



AASF Data Ecosystem Project

Insights Report
V1.0

David Lemon and Laura Kostanski
March 2024

Australia's National Science Agency



Citation

Lemon D. and Kostanski L. (2024). AASF Data Ecosystem Design – Insights Report. CSIRO, Australia.

Copyright

© Commonwealth Scientific and Industrial Research Organisation 2024. To the extent permitted by law, all rights are reserved and no part of this publication covered by copyright may be reproduced or copied in any form or by any means except with the written permission of CSIRO.

Important disclaimer

CSIRO advises that the information contained in this publication comprises general statements based on scientific research. The reader is advised and needs to be aware that such information may be incomplete or unable to be used in any specific situation. No reliance or actions must therefore be made on that information without seeking prior expert professional, scientific and technical advice. To the extent permitted by law, CSIRO (including its employees and consultants) excludes all liability to any person for any consequences, including but not limited to all losses, damages, costs, expenses and any other compensation, arising directly or indirectly from using this publication (in part or in whole) and any information or material contained in it.

CSIRO is committed to providing web accessible content wherever possible. If you are having difficulties with accessing this document please contact csiro.au/contact.

About this document

This document is the first significant output from the **Discovery Phase** for the design of the data ecosystem to support the *Australian Agricultural Sustainability Framework (AASF)*. It provides insights captured from a desktop review, and from a series of interviews with key stakeholders with an interest in the AASF and agricultural sustainability reporting more generally.

Section 1 of the report describes the background to the AASF data ecosystem design project and the methods used to acquire the information reported.

Section 2 gives an overview of the data ecosystem concept - the benefits and limitations. Also included is a set of key insights structured around the lines of inquiry used for the activity, and a set of variables to be considered for the AASF data ecosystem.

Section 3 provides a set of recommended options for AASF data ecosystem vision, objectives, principles, use cases, and structures. These will be utilised as inputs to a co-design session of AASF data ecosystem stakeholders in March 2024.

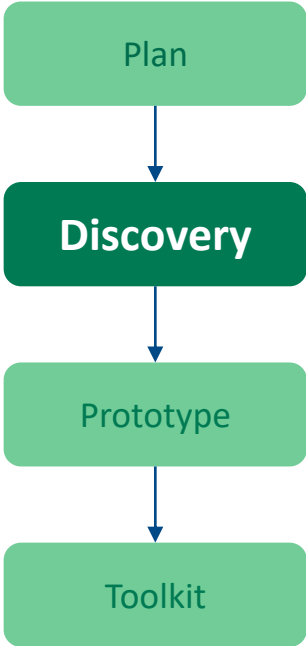


Table of Contents

Executive Summary	4
1.0 About this project	6
<i>Background</i>	7
<i>Methods & Lines of Enquiry</i>	8
<i>Participants & Other Inputs</i>	11
2.0 Findings	14
2.1 What is a data ecosystem?	16
<i>Components of a data ecosystem</i>	18
<i>Best-practice data ecosystem examples</i>	19
2.2 How do agriculture data ecosystems function in Australia?	20
<i>Australian agricultural sustainability reporting stakeholder landscape</i>	22
<i>Insights – behaviours, opportunities, blockers, risks, benefits, concerns</i>	23
2.3 What variables need to be accounted for?	30
<i>Stakeholder cohorts and data supply chains</i>	31
<i>Use cases</i>	32
3.0 Our Recommendations	34

Executive Summary

This Report

This report is the major output of the Discovery Phase of the AASF Data Ecosystem Project. It describes the findings from a series of activities undertaken to understand:

- *how agricultural sustainability data is currently collected, managed and shared,*
- *the perspectives, hopes, concerns of stakeholders towards the AASF Data Ecosystem*
- *Possible use cases that the AASF Data Ecosystem might support*

Information was gathered from:

1. Interviews with stakeholders (33 interviews, 58 interviewees – October 2023 to February 2024)
2. Review of published (online) content where available (ongoing)
3. Mini workshop with AASF Community of Practice members (November 2023)

Findings

Stakeholder Landscape

There are numerous cohorts of stakeholders with some interest in agricultural sustainability data. This interest might be: accessing data to support some form of reporting; providing data to others for reporting purposes; or enabling the sharing of data. Unfortunately, there is no coordination to this data sharing, no agreements on what is collected and shared, and little focus on who is bearing the burden of data collection activities.

