Review
The Role of FAIR Data towards Sustainable Agricultural Performance: A Systematic Literature Review

Basharat Ali 1,2,* and Peter Dahlhaus 1,2

1 Centre for eResearch and Digital Innovation (CeRDI), Greenhill Enterprise Centre Ballarat Technology Park, Federation University, Mount Helen, VIC 3350, Australia; p.dahlhaus@federation.edu.au
2 Food Agility Cooperative Research Centre, Ultimo, NSW 2007, Australia
* Correspondence: basharat.ali2@federation.edu.au

Abstract: Feeding a growing global population requires improving agricultural production in the face of multidimensional challenges; and digital agriculture is increasingly seen as a strategy for better decision making. Agriculture and agricultural supply chains are increasingly reliant on data, including its access and provision from the farm to the consumer. Far-reaching data provision inevitably needs the adoption of FAIR (Findable, Accessible, Interoperable, and Reusable) that offer data originators and depository custodians with a set of guidelines to safeguard a progressive data availability and reusability. Through a systematic literature review it is apparent that although FAIR data principles can play a key role in achieving sustainable agricultural operational and business performance, there are few published studies on how they have been adopted and used. The investigation examines: (1) how FAIR data assimilate with the sustainability framework; and (2) whether the use of FAIR data by the agriculture industry, has an impact on agricultural performance. The work identifies a social science research gap and suggests a method to guide agriculture practitioners in identifying the specific barriers in making their data FAIR. By troubleshooting the barriers, the value propositions of adopting FAIR data in agriculture can be better understood and addressed.

Keywords: FAIR data; findable; accessible; interoperable; reusable; sustainability

1. Introduction

The global need for agricultural production has been increasing [1], and most food production remains soil-based. With nearly all arable land under cultivation [2], agriculture is projected to encounter several challenges including: sustaining maximum production, limited natural resources, endangered environments and ecosystems, soil degradation and erosion [3,4]. It is quantitatively confirmed that farming strategies, methods and decision making are key factors in the future of sustainable and enhanced agricultural production [5,6]. Operationally, digital agriculture and data/information mutually rely on each other [3,7,8]. The volume of digital data in agricultural landscapes has grown exponentially, much of it collected by sensors (both remote sensing and the Internet of Things). Agricultural knowledge building, appropriate management responses [9] and farm management decisions [10] highly depend on the data collected through the use of digital technologies, and ubiquitous internet technologies provide access to all these data, delivered on demand via high-speed broadband to mobile tablet devices [11]. Precision agriculture technology employs data [12] to perform operations such as economising crop inputs, optimising machinery performance and appropriate location finding [13]. The digital agricultural revolution has led to a plethora of websites and mobile applications (Apps) that are now available to assist the farmer, agronomist, agribusiness investor, landscape manager and researcher in decision making. However, the applications and tools are only as good as the data they use, and because of the disparity of data collection, formats, and storage only a fraction of the required data are utilised. Current common
limitations in system models for decision support are: (1) data scarcity (quantity, resolution, and quality) and (2) inadequate knowledge systems to effectively communicate the results to the end-user. These limitations are greater obstacles to the use of the tools than gaps in theory or technology [14,15]. Seamless automated data collection (from both public and private sources), data interoperability and the federation of multidisciplinary data (plant, animal, soil, land, climate, weather, machinery, farm business, economics, marketing, trade, etc.) are required, preferably utilising open cloud-based systems for data storage and open standards for data exchange. Combining these data in new technologies, such as those deploying data mining, machine learning, artificial intelligence algorithms and digital twins, will ultimately provide the holistic viewpoint needed for sustainable agricultural production [3,7,16].

Improving agricultural knowledge, appropriate management responses [9] and farm management decisions [10] requires the stakeholders to increasingly depend on data collected through the use of digital technologies, such as the internet, sensors and mobile computing at a more localised scale [11]. Precision agricultural technologies massively employ data [12] to perform operations such as economised crop inputs based on high resolution location finding [3,16–18], hence requiring a robust database management system with far-reaching data provision [19]. Since financial benefits highly depend upon the suitability of the chosen farming techniques and technology, access to data, data literacy and/or technical support are important needs in the adoption of innovative practices in the agricultural sector [20–22]. FAIR data principles, first published by Wilkinson [23], provide data originators and data depository custodians with a set of guidelines to ensure data availability and reusability. While recognising an ever-growing need for automation of wide-ranging data encounter, recovery, integration and analysis, the FAIR data principles make this goal possible, by combining multidisciplinary, cross-disciplinary data, with disparate data formats, from different sources, as well as by emphasizing that each of the principles should be equally valid to both humans and machines [23–25].

FAIR (findable, accessible, interoperable, reusable) data principles add more value to data by enhancing data utility, especially for legacy data sets. Data findability focuses on maintaining its worldwide uniqueness through persistent identifiers that are machine-readable and index-able to assist individuals and artificially intelligent systems. It is an exclusive, persistent way to refer to the data using standard digital object identifiers (DOI), uniform resource identifiers (URI) or uniform resource locators (URL). Accessibility ensures that the data and other digitised information are available subject to specified conditions of access. It has three major components, i.e., access protocol, access permission and metadata permanency. Accessibility confirms that the data can be accessed by humans and devices using standard internet protocols, provided that the access controls allow that. The controls may require managing data licensing in a convenient way. Interoperability makes sure that the data and other digitised information is unambiguously understood machine-to-machine. In accordance with FAIR principles, data and metadata should be conveyed using syntactic and semantic data structures covering the raw and highly processed data. Ontologies, communicated through a resource description framework (RDF) or other open-source frameworks, can carry data integration across the board. Through interoperability, heterogeneous data distributed across disparate databases and devices can be brought together in standardised and harmonised formats. Reusability segregates conventional data management from FAIR data stewardship, which requires a multi-layered approach that addresses the demand of data to be reusable. Fully described contextual and descriptive machine-readable metadata is required to allow new consumers to reuse data for new needs and applications, decades after those data were collected. It also requires enhanced provenance metadata for tracing changes to names, editions, and parameters of analysis. Principally, if data are not comprehensively and unambiguously described in a machine-readable form, they have little benefit, regardless of whether they are published with open access [23–28]. The adoption of FAIR data principles, a thorough review of literature reveals,
is in its infancy in agricultural systems despite recognition of its value and development of guidelines.

