into the specific requirements and considerations for achieving them in the context of agricultural data. Overall, it was demonstrated that dairy and fish farming stakeholders have similar requirements for achieving FAIRification. All experts in both domains highly valued the significance of metadata in explaining and obtaining datasets. Furthermore, they emphasized the importance of a strict and secure data access protocol, semantic interoperability, and rich information of data documentation to implement FAIR principles. This insight is valuable for guiding FAIR principles implementation within data management systems for dairy and fish farming to ensure that data can be effectively managed, shared, and reused. However, further research is needed to examine FAIR principles in different domains. This study adopts the DSR method to identify potential FAIR approaches, which offers a highly adaptable and reusable framework applicable across various fields. The DSR method's systematic steps enables researchers or practitioners from other domains to replicate the process when investigating FAIRification within their specific contexts. Thus, this study encourages a further investigation and implementation of FAIR data principles, particularly in smart farming systems.

CRediT authorship contribution statement

Ngakan Nyoman Kutha Krisnawijaya: Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Bedir Tekinerdogan:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Cagatay Catal:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Rik van der Tol:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yeni Herdiyeni:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

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.

Appendix A. Supplementary data

Supplementary data associated with the survey instruments and results used in this article can be found in the online version at https://doi.org/10.1016/j.compag.2024.109855

Data availability

No data was used for the research described in the article.

References

Ali, B., Dahlhaus, P., 2022. The Role of FAIR Data towards Sustainable Agricultural Performance: A Systematic Literature Review. Agriculture 2 (12). https://doi.org/10.3390/agriculture12020309.

Allemang, D., Teegarden, B., 2017. A global data ecosystem for agriculture and food. F1000Research 6 (1844), 1844. https://doi.org/10.1079/CABICOMM-79-12.

Andrade, R., Urioste, S., Rivera, T., Schiek, B., Nyakundi, F., Vergara, J., Mwanzia, L., Loaiza, K., Gonzalez, C., 2021. Where Is My Crop? Data-Driven Initiatives to Support Integrated Multi-Stakeholder Agricultural Decisions. Front. Sustainable Food Syst. 5. https://doi.org/10.3389/fsufs.2021.737528.

Arefolov, A., Adam, L., Brown, S., Budovskaya, Y., Chen, C., Das, D., Farhy, C., Ferguson, R., Huang, H., Kanigel, K., Lu, C., Polesskaya, O., Staton, T., Tajhya, R., Whitley, M., Wong, J.-Y., Zeng, X., McCreary, M., 2021. Implementation of the FAIR Data Principles for Exploratory Biomarker Data from Clinical Trials. Data Intell. 3 (4), 631–662. https://doi.org/10.1162/dint_a_00106.

Australian Research Data Commons (ARDC). (2022). FAIR Data Self Assessment Tool. Retrieved 7 May from https://ardc.edu.au/resource/fair-data-self-assessment-tool/.

Bahim C, Dekkers M, & Wyns B. (2019). Results of an Analysis of Existing FAIR Assessment Tools.

Bahlo, C., Dahlhaus, P., Thompson, H., Trotter, M., 2019. The role of interoperable data standards in precision livestock farming in extensive livestock systems: A review. Comput. Electron. Agric. 156, 459–466. https://doi.org/10.1016/j.compag.2018.12.007.

Borycz, J., Carroll, B., 2020. Implementing FAIR data for people and machines: Impacts and implications - results of a research data community workshop. Inf. Serv. Use 40 (1–2), 71–85. https://doi.org/10.3233/isu-200083.

Celebi, R., Rebelo Moreira, J., Hassan, A.A., Ayyar, S., Ridder, L., Kuhn, T., Dumontier, M., 2020. Towards FAIR protocols and workflows: the OpenPREDICT use case. PeerJ Comput Sci 6, e281. https://doi.org/10.7717/peerj-cs.281.