Other Insights

- The current agriculture sustainability data ecosystem is anarchic in nature
- Different drivers are informing how organisations develop their data practices, frameworks and governance arrangements
- Data sharing with respect to agricultural sustainability is undertaken on an ad-hoc basis
- Different users will engage with and use the AASF and hence the AASF Data Ecosystem in different ways
- In general, stakeholders can see a range of benefits coming from the AASF Data Ecosystem
- The greatest opportunity of, and the greatest risk to, the data ecosystem is trust

Recommendations

It is recommended that the AASF Community next:

1. Agree on a Vision, Goals and Principles for the AASF Data Ecosystem so that the project can move forward with a clear understanding of what is intended to be achieved by it
2. Identify and prioritise the key set of use cases that the data ecosystem will initially support to guide the next phase of the current project

These will be addressed in the AASF Community of Practice face-to-face workshop (20/21 March 2024).

1.0 About this project

The Australian Agricultural Sustainability Framework

The Australian Agricultural Sustainability Framework (AASF) is the first country-specific framework to address sustainability from a whole-of-agriculture perspective. It is being developed by the National Farmer’s Federation (NFF) to provide benefits to farmers and the community by promoting best practice in agricultural sustainability and ensuring these efforts can be recognised by international markets and the community.

Development of the AASF is being undertaken using a phased approach with each phase consisting of a set of discrete tasks. Previous phases have:

- Undertaken foundational research
- Develop the initial framework of the AASF
- Explored incentives for use
- Undertaken an initial data analysis
- Benchmarked existing frameworks
- Explored legal and policy barriers
- Create a AASF Community of Practice (COP)

The current stage (Stage 2) is continuing to develop the AASF towards operationalisation. It includes tasks to:

- Complete a materiality assessment against the AASF’s 17 principles
- Develop a model report for the AASF
- Undertake a series of pilots using the framework
- Continue the activities of the AASF COP
- Design a data sharing ecosystem to support the Australian agricultural sustainability sector.

Further information on the framework and its various programs of work can be found at aasf.org.au

The AASF Data Ecosystem

Analysis conducted by CSIRO in 2021/22 found that, within the Australian agricultural sustainability sector, there are no uniform data sharing arrangements; no standards for data interoperability; and varying levels of governance capability maturity amongst data providers and users. It further found that publicly available data that might be of relevance to the AASF, have varying levels of sustainability, usability, and accessibility.^{1,2}

Discussions with the AASF Community of Practice (CoP) led to the realisation that the AASF needs to be supported with robust and logically connected mechanisms for governing data sharing activities. This would require the definition of use cases, appropriate institutional arrangements, and instilling a culture of trust and collaboration that enables the AASF community, data policies, and information systems to effectively function.

Building on previous work, the AASF Data Ecosystem project will:

- Help the primary AASF stakeholders understand the issues that the AASF data ecosystem will need to address
- Propose a governance model and set of institutional arrangements for the data ecosystem that have been co-designed, tested and endorsed by the AASF community
- Provide a clear path for implementing and supporting the AASF data ecosystem

These are intended to be achieved by December 2024.

Ultimately, as a result of this project, stakeholders will understand the mechanisms by which the data needed to support use of the AASF can be identified and assessed for suitability; and, have a robust approach to developing appropriate supply arrangements which support ongoing availability and use of this data.



1. Lemon, D; Tetreault Campbell, S; Whitten, S.(2022) Australian Agricultural Sustainability Framework - Data Analysis. National Farmers Federation: CSIRO. csiro:EP2022-0856.
2. Lemon D (2022) Australian Agricultural Sustainability Framework – Review of Publicly Available Data Sets. CSIRO, Australia

The AASF Data Ecosystem project has 4 distinct phases, as described below. This report includes information relevant to components of Phase 2 – Discovery.