This study conducts a systematic literature review aimed at: (1) understanding the adoption and use of FAIR data principles in the agriculture industry; (2) demonstrating by what means the sub-indicator/characteristics of FAIR data complement sustainable agricultural performance indicators, and the way the sustainability indicators assimilates with the sustainable agricultural performance indicators; and (3) whether FAIR data have impacted agricultural industry performance. Furthermore, this study uniquely describes how the adoption of FAIR data contribute to sustainable agricultural (operational and business) performance, by summarising the literature with comprehensive methodological approaches that demonstrate FAIR data implementation processes. The research identifies a social science research gap and suggests a method to guide agriculture practitioners in identifying the specific barriers to making their data FAIR, and to duly analyse the barriers to better understand and address the value propositions of adopting FAIR data in agriculture.

2. Methodology

This systematic literature review follows more rigorous and transparent methodology, following guidelines by Van der Knaap et al. [29], Moher et al. [30] and Koutsos et al. [31].

2.1. Scoping

This study attempts to answer two related research questions: To what extent have FAIR data principles been adopted by the agricultural industry? and what role do FAIR data principles play in agricultural performance?

The research uses the Web of Science digital database, initially created by the Institute for Scientific Information (ISI) but now maintained by Clarivate [32], to search for more authentic and comprehensive scientific literature. Considering the multidisciplinary nature of FAIR data in agriculture, the search includes all the relevant Web of Science categories and pertinent publications/journals. The broad literature search shows a clearer research gap in this area, with no previous studies found that address the above research questions.

2.2. Planning

To find the qualified publications from the Web of Science, a far-reaching search approach was carried out by using the most appropriate terms or keywords (given below), combining with the Boolean operators (AND and OR). The search query we used is as follows:

(“FAIR data” OR “FAIR data principles” OR “FAIR data guidelines” OR “FAIR principles” OR “FAIR guidelines” OR “findability” OR “accessibility” OR “interoperability” OR “reusability” OR “findable” OR “accessible” OR “interoperable” OR “reusable” OR “datasets” OR “data sources”) AND (“metadata standards” OR “metadata schema” OR “metadata schemes” OR “big data” OR “data management” OR “database”)

The search was limited by document type (article, review), years (2016–present) and language (English). Following eligibility criteria, an additional search was conducted through Google Scholar [33], applying the ‘snowballing’ (or ‘backward and forward snowball’) technique [34]. The search was broadened because: (1) FAIR data are a relatively new subject in the agriculture disciplines; and (2) agriculture is a multidimensional and multidisciplinary subject that encompasses sociology, data science, economic, environmental sciences, etc. The suitability criterium of the additional literature was assessed as being peer-reviewed (approved by two independent reviewers), following Van derWindt et al. [35] guidelines.

The selection criteria of the selected studies through the Web of Science were based on: (1) the role of FAIR data towards agricultural performance; (2) the implementation of FAIR data principles: i.e., applied research; (3) a comprehensive methodological approach, tools
and/or rules, etc. To exclude studies not related to agricultural research (exclusion criteria), the query results were refined by choosing most appropriate Web of Science categories: (1) Agriculture Multidisciplinary; (2) Agriculture Dairy Animal Science; (3) Soil Science; (4) Plant Sciences; (5) Green Sustainable Science Technology; and (6) Computer Science Interdisciplinary Applications.

This review includes all the studies (with document type article, review) from the last six years (2016–present) since the FAIR data principles were first published in 2016. The search was confined to English language papers published by the top five publishing houses, i.e., Elsevier, Springer Nature, MDPI, Wiley and Taylor & Francis, since that covered most high-quality peer-reviewed papers.

At the initial stage, the total number of studies obtained was 1042. To further confine the research focus to agriculture, the selection excluded numerous journals having entirely different scope areas, for instance medical sciences, astronomy, industrial engineering, computation and mathematics, etc. bringing the total number of studies to be assessed to 469. There were five studies included through the ‘Google Scholar’ search. More details are provided in the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) statement (Figure 1).

![Systematic review based on PRISMA flowchart](image)

2.3. Identification/Search

Details of the studies found through the ‘advanced search option’ of the Web of Science were exported into a reference manager (Endnote) and spreadsheet (Excel) for further assessments. A Google Scholar search was again used to ensure that the maximum number of relevant papers were captured, given the multidisciplinary and multidimensional nature of the research objective. The snowballing technique was helpful for additional searching as there was a limit on searching imposed in the Web of Science database. All studies obtained from the search queries were fully checked based on their titles, abstracts, keywords, year of publication and the research area and so on, to confirm their eligibility for inclusion.
2.4. Screening

All the eligible literature found though the Web of Science and Google Scholar searches were saved (with title, keywords, abstracts) in two separate files. Duplicates (3 studies) were deleted. That resulted in 42 studies (full text) downloaded for a thorough assessment.

2.5. Eligibility/Assessment

The inclusion criteria comprise studies that: (1) cover the role of FAIR data in relation to agricultural performance; (2) take a FAIR data approach to agricultural data; and (3) include a comprehensive methodological segment on the practical implementation of FAIR data principles. As shown in Figure 1, out of 469 studies, 45 records remained after being refined. After deletion of duplicate studies, abstract reading and full-text skimming, there were 12 studies to be considered for full-text assessment.