Dey, K., Shekhawat, U., 2021. Blockchain for sustainable e-agriculture: Literature review, architecture for data management, and implications. J. Clean. Prod. 316. https://doi.org/10.1016/j.jclepro.2021.128254.

Dorich, C.D., Conant, R.T., Albanito, F., Butterbach-Bahl, K., Grace, P., Scheer, C., Snow, V.O., Vogeler, I., van der Weerden, T.J., 2020. Improving N2O emission estimates with the global N2O database. Curr. Opin. Environ. Sustain. 47, 13–20. https://doi.org/10.1016/j.cosust.2020.04.006.

FAIR Data Maturity Model Working Group. (2020). FAIR Data Maturity Model. Specification and Guidelines.

Food and Agriculture Organization of the United Nations (FAO). (2021). AGROVOC – Semantic data interoperability on food and agriculture. https://doi.org/10.4060/cb2838en.

Gacenga, F., An-Vo, D.-A., McCulloch, J., Young, R., Cobon, D., 2024. Making Australian Drought Monitor dataset findable, accessible, interoperable and reusable. Comput. Electron. Agric. 226. https://doi.org/10.1016/j.compag.2024.109381.

Garcia, L., Bolleman, J., Gehant, S., Redaschi, N., Martin, M., UniProt, C., 2019. FAIR adoption, assessment and challenges at UniProt. Sci Data 6 (1), 175. https://doi.org/10.1038/s41597-019-0180-9.

Giray, G., Catal, C., 2021. Design of a Data Management Reference Architecture for Sustainable Agriculture. Sustainability 13, 1–17. https://doi.org/10.3390/su13137309.

GO FAIR. (2022). *FAIR Principles*. Retrieved 18 November from https://www.go-fair.org/fair-principles/.

Goyal, S., 2014. Public vs Private vs Hybrid vs Community - Cloud Computing: A Critical Review. International Journal of Computer Network and Information Security 6 (3), 20–29. https://doi.org/10.5815/ijcnis.2014.03.03.

Gregory, K., Khalsa, S.J., Michener, W.K., Psomopoulos, F.E., de Waard, A., Wu, M., 2018. Eleven quick tips for finding research data. PLoS Comput. Biol. 14 (4), e1006038. https://doi.org/10.1371/journal.pcbi.1006038.

Hauschke, C., Nazarovets, S., Altemeier, F., Kaliuzhna, N., 2021. Roadmap to FAIR Research Information in Open Infrastructures. J. Libr. Metadata 21 (1–2), 45–61. https://doi.org/10.1080/19386389.2021.1999156.

Hevner, A.R., March, S.T., Park, J., Ram, S., 2004. Design Science in Information Systems Research. MIS q. 28 (1), 75–105. https://doi.org/10.2307/25148625.

Himanen, L., Geurts, A., Foster, A.S., Rinke, P., 2019. Data-Driven Materials Science: Status, Challenges, and Perspectives. Adv Sci (Weinh) 6 (21), 1900808. https://doi.org/10.1002/advs.201900808.

Jacobsen, A., de Miranda Azevedo, R., Juty, N., Batista, D., Coles, S., Cornet, R., Courtot, M., Crosas, M., Dumontier, M., Evelo, C.T., Goble, C., Guizzardi, G., Hansen, K.K., Hasnain, A., Hettne, K., Heringa, J., Hooft, R.W.W., Imming, M., Jeffery, K.G., Schultes, E., 2020a. FAIR Principles: Interpretations and Implementation Considerations. Data Intell. 2 (1–2), 10–29. https://doi.org/10.1162/dint_r_00024.

Jacobsen, A., Kaliyaperumal, R., da Silva Santos, L.O.B., Mons, B., Schultes, E., Roos, M., Thompson, M., 2020b. A Generic Workflow for the Data FAIRification Process. Data Intell. 2 (1–2), 56–65. https://doi.org/10.1162/dint_a_00028.