1.Planning (Aug – Sept 2023)

Purpose

- *agree with primary stakeholders on timing, activities, outputs and outcomes for the project*

Activities

- *Intent workshop with primary stakeholder (August 2023)*
- *Seek ethical approval from CSIRO Ethics (Approval: 144-17)*

Outputs

- *Project Intent and Plan – including: project impact strategy, project timeline, research plan, initial lines of enquiry*

2.Discovery (Oct 2023 – Mar 2024)

Purpose

- *Engage with key stakeholders to understand current approaches, needs, expectations, concerns and desires for the AASF Data ecosystem*
- *Identify key use cases; determine requirements, agree on vision, objectives and principles*

Activities

- *one-on-one Interviews with key stakeholders (Oct 23 – Feb 24)*
- *review of published literature (academic, company websites, reports, etc.)*
- *minor workshop with key stakeholders (Nov 15 2023)*
- *major workshop (March 20/21 2024)*

Outputs

- *Discovery Report [**this report**]*
- *Major Workshop Report – describing agreed vision*

3. Prototyping (Apr – Dec 2024)

Purpose

- *Develop and test different institutional arrangements and system mechanics to support 5 priority use cases*
- *Provide recommendations to primary stakeholders on next steps*

Activities

- *Focus group activities to test individual prototypes*

Outputs

- *Recommendations report*

4.Toolkit (Jan – June 2025)

Purpose

- *Capture and publish knowledge gained from the project*

Activities

- *Paper and report writing*

Outputs

- *Data Ecosystem Design toolkit*

The primary activity during the discovery phase is to engage the key stakeholders in the AASF data ecosystem to understand:

- their current practices with respect to sustainability data collection, access and use
- their thoughts on how the AASF data ecosystem might operate
- any concerns they have about the data ecosystem

Data Collection

Data collection for the discovery phase was undertaken in three ways:

1. Interviews with stakeholders
2. Review of published (online) content
3. Mini workshop with AASF Community of Practice members

Interviews

Initial interview participants were selected in consultation with the project sponsor who then provided introductions for the project team.

After the initial interviews, the project team identified and approached specific candidates for interviews from the membership of the AASF Community of Practice. Selections were made so as to gain as broad a coverage of the different cohorts of stakeholders in the data ecosystem, as possible.

Each interview¹ as conducted online (using MS Teams) using open ended questions aligned with the project Lines of Enquiry (over). Interviews were recorded so that the project team could review them later to ensure insights were collected accurately. Only the project team had access to the recordings. In most cases, interviews involved two project team members.

Literature Review

The project team reviewed select published academic literature on data ecosystems as well as websites associated with the organisations interviewed. In some interviews, interviewees directed the project to other relevant online content which was also reviewed.

1. One interview was conducted in person at the interviewees place of work.

Data Analysis

Once the data collection phase was complete, the data analysed to look for themes and patterns, and key insights aligned with the project lines of enquiry (see over) extracted by each project team member individually. In some cases, these messages were direct quotes from a workshop or interview. In other cases, messages were summarised from the conversation.

Analysis of the collected insights was then undertaken as a team using an affinity mapping process, whereby insights about similar concepts or subject were collated. This was undertaken using the online tool Miro.

Once this mapping process was complete, it was possible to summarise the key themes within each line of enquiry. These results of this analysis are presented in this report.

Project Lines of Enquiry – Phase 1

In reviewing the project activity outcomes and broader project outcomes to be achieved, the project team considered what needs to be explored, understood and communicated through various project activities. A set of high-level Lines of Enquiry were developed for each of the project outcome areas. Each of these lines are being explored through relevant project activities. It should be noted that answers to some of these questions are already available from previous projects which the project team is building upon. Therefore, only a subset of these questions are being taken into stakeholder engagement activities, while others are posed to document a comprehensive literature review. The broader set of Lines of Enquiry for the whole-of-project can be found in the appendices.

Project Outcomes	Lines of Enquiry
Engagement from CoP and broader stakeholder community, in the development of insights, case studies and requirements backlog	<ul style="list-style-type: none"> Why CoP and AASF were developed, what their goals are, and where they are planned to go in the future. Understand the structure and governance processes of the CoP (now and planned) Understand how to effectively engage CoP members in co-design processes. Clarify how to avoid stakeholder fatigue. Understand the broader stakeholder landscape and who is a priority to engage and why. What are processes and how does the community implement them?
Clarity on what practices, frameworks and governance arrangements might be required and need to be tested with stakeholders	<ul style="list-style-type: none"> What existing practices, frameworks and governance arrangements are in place, and what, if any, analogous examples exist in Australia and globally. How stakeholders engage with AASF currently and/or would like to utilise AASF (or other mechanism) as a guide for data collection, usage and reporting in future. How data supply chains currently function for agricultural industries and the institutional arrangements which inform their design. What the opportunities, blockers and potential risks are for designing a data ecosystem which supports stakeholders to use AASF.