Since the primary process of this systematic review has been to find studies that clearly describe the role of FAIR data in relation to agricultural performance, as well as evidence (i.e., a methodology approach) of the practical implementation of FAIR data principles, the selected papers were assigned a strength of evidence based on a grading system, described in Table 1:

Table 1. Strength of evidence based on a grading system.

| Grade             | Criteria                                                                 |
|-------------------|--------------------------------------------------------------------------|
| Substantiated     | include a clear role of FAIR data in relation to agricultural performance |
|                   | include a comprehensive methodological approach that demonstrates        |
|                   | FAIR data implementation processes                                        |
|                   | scientific, evidence based, empirical, quantitative and/or case study     |
| Partially substantiated | include a clear role of FAIR data in relation to agricultural performance |
|                   | include a comprehensive methodological approach that demonstrates        |
|                   | FAIR data implementation processes                                        |
|                   | scientific, evidence based, empirical, quantitative and/or case study     |
| Unsubstantiated   | studies discussing the role of FAIR data in other contexts and do not    |
|                   | qualify for the eligibility criteria                                      |

Finally, eight studies met the final appraisal criteria. The eight finalists were examined and assessed for their strength of evidence. Of the final eight studies, two (N = 2 or 25%) were graded as partially substantiated with the strength of evidence II, and six (N = 6 or 75%) were graded as substantiated with the strength of evidence I (Table 2). Average citation rate of the included studies is calculated as 7.
Table 2. Assessment of the selected studies based on their strength of evidence.

| #  | Citation a | Research Areas b | Methodology | Scientific | Empirical | Case Study | Descriptive | Evidence | Strength of Evidence d | Cited by e |
|----|------------|------------------|-------------|------------|-----------|------------|-------------|----------|------------------------|------------|
| 1  | Wijk et al. [36] | Sci. and Tech; Other topics | * | * | * | * | +++ | I | 03 |
| 2  | Harrison et al. [37] | Agri.; Genetics and Heredity | * | * | * | * | +++ | I | 14 |
| 3  | Dorich et al. [38] | Sci. and Tech; Env. Sci. & Ecology | * | * | * | * | +++ | I | 05 |
| 4  | Giuliani et al. [39] | Remote Sensing | * | * | * | * | +++ | I | 19 |
| 5  | Specka et al. [40] | Computer Science; Geology | * | * | * | * | +++ | I | 02 |
| 6  | Arnaud et al. [41] | Computer Science | * | * | * | * | +++ | I | 04 |
| 7  | Hackett et al. [42] | Plant Sciences | * | * | * | * | ++ | II | 01 |
| 8  | Singh et al. [43] | Plant Sciences | * | * | * | * | ++ | II | 11 |

* Shows study descriptions (with respect to methodology, types) a Authors; b Web of Science research areas of the included studies; c Substantiated (+++); Partially substantiated (++); Unsubstantiated (+); d Strength; e Citations.

In this literature review, the included studies were shortlisted based on the inclusion criteria that were confined to the studies: (1) classified under the six Web of Science categories: Agriculture Multidisciplinary, Agriculture Dairy Animal Science, Soil Science, Plant Sciences, Green Sustainable Science Technology and Computer Science Interdisciplinary Applications; (2) listed as article and/or review; (3) written in the English language; (4) comprise applied implications of FAIR data, its implementation and comprehensive methodological approach; (5) published with the top-ranked publishers, i.e., Elsevier, Springer Nature, MDPI, Wiley and Taylor & Francis, in all the journals nominated for the research area; and (6) published within last six (2016–present) years. There were three studies selected through Google Scholar, using the snowballing search technique, based on the reference lists of the studies obtained from Web of Science search. Notably, during our literature review process, we found three studies, i.e., Capalbo et al. [15], Weersink et al. [11] and Wolfert et al. [44], describing the importance of data/big data in the agricultural performance without relating to FAIR data principles. Our search also found four research papers, i.e., Koers et al. [45], Robinson et al. [46], Roitsch et al. [47], Ingram et al. [48] and Bahlo et al. [49], recommending the adoption of FAIR data principles to enhance agricultural performance, without explaining how or why. These papers do not meet the selections criteria; hence, they were excluded during screening and eligibility processes.

It is acknowledged that farming is a multifaceted business with no ‘best bet’ or ‘one method fits all’ solution due to the multiplicity of unforeseen factors. All the above criteria may introduce bias in this systematic review that may influence or contrast with the findings.
2.6. Presentation/Interpretation

According to the Web of Science research areas, two studies were categorised under ‘Plant Sciences’, one listed under ‘Remote Sensing’, and five studies relates to Agriculture, Science and Technology, Genetics & Heredity, Environmental Science, Ecology, and Geology categories. Moreover, the included eight studies have been published to different prominent journals (Table 3).

Table 3. Web of science categories and journals included.

| ID | Web of Science Categories | Journal                                      |
|----|----------------------------|-----------------------------------------------|
| 1  | Multidisciplinary Sciences | Scientific Data                               |
| 2  | Agri., Dairy & Animal Science; Genetics & Heredity | Animal Genetics                             |
| 3  | Green & Sustainable Science & Technology; Env. Sci. | Current Opinion in Env. Sustainability       |
| 4  | Remote Sensing              | Int’l Journal of Applied Earth Observation & Geoinformation |
| 5  | Computer Sci., Interdisciplinary Applications; Geosciences, Multidisciplinary Computer Sci., Artificial Intelligence; | Computers & Geosciences                      |
| 6  | Computer Sci., Information Systems; Computer Sci., Interdisciplinary Applications | Patterns                                     |
| 7  | Plant Sciences              | Applications in Plant Sciences                 |
| 8  | Plant Sciences              | Trends in Plant Science                        |

Although the trend is increasing (Figure 2), the minimal number of studies demonstrates very slow adoption of FAIR data principles reported in agricultural research publications, which may reflect on the slow uptake by the agricultural researchers and practitioners.

![Figure 2. Total number of the studies included and growing trendline.](image)

3. Findings

Since the publication of the FAIR data guidelines [23], there are only eight published studies found through this systematic literature review approach of their adoption in agricultural data. Though these studies confirm the significant role that FAIR data can have on agricultural operational and business performance, the actual implementation of FAIR data principles in the agricultural industry appears to be minimal. Figure 3 demonstrates indicators, sub-indicators of FAIR data, their impacts on sustainable agriculture performance and their assimilation with the sustainability framework [50,51]. The detailed analysis of the final selected papers with respect to their objectives, roles towards agricultural performance, scope and the sources or tools employed, is listed in Table 4.
Figure 3. Indicators, sub-indicators, their impacts on sustainable agri. performance, & assimilation with sustainability framework.