Kamilaris, A., Kartakoullis, A., Prenafeta-Boldú, F.X., 2017. A review on the practice of big data analysis in agriculture. Comput. Electron. Agric. 143, 23–37. https://doi.org/10.1016/j.compag.2017.09.037.

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 (5), 4498–4510. https://doi.org/10.1002/agj2.20285.

Khatoon, P.S., Ahmed, M., 2022. Importance of semantic interoperability in smart agriculture systems. Trans. Emerg. Telecommun. Technol. 33 (5). https://doi.org/10.1002/ett.4448.

Koers, H., Bangert, D., Hermans, E., van Horik, R., de Jong, M., Mokrane, M., 2020. Recommendations for Services in a FAIR Data Ecosystem. Patterns (N Y) 1 (5), 100058. https://doi.org/10.1016/j.patter.2020.100058.

Krisnawijaya, N.K.N., Tekinerdogan, B., Catal, C., Tol, R.v.d., 2022. Data analytics platforms for agricultural systems: A systematic literature review. Comput. Electron. Agric. 195. https://doi.org/10.1016/j.compag.2022.106813.

Krisnawijaya, N.N.K., Tekinerdogan, B., Catal, C., van der Tol, R., 2024. Reference architecture design for developing data management systems in smart farming. Eco. Inform. 81, 102613. https://doi.org/10.1016/j.ecoinf.2024.102613.

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. https://doi.org/10.1016/j.agsy.2023.103822.

Lehne, M., Sass, J., Essenwanger, A., Schepers, J., Thun, S., 2019. Why digital medicine depends on interoperability. NPJ Digit Med 2, 79. https://doi.org/10.1038/s41746-019-0158-1.

Lin, Y., Purnama Jati, P.H., Aktau, A., Ghardallou, M., Nodehi, S., van Reisen, M., 2022. Implementation of FAIR Guidelines in Selected Non-Western Geographies. Data Intell. 4 (4), 747–770. https://doi.org/10.1162/dint_a_00169.

Lupu, V., Cujba, R., & Sobetchi, V. (2023). The Attitudes of Agricultural Researchers Towards Data Sharing: Case Study of the Republic of Moldova Central and Eastern European eDem and eGov Days 2023.

MacLeod, A., Wong, M., Gregory, L., Schneider, D., Williams, A., Castleden, I, Simons, B., Levett, K., Box, P., (2020). The Agricultural Research Federation (AgReFed) Technical and Information Policy Suite. Version 1.0. Endorsed by the AgReFed Council at 13 May 2020.

Mehrabi, Z., McDowell, M.J., Ricciardi, V., Levers, C., Martinez, J.D., Mehrabi, N., Wittman, H., Ramankutty, N., Jarvis, A., 2020. The global divide in data-driven farming. Nat. Sustainability 4 (2), 154–160. https://doi.org/10.1038/s41893-020-00631-0.

Oliveira, S.M.I., Barros Lima, G.d.F., Farias Lóscio, B., 2019. Investigations into Data Ecosystems: a systematic mapping study. Knowl. Inf. Syst. 61 (2), 589–630. https://doi.org/10.1007/s10115-018-1323-6.

Rix, C., Stein, H., Chen, Q., Frank, J., Maass, W., 2021. Conceptualizing data ecosystems for industrial food production. In: 2021 IEEE 23rd Conference on Business Informatics (CBI), Vol. 1. IEEE, pp. 201–210.

Roussaki, I., Doolin, K., Skarmeta, A., Routis, G., Lopez-Morales, J.A., Claffey, E., Mora, M., Martinez, J.A., 2023. Building an interoperable space for smart agriculture. Digital Communications and Networks 9 (1), 183–193. https://doi.org/10.1016/j.dcan.2022.02.004.

Runeson, P., Höst, M., 2009. Guidelines for conducting and reporting case study research in software engineering. Empir. Softw. Eng. 14, 131–164. https://doi.org/10.1007/s10664-008-9102-8.