Project Outcomes	Lines of Enquiry
Understand the motivations for continuing with existing practices and the incentives which might build trust and enable change	<ul style="list-style-type: none"> Stakeholder behaviours (at cohort, industry and systems levels) which inform how they currently collect or source, assess and utilise data for sustainability reporting purposes Data supply chain touchpoints for stakeholders and the opportunities, problems and pain points they experience Perceived social, economic and environmental benefits of AASF data ecosystem and stakeholder (at cohort, industry and systems levels) perspectives of what the future needs to contain (ie what their visions are and what success looks like for AASF data ecosystem) Stakeholder concerns about implementation of AASF data ecosystem in the short versus long term and what changes are most likely to succeed at individual, cohort and systems levels What infrastructure changes are immediately required to build a trusted supply chain of data for indicators, and what is required to ensure trust in the data supply chain and incentivise stakeholder participation

Project Outcomes	Lines of Enquiry
Clear direction on priority requirements to be prototyped in Phase 2	<ul style="list-style-type: none"> The variations of stakeholder requirements, and the priority cohorts to design for Who the priority stakeholders are in both the short and long terms for development of the AASF data ecosystem What priority stakeholders require, and prioritise, for the design of AASF data supply chains that support application of indicators which measure trends in Australian agricultural sustainability practices What needs to be tested as a first-order priority for the AASF data ecosystem processes, governance and incentives, and what can be left for later and/or de-prioritised
Confidence in AASF data use cases to meet stakeholder requirements	<ul style="list-style-type: none"> What the use cases are for AASF data ecosystem, and which ones are a priority to design for in the immediate term Why certain use cases can or should be prioritised over others (ie the metrics by which “importance” is determined for use cases) What stakeholders anticipate the minimum data requirements, quality and accessibility levels will be for acceptance and/or buy in to the AASF data ecosystem

Participants

To explore the current context of agricultural sustainability reporting, a set of exploratory interviews were conducted with stakeholders between October and February 2023. In total 33 interviews of 58 individuals representing 31 organisations were conducted.

Candidate interviewees were initially selected in conversation with the client with introductions provided to interviewees. As the interview phase proceeded, the project team identified individuals and organisations they were keen to talk to and reached out directly. In these cases, the team had either: discovered that the new interviewee potentially had a unique perspective or experience of interest; or the team identified a gap in knowledge and was ensuring that a broad breadth of stakeholder cohorts were engaged in the research process.

Participants represented different cohorts of stakeholders with respect to agricultural supply chains and sustainability reporting. These cohorts included:

- Government (State and Federal, Policy and Operational)
- Agricultural Research and Development Corporations (RDCs)
- Existing commodity specific sustainability frameworks
- Agricultural industry service providers
- Finance
- Producers
- Processors
- Retailers
- Non-Government Organisations
- Research Organisations
- Agricultural Technology companies
- Agricultural Consultants



Other Inputs

The discovery phase has also collected data and insights from a range of other sources

Organisational and project websites and reports

Sustainability activities and associated reports are often published by individual organisations on their websites. These give a good insight into how the organisation views these activities, the drivers for them and the protocols and processes they are following. Where it was available, the project team reviewed this information for each organisation interviewed prior to interview.

Interviewees also often directed the project team to other organisations and projects of interest. In all cases, the websites for these activities were reviewed as a minimum

COP Workshop

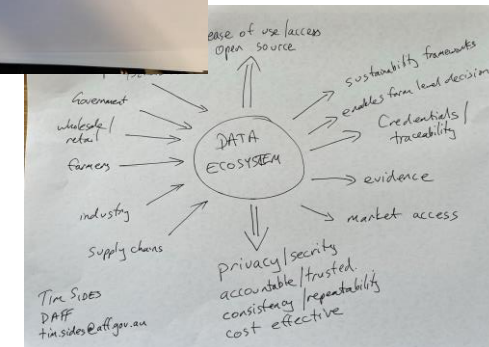
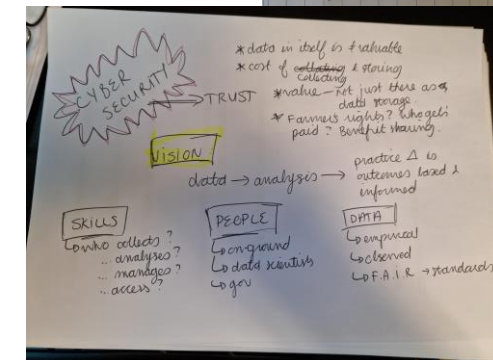