Table 4. Authors, objectives, and roles of FAIR data towards agricultural performance, scopes and sources or tools in the included studies.

| Author, Objective & Scope | FAIR Data Role Towards Agricultural Performance |
|---------------------------|------------------------------------------------|
| 1. Wijk et al. [36]       | Rural Household Multiple Indicator Survey (RHoMIS) aims to: |
| **Objective**: To list a well-coherent and interoperable dataset (following FAIR data principles) to help standardization of agricultural household surveys approach by collecting information on 758 variables, to better quantify more than 40 different indicators on farm and household characteristics, welfare, productivity and economic performance. |
| **Scope**: |
| - In line with the rural household multiple indicator survey (RHoMIS). |
| - A total of 21 countries in Central America, sub-Saharan Africa and Asia. |
| - promptly characterise a sequence of key indicators across the range of agricultural products and off farm activities, together with marketplace integration, nutrition, food security, poverty and greenhouse gas (GHG) emissions. |
| - measure on- and off-farm paths to food security, various diets and variations in poverty for rural smallholder farm households. |
| - attain sustainable development goals that require more improved sustainable food production and development of rural economies. |
| - better understand the relationships between farming practices, livelihood and the influences on farm performance and household welfare that help with developing targeted investment to advance in agronomic development. Trustworthy indicators at farm-household level of both agri. performance and household wellbeing help to better understand and model these relationships, and to review, revise and update the strategy and execution of intermediations involved officials through an extensive array of changing geographies and socio-economic features. |
| - manage propagation of survey tools and indicators heading to illogical, incoherent and un-interoperable datasets. |
2. Harrison et al. [37]

Objective: To develop high-quality and rich-supporting metadata (in line with FAIR data guidelines) to describe the project’s animals, specimens, cell cultures and experimental assays.

Scope:
- Functional annotation of animal genomes (FAANG) with an initial focus on farmed and companion animals.
- UK, USA.

The functional annotation of animal genomes (FAANG) metadata helps in:
- Creating sample and experiment metadata standards to enhance data recording.
- Standardising global vocabularies and or expressions by using ontologies.
- Utilising the wide-ranging livestock datasets produced outside of the project by employing less rigorous legacy guidelines.
- Developing authentication software to help the community in fulfilling the metadata standards and to input their data to the public archives, hence actively support the community.
- Providing a community data portal that classifies all sample and experimental datasets by using a single-focused user interface.
- Successfully dealing with the challenges faced by the consumers in creating infrastructure that mutually and efficiently coordinate genome-to-phenotype research activities.
- Maximising the usefulness and inter-comparison of assay data.
- Creating a powerful genome-to-phenotype resource and supports on-going developments in animal data standards as a whole, to support the community.

3. Dorich et al. [38]

Objective: To create the Global N₂O Database (following FAIR data principles) to serve as a repository for N₂O datasets to be publicly available data and for analytical advances.

Scope:
- Global N₂O Database.
- Nitrous oxide emissions.
- Global.

The Global N₂O Database deals with farming-oriented (nearly 20% of the total global) GHG emissions and is likely to improve evaluations level by improving annual N₂O estimates. The Global Nitrous oxide (N₂O) Database aims to:
- Improve Nitrous oxide (N₂O) emission estimates that are primarily obtained from agriculture and is an intoxicating greenhouse gas (GHG) that is roughly 300 times more intense than CO₂ and is the most hazardous ozone-draining material.
- Improve N₂O assessments, detection of hotspots and alleviation priority zones, and better interpretation of climate change feedbacks.
- Help with providing an opportunity for methods comparisons by collecting exact data from all practices.
- Provide value-added computation of annual emissions within a monitoring, reporting and verification process (MRV) scheme for enhanced policy making, for instance greenhouse gas exchanges and to reduce N₂O emissions.
- Help to moderate and better prepare for climate change by developing and verifying practical abatement strategies in the land use sector, as agriculture has a cause and effect’ correlation with the climate change.
4. Giuliani et al. [39]

**Objective:** To develop an innovative, scalable and flexible framework to monitor land degradation at various scales by using various components of the Global Earth observation system of systems (GEOSS) platform to leverage EO resources.

**Scope:**
- Monitoring land degradation at various scales system.
- Land degradation.
- Global.

Monitoring of land degradation at various (national, regional, global) scales system, in accordance with the UN SDG 15.3.1 framework, is a successful milestone that effectively embed science into the decision-making process. This system enables users to use EO-based resources more effectively and efficiently. It further aims to:

- reduce climate change and biodiversity losses, as well as ensuring food security and sufficient provision of ecosystem services, at the same time.
- produce multidisciplinary reliable knowledge on quantifiable objectives, at different scales, in order to proficiently support applied policymaking to ensure balanced functioning of the ecosystem.
- achieve sustainable development goals (SDG) 15.3.1, by following the data-information-knowledge pattern using the Trends.Earth model [52] and several data sources to produce the indicator.
- provide more flexible and scalable version of Trends.Earth to enhance decision-making processes and to scope of our planet, for instance natural resources, etc.
- strengthen respective regional capacities to effectively assess and map the degraded lands as per the UN sustainable development goals (SDGs).
- institute ‘data analytics’ podia that can potentially help nations to discover, access and use the necessary datasets to evaluate land degradation.

5. Specka et al. [40]

**Objective:** To frame new model, BonaRes metadata schema (following FAIR principles), by integrating the INSPIRE and DataCite metadata schemas.

**Scope:**
- Model based on two schemas, i.e., INSPIRE and DataCite metadata.
- The BonaRes metadata schema for geospatial soil-agricultural research.
- Global.

In compliance with the INSPIRE and DataCite metadata schemes and FAIR data principles, a modern research data management, BonaRes metadata:

- supports cross-portal metadata interoperability with other INSPIRE-compliant spatial data infrastructures (SDIs).
- increases the visibility and findability of researchers’ investigation, as the data can be assigned a digital object identifier (DOI), which is essential for data publications and data citations.
- allows targeted dataset queries and to better the discovery and reusability of research data.

6. Arnaud et al. [41]

**Objective:** To annotate multidisciplinary research data with the appropriate ontologies to stimulate the ontology content to fill the gap rather than developing completely new ontologies.

**Scope:**
- Consultative group on international agricultural research (CGIAR) methodology.
- Multidisciplinary ontologies—Focuses on ontologies in the fields of agronomy, crop, environment, plant, and socio-economic.
- Global.