Saiz-Rubio, V., Rovira-Más, F., 2020. From Smart Farming towards Agriculture 5.0: A Review on Crop Data Management. Agronomy 10, 1–21. https://doi.org/10.3390/agronomy10020207.

Science Europe. (2021). Practical Guide to the International Alignment of Research Data Management. https://www.scienceeurope.org/media/4brkxxe5/se_rdm_practical_guide_extended_final.pdf.

Singh, A., Jones, S., Ganapathysubramanian, B., Sarkar, S., Mueller, D., Sandhu, K., Nagasubramanian, K., 2021. Challenges and Opportunities in Machine-Augmented Plant Stress Phenotyping. Trends Plant Sci 26 (1), 53–69. https://doi.org/10.1016/j.tplants.2020.07.010.

Tantalaki, N., Souravlas, S., Roumeliotis, M., 2019. Data-Driven Decision Making in Precision Agriculture: The Rise of Big Data in Agricultural Systems. Journal of Agricultural & Food Information 20 (4), 344–380. https://doi.org/10.1080/10496505.2019.1638264.

Tekinerdogan, B., 2022. Software Ecosystems for Precision Agriculture. In: Zhang, Q. (Ed.), Encyclopedia of Smart Agriculture Technologies. Springer International Publishing, pp. 1–10. https://doi.org/10.1007/978-3-030-89123-7_269-1.

Tekinerdogan, B., & Öztürk, K. (2013). Feature-Driven Design of SaaS Architectures. In Z. Mahmood & S. Saeed (Eds.), Software Engineering Frameworks for the Cloud Computing Paradigm (pp. 189-212). Springer London. https://doi.org/10.1007/978-1-4471-5031-2_9.

Thompson, M., Burger, K., Kaliyaperumal, R., Roos, M., da Silva Santos, L.O.B., 2020. Making FAIR Easy with FAIR Tools: From Creolization to Convergence. Data Intell. 2 (1–2), 87–95. https://doi.org/10.1162/dint_a_00031.

Top, J., Janssen, S., Boogaard, H., Knapen, R., Şimşek-Şenel, G., 2022. Cultivating FAIR principles for agri-food data. Comput. Electron. Agric. 196. https://doi.org/10.1016/j.compag.2022.106909.

van Geest, M., Tekinerdogan, B., Catal, C., 2021. Design of a reference architecture for developing smart warehouses in industry 4.0. Comput. Ind. 124, 103343. https://doi.org/10.1016/j.compind.2020.103343.

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Mons, B., 2016. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018. https://doi.org/10.1038/sdata.2016.18.

Wilkinson, M.D., Sansone, S.A., Schultes, E., Doorn, P., da Silva, B., Santos, L.O., Dumontier, M., 2018. A design framework and exemplar metrics for FAIRness. Sci Data 5, 180118. https://doi.org/10.1038/sdata.2018.118.

Wise, J., de Barron, A.G., Splendiani, A., Balali-Mood, B., Vasant, D., Little, E., Mellino, G., Harrow, I., Smith, I., Taubert, J., van Bochove, K., Romacker, M., Walgemoed, P., Jimenez, R.C., Winnenburg, R., Plasterer, T., Gupta, V., Hedley, V., 2019. Implementation and relevance of FAIR data principles in biopharmaceutical R&D. Drug Discov Today 24 (4), 933–938. https://doi.org/10.1016/j.drudis.2019.01.008.

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 Sci. J. 21. https://doi.org/10.5334/dsj-2022-013.

Zhang, C., Di, L., Lin, L., Zhao, H., Li, H., Yang, A., Guo, L., Yang, Z., 2023. Cyberinformatics tool for in-season crop-specific land cover monitoring: Design, implementation, and applications of iCrop. Comput. Electron. Agric. 213. https://doi.org/10.1016/j.compag.2023.108199.