Annotation of and integrative, multifaceted, versatile, associative research data with the most suitable ontologies aims to comply with the FAIR data principles, and to strengthen the findability of data for further reuse, hence adding to the return on investment (RoI) for information collection and storage. It further aims to:

- provide most appropriate ontologies with more applicable and reliable data control, including data stewardship, ownership and robust policy.
- enhance technological structure (essential softwares, connectivity and servers) that plays key role in arranging and classifying the tangible data structures, i.e., ontologies, taxonomies and structured vocabularies to finally form a supportable data management system.
- exclusively increase the influence of agronomic research and development (R&D) by following FAIR data guidelines to build community trust level.
Table 4. Cont.

| Author, Objective & Scope | FAIR Data Role Towards Agricultural Performance |
|---------------------------|-----------------------------------------------|
| 7. Hackett et al. [42]    | Global biodiversity information facility (GBIF) data sets: |
| **Objective**: To present biodiversity data workflow, i.e., data collection and curation, from multiple sources while abiding by FAIR data principles, to enable researchers, managers, and policymakers to address issues of global and future concern. |
| **Scope:** |
| - Biodiversity associated with globally vulnerable prairie fen wetlands. Plant diversity research and species-focused studies concerning the biology, ecology and behaviour of the federally endangered Poweshiek skipperling. |
| - Plant communities and the federally endangered Poweshiek skipperling (Oarisma poweshiek). |
| - Michigan prairie Fens, USA. |

8. Singh et al. [43]

**Objective:** To develop a single database for annotated plant stress images that supports FAIR principles (of accessibility and reusability) to propose an overarching strategy for utilizing ML techniques that methodically enables the application of plant stress phenotyping at multiple scales across different types of stresses, program goals, and environments.

**Scope:**
- Plant stress phenotyping ‘plant stress severity’ to encompass both biotic and abiotic stresses.
- Plant stress phenotyping, Maize plants.
- Iowa, USA.

Plant stress evaluations measure the visible signs and/or indications of stress and its progress on different plant units (e.g., leaf, stem, or roots) at the leaf, canopy, plot and field levels. A comprehensive database for annotated plant stress images, embedded with FAIR data principles, aims to:
- choose stress-resistant varieties and to develop better stress-management schemes.
- standardise visual evaluations and to utilise imaging techniques to better enhance the precision and trustworthiness of stress assessment in contrast with single-handed visual measurement.
- enhance machine learning (ML) approaches combined with image-based phenotyping to get up-to-date insights from highly organised, annotated (supported with explanations and/or comments), and high-dimensional (provided with the staggeringly higher number of dimensions) datasets across wide-ranging stresses and crops.
- concurrently build up the pace, precision, trustworthiness and scalability of stress phenotyping, and agility for highly varying program objectives; whereas innovative ML algorithms offer extended plant stress phenotyping techniques to deal with these challenges.
- advance mechanisation and accuracy of plant stress gravity evaluations that improve the proportion of genetic gain within crops, providing comprehensive management approaches. Notably, stress gravity evaluations in plants are important for appraising management strategies, plant breeding choice approaches and checking novel varieties for their capacity to alleviate crop damages. Additionally, efficiently quantify plant traits under different environmental circumstances with a stipulated precision and accuracy at various scales from organs to canopies.

Sustainable and enhanced agricultural production highly rely on the farming strategies, methods and decision making made by the stakeholders [5,6,53]. Accomplishment of FAIR data principles ensures data findability, accessibility, interoperability and reusability that extensively contributes better decision making towards both the operations and business management. Figure 3 demonstrates on how the sub-indicator/characteristics of FAIR data lead to sustainable agricultural performance indicators, and the way the sustainability indicators assimilate with the sustainable agricultural performance indicators, ultimately complementing the sustainability triple bottom line framework.
4. Discussion

The analysis of papers found through this systematic literature review clearly demonstrate that the provision of FAIR data can help improve agriculture operational and business performances. The multidimensional challenges in agriculture demand better decision making, which in turn relies on better access to data for wiser decision making. In this regard, a big responsibility lies on the shoulders of the stakeholders, i.e., farmers, farm managers, agronomists, service providers, researchers, etc. to make their data FAIR.

However, the systematic literature review also found only a few documented cases of practical examples that relate FAIR data to agricultural performance. While there may be many reasons for the low number of published case studies of adoption of FAIR data in agriculture, this review strongly indicates a need for social science research to explore what those reasons may be. Clearly, the stakeholders’ value propositions for the adoption of FAIR data in agriculture need to be better understood.

In accordance with Rogers’ [54], diffusion of innovation theory simply having knowledge of a new idea is not enough. By and large, informative individuals go through a persuasion stage to build either positive or negative attitudes towards the new idea, in accordance with innovation decision processes. Every novel idea carries a certain level of hesitation as the individual consumers of various personality types weigh up their appetite for risk and capacity for adopting change and so on. It is keenly realised that researchers should incorporate and comprehend the farmers’ social perspectives, seeing that the comprehension of attitude and social capacity are the key indicators towards viable agriculture and central to better explore the community attitudes and behaviour [55,56].

Firstly, agriculture practitioners’ attitude towards data and data sources in addition to their capability to employ the information are critical factors for the useful utilisation of those data [57,58]. Data are key in decision making [27], whereas the practitioners’ attitude towards data authenticity, data source, its genuineness and applicability are critical factors towards its usefulness [57,58]. Secondly, practitioners and stakeholders play their respective roles towards the betterment of agriculture. If agricultural data—as a key input in decision-making—are not findable, and accessible, they will be of no use even though they are well researched. Stakeholders’ data comprehension and knowledge exchange at all levels are important factors towards agricultural improvement [59]. If the data are not comprehendible to the end-users, they will not be used in accordance with the notion of the behavioural economics [60]. Thirdly, the available agriculture data are found disparate, disorganised and disarrayed [61]. The required datasets are either not available [62,63] or the consumers are provided with restricted and/or limited access [19].

Theory of reasoned action (TRA) uncovers that an individual’s objective and decision making depends upon the level of information or data they have [64]. Stakeholders’ knowledge building, including that of FAIR data, is essential. In this way, those with limited or no technical knowledge are asked to educate themselves with the emerging technology to better adapt with the modern advancements. Adoption of digital agriculture in general, and the adoption of FAIR data principles in particular, are facing several challenges. Politically, in Australia for instance, digital agriculture needs more proactive steps towards policy making, cooperation, administration and cross industry collaboration. Socially, digital literacy among all the key stakeholders can play a vital role to better comprehend value proposition of digital agriculture, and to bridge the trust gap between consumers, technology providers and data custodians. While device connectivity and the required data input are key prerequisites to make the technology work, the lack of mobile connectivity and internet telecommunications infrastructure, and limited or no access to the required data, curtail practitioners to use the technology at its best [65]. Economically, the Australian agriculture sector estimates an increase by 42% in its technology-oriented capital by 2030, whereas the farming sector needs to adopt the latest farming technologies and techniques to meet future production demands. Digital agriculture can add a gross value of AUD 20.3 billion [63], while better agricultural planning and appropriate farming
techniques have a potential to further supplement AUD 100 billion to the Australian economy by 2030 [66].

The provision of FAIR data is a well-documented constraint in modern agriculture [19,67,68], but there is little to no research on why FAIR data are not made available. If FAIR data are widely regarded as important, why is it not widely adopted? Is stakeholders' knowledge and perception towards FAIR data a barrier? To comprehend these questions, a FAIR data process flowchart has been devised as a guiding template to direct the social research. Using plain language, this flowchart attempts to step an agricultural practitioner through the requirements for FAIR data to ascertain where they perceive barriers and difficulties in adopting FAIR data processes.

The flowchart (Figure 4) intends to test the extents to which each of the required FAIR data components (findability, accessibility, interoperability, reusability) present barriers to adoption, whether these barriers are equally distributed across the four components and whether there are specific steps within each component that present higher barriers than others. By troubleshooting the specific barriers of each requirement, the value propositions of adopting FAIR data in agriculture can be better understood and addressed.

![Flowchart](image-url)

Figure 4. Cont.
5. Conclusions

The advent of digital agriculture has provided agricultural practitioners with access to a plethora of internet resources, sensors and applications for decision support in their operations and businesses. The FAIR data principles were designed to address the challenge of how to harness the increasing volume of disparate, but relevant data to improve decision making and enhance agricultural performance. This systematic literature review reinforces this hypothesis but finds very few published examples of how adopting the principles are related to agricultural performance, indicating that the value proposition is yet to be realised. To explore this tardiness will require social science research to find the specific barriers in making data findable, accessible, interoperable and reusable, so that benefits of digital agricultural can be more broadly gained.

Author Contributions: This project is supervised by P.D.; B.A. developed research framework, methodology and strategy; original draft preparation was prepared by B.A.; P.D. reviewed and edited the draft. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Centre for eResearch and Digital Innovation (CeRDI), Federation University and Food Agility Cooperative Research Centre, Australia, under Grant Project CERDI-19-136.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.
Acknowledgments: We thank the Centre for eResearch and Digital Innovation (CeRDI), Federation University, and Food Agility Cooperative Research Centre, Australia. We are grateful to the whole CeRDI research team for their support in completing this project.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. FAO, Food and Agriculture Organization of the United Nations. The State of Food Security and Nutrition in the World 2018; Food and Agriculture Organization of the United Nations: Rome, Italy, 2018.
2. Baligar, V.; Fageria, N.; He, Z. Nutrient use efficiency in plants. Commun. Soil Sci. Plant Anal. 2001, 32, 921–950. [CrossRef]
3. Capmourteres, V.; Adams, J.; Berg, A.; Fraser, E.; Swanton, C.; Anand, M. Precision conservation meets precision agriculture: A case study from southern Ontario. Agric. Syst. 2018, 167, 176–185. [CrossRef]
4. Rijswijk, K.; Klerkx, L.; Turner, J.A. Digitalisation in the New Zealand Agricultural Knowledge and Innovation System: Initial understandings and emerging organisational responses to digital agriculture. NJAS-Wagening. J. Life Sci. 2019, 90–91, 100313. [CrossRef]
5. Kusunose, Y.; Mahmood, R. Imperfect forecasts and decision making in agriculture. Agric. Syst. 2016, 146, 103–110. [CrossRef]
6. Acosta, M.; van Wessel, M.; van Bommel, S.; Ampaire, E.L.; Twyman, J.; Jassogne, L.; Feindt, P.H. What does it Mean to Make a ‘Digital’ Decision? Unpacking Intra-household Decision Making in Agriculture: Implications for Policy and Practice. J. Dev. Stud. 2020, 56, 1210–1229. [CrossRef]
7. Shepherd, M.; Turner, J.A.; Small, B.; Wheeler, D. Priorities for science to overcome hurdles thwarting the full promise of the ‘digital agriculture’ revolution. J. Sci. Food Agric. 2020, 100, 5083–5092. [CrossRef]
8. Birner, R.; Daum, T.; Pray, C. Who drives the digital revolution in agriculture? A review of supply-side trends, players and challenges. Appl. Econ. Perspect. Policy 2021, 43, 1260–1285. [CrossRef]
9. Whitelock, V. Business analytics and firm performance: Role of structured financial statement data. J. Bus. Anal. 2018, 1, 81–92. [CrossRef]
10. Faustian, M.; Theuven, L. Adoption of precision agriculture technologies by German crop farmers. Precis. Agric. 2017, 18, 701–716. [CrossRef]
11. Weersink, A.; Fraser, E.; Pannell, D.; Duncan, E.; Rotz, S. Opportunities and challenges for Big Data in agricultural and environmental analysis. Annu. Rev. Resour. Econ. 2018, 10, 19–37. [CrossRef]
12. Kitchen, N.; Sudduth, K.; Myers, D.; Massey, R.; Sadler, E.; Lerch, R.; Hummel, J.; Palm, H. Development of a conservation-oriented precision agriculture system: Crop production assessment and plan implementation. J. Soil Water Conserv. 2005, 60, 421–430.
13. Fountas, S.; Blackmore, S.; Ess, D.; Hawkins, S.; Blumhoff, G.; Lowenberg-Deboer, J.; Sorensen, C. Farmer experience with precision agriculture in Denmark and the US Eastern Corn Belt. Precis. Agric. 2005, 6, 121–141. [CrossRef]
14. Antle, J.M.; Basso, B.; Conant, R.T.; Godfray, H.C.J.; Jones, J.W.; Herrero, M.; Howitt, R.E.; Keating, B.A.; Munoz-Carpena, R.; Rosenzweig, C. Towards a new generation of agricultural system data, models and knowledge products: Design and improvement. Agric. Syst. 2017, 155, 255–268. [CrossRef]
15. Capalbo, S.M.; Antle, J.M.; Seavert, C. Next generation data systems and knowledge products to support agricultural producers and science-based policy decision making. Agric. Syst. 2017, 155, 191–199. [CrossRef]
16. Yost, M.; Kitchen, N.; Sudduth, K.; Sadler, E.; Drummond, S.; Volkmann, M. Long-term impact of a precision agriculture system on grain crop production. Precis. Agric. 2017, 18, 823–842. [CrossRef]
17. Delgado, J.A.; Khosla, R.; Mueller, T. Recent advances in precision (target) conservation. J. Soil Water Conserv. 2011, 66, 167A–170A. [CrossRef]
18. Berry, J.K.; Delgado, J.; Khosla, R.; Pierce, F. Precision conservation for environmental sustainability. J. Soil Water Conserv. 2003, 58, 332–339.
19. Barry, S.; Darnell, R.; Grundy, M.; Moore, A.; Robertson, M.; Brown, J.; Gaire, R.; George, A. Precision to Decision—Current and Future State of Agricultural Data for Digital Agriculture in Australia; CSIRO: Canberra, Australia, 2018.
20. Bronson, K.; Knezovic, I. Big Data in food and agriculture. Big Data Soc. 2016, 3, 2053951716648174. [CrossRef]
21. Leite, A.E.; De Castro, R.; Jabbour, C.J.C.; Batalha, M.O.; Govindan, K. Agricultural production and sustainable development in a Brazilian region (Southwest, São Paulo State): Motivations and barriers to adopting sustainable and ecologically friendly practices. Int. J. Sustain. Dev. World Ecol. 2014, 21, 422–429. [CrossRef]
22. Delgado, J.A.; Bausch, W. Potential use of precision conservation techniques to reduce nitrate leaching in irrigated crops. J. Soil Water Conserv. 2005, 60, 379–387.
23. 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. The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 2016, 3, 1–9. [CrossRef] [PubMed]
24. Boeckhout, M.; Zielhuis, G.A.; Bredenoord, A.L. The FAIR guiding principles for data stewardship: Fair enough? Eur. J. Hum. Genet. 2018, 26, 931–936. [CrossRef] [PubMed]
25. Wilkinson, M.D.; Verhough, R.; da Silva Santos, L.O.B.; Clark, T.; Swartz, M.A.; Kelpin, F.D.; Gray, A.J.; Schultz, E.A.; van Mulligen, E.M.; Ciccarese, P. Interoperability and FAIRness through a novel combination of Web technologies. PeerJ Comput. Sci. 2017, 3, e110. [CrossRef]

26. Foundation, G.F. GO FAIR Foundation. Available online: https://www.gofairfoundation.org/ (accessed on 1 November 2021).

27. Mons, B.; Neylon, C.; Velterop, J.; Dumontier, M.; da Silva Santos, L.O.B.; Wilkinson, M.D. Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud. Inf. Serv. 2017, 37, 49–56. [CrossRef]

28. Commons, A.R.D. FAIR Self Assessment Tool. Available online: https://ardc.edu.au/resources/working-with-data/fair-data/fair-self-assessment-tool/ (accessed on 20 August 2021).

29. Van der Windt, D.; Thomas, E.; Pope, D.P.; de Winter, A.F.; Macfarlane, G.J.; Bouter, L.M.; Silman, A.J. Occupational risk factors for shoulder pain: A systematic review. Occup. Environ. Med. 2000, 57, 433–442. [CrossRef]

30. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; Grp, P. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med. 2009, 6, e1000097. [CrossRef]

31. Koutsos, T.M.; Menexes, G.C.; Dordas, C.A. An efficient framework for conducting systematic literature reviews in agricultural sciences. Sci. Total Environ. 2019, 682, 106–117. [CrossRef]

32. Clarivate. Available online: https://clarivate.com/webofsciencegroup/solutions/web-of-science/ (accessed on 16 February 2022).

33. Google Scholar. Available online: https://scholar.google.com.au/ (accessed on 16 February 2022).

34. Wohlin, C. Guidelines for Snowballing in Systematic Literature Studies and a Replication in Software Engineering. In Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering, London, UK, 13–14 March 2014; pp. 1–10.

35. van der Windt, D.; Thomas, E.; Pope, D.P.; de Winter, A.F.; Macfarlane, G.J.; Bouter, L.M.; Silman, A.J. Occupational risk factors for shoulder pain: A systematic review. Occup. Environ. Med. 2000, 57, 433–442. [CrossRef]

36. Wijk, v.M.; Hammond, J.; Gorman, L.; Adams, S.; Ayantunde, A.; Baines, D.; Bolliger, A.; Bosire, C.; Carpena, P.; Chesterman, S. The Rural Household Multiple Indicator Survey, data from 13,310 farm households in 21 countries. Sci. Data 2020, 7, 1–9. [CrossRef]

37. Harrison, P.W.; Fan, J.; Richardson, D.; Clarke, L.; Zerbino, D.; Cochrane, G.; Archibald, A.L.; Schmidt, C.J.; Flicek, P. FAANG, establishing metadata standards, validation and best practices for the farm and companion animal community. Anim. Genet. 2018, 49, 520–526. [CrossRef]

38. 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. Improving N2O emission estimates with the global N2O database. Curr. Opin. Environ. Sustain. 2020, 47, 13–20. [CrossRef]

39. Giuliani, G.; Mazzetti, P.; Santoro, M.; Nativi, S.; Van Bemmelen, J.; Colangelo, G.; Lehmann, A. Knowledge generation using satellite earth observations to support sustainable development goals (SDG): A use case on Land degradation. Int. J. Appl. Earth Obs. Geoinf. 2020, 88, 102068. [CrossRef]

40. Specka, X.; Gärtnert, P.; Hofmann, C.; Svooboda, N.; Stecker, M.; Einspanner, U.; Senkler, K.; Zoarder, M.A.M.; Heinrich, U. The BonaRes metadata schema for geospatial soil-agricultural research data—Merging INSPIRE and DataCite metadata schemes. Comput. Geosci. 2019, 132, 33–41. [CrossRef]

41. Arnaud, E.; Laporte, M.-A.; Kim, S.; Aubert, C.; Leonelli, S.; Miro, B.; Cooper, L.; Jaiswal, P.; Kruseman, G.; Shrestha, R. The ontologies community of practice: A CGIAR initiative for big data in agrifood systems. Patterns 2020, 1, 100105. [CrossRef]

42. Hackett, R.A.; Belitz, M.W.; Gilbert, E.E.; Monfils, A.K. A data management workflow of biodiversity data from the field to data users. Appl. Plant Sci. 2019, 7, e1310. [CrossRef]

43. Singh, A.; Jones, S.; Ganapathysubramanian, B.; Sarkar, S.; Mueller, D.; Sandhu, K.; Nagasubramanian, K. Challenges and Opportunities in Machine-Augmented Plant Stress Phenotyping. Trends Plant Sci. 2021, 26, 56–69. [CrossRef]

44. Wolfert, S.; Ge, L.; Verdouw, C.; Bogaard, M.-J. Big data in smart farming—a review. Agric. Syst. 2017, 153, 69–80. [CrossRef]

45. Koers, H.; Bangert, D.; Hermans, E.; van Horik, R.; de Jong, M.; Mokrane, M. Recommendations for Services in a FAIR Data Ecosystem. Patterns 2020, 1, 100058. [CrossRef]

46. Robinson, N.J.; Dahlhaus, P.G.; Wong, M.; MacLeod, A.; Jones, D.; Nicholson, C. Testing the public–private soil data and information sharing model for sustainable soil management outcomes. Soil Use Manag. 2019, 35, 94–104. [CrossRef]

47. Roitsch, T.; Cabrera-Bosquet, L.; Fournier, A.; Ghamkhar, K.; Jiménez-Berni, J.; Pinto, F.; Ober, E.S. Review: New sensors and data-driven approaches—A path to next generation phenomics. Plant Sci. 2019, 282, 2–10. [CrossRef]

48. Ingram, J.; Maye, D.; Baily, C.; Barnes, A.; Bear, C.; Bell, M.; Cutress, D.; Davies, L.; de Boon, A.; Dinnie, L. What are the priority research questions for digital agriculture? Land Use Policy 2022, 114, 105962. [CrossRef]

49. Bahlo, C.; Dahlhaus, P.; Thompson, H.; Trotter, M. The role of interoperable data standards in precision livestock farming in extensive livestock systems: A review. Comput. Electron. Agric. 2019, 156, 459–466. [CrossRef]

50. Elkington, J. Towards the sustainable corporation: Win-win-win business strategies for sustainable development. Calif. Manag. Rev. 1994, 36, 90–100. [CrossRef]

51. Keeble, J.J.; Topiol, S.; Berkeley, S. Using indicators to measure sustainability performance at a corporate and project level. J. Bus. Ethics 2003, 44, 149–158. [CrossRef]

52. Trends.Earth. Available online: https://trends.earth/docs/en/ (accessed on 16 February 2022).

53. Tilman, D.; Balzer, C.; Hill, J.; Befort, B.L. Global food demand and the sustainable intensification of agriculture. Proc. Natl. Acad. Sci. USA 2011, 108, 20260–20264. [CrossRef]
54. Rogers, E.M. *Diffusion of Innovations*, 5th ed.; The Free Press: New York, NY, USA, 2003.
55. CCMA. *North Central Victoria Regional Sustainable Agriculture Strategy*; Central Catchment Management Authority: Huntly, Australia, 2016.
56. Tomer, M. How do we identify opportunities to apply new knowledge and improve conservation effectiveness? *J. Soil Water Conserv.* 2010, 65, 261–265. [CrossRef]
57. Milovanović, S. The role and potential of information technology in agricultural improvement. *Econ. Agric.* 2014, 61, 471–485. [CrossRef]
58. Zhang, Y.; Wang, L.; Duan, Y. Agricultural information dissemination using ICTs: A review and analysis of information dissemination models in China. *Inf. Process. Agric.* 2016, 3, 17–29. [CrossRef]
59. Webb, N.P.; Marshall, N.A.; Stringer, L.C.; Reed, M.S.; Chappell, A.; Herrick, J.E. Land degradation and climate change: Building climate resilience in agriculture. *Front. Ecol. Environ.* 2017, 15, 450–459. [CrossRef]
60. Teitelbaum, J.C.; Zeiler, K. *Research Handbook on Behavioral Law and Economics*; Edward Elgar Publishing: Northampton, MA, USA, 2018.
61. Leonard, E.; Rainbow, R.; Laurie, A.; Lamb, D.; Llewellyn, R.; Perrett, E.; Sanderson, J.; Skinner, A.; Stollery, T.; Wiseman, L. *Accelerating Precision Agriculture to Decision Agriculture: Enabling Digital Agriculture in Australia*; Cotton Research and Development Corporation: Narrabri, Australia, 2017.
62. Keogh, M.; Henry, M. *The Implications of Digital Agriculture and Big Data for Australian Agriculture*; Australian Farm Institute: Sydney, Australia, 2016; pp. 1–51. ISBN 978-1-921808-38-8.
63. Nolet, S. *Seeds of Success: Advancing Digital Agriculture from Point Solutions to Platforms*; United States Studies Centre at the University of Sydney: Sydney, Australia, 2018; pp. 20–25.
64. Ajzen, I.; Fishbein, M. *Understanding Attitudes and Predicting Social Behavior*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1980.
65. Wiseman, L.; Sanderson, J. *P2D Project: Accelerating Precision Agriculture to Decision Agriculture*; Griffith University: Gold coast, Australia, 2017; ISBN 978-0-6482462-1-3.
66. NFF. *Budget Roadmap Charts Course for $100 Billion in Farm Production by 2030*, 2018.
67. Allemang, D.; Teegarden, B. A global data ecosystem for agriculture and food. *FResearch* 2017, 6. [CrossRef]
68. CRDC, Cotton Research & Development Corporation. *A Big Data Reference Architecture for Digital Agriculture in Australia*; Cotton Research & Development Corporation: Narrabri, Australia, 2017; ISBN 978-0-6481983-0-7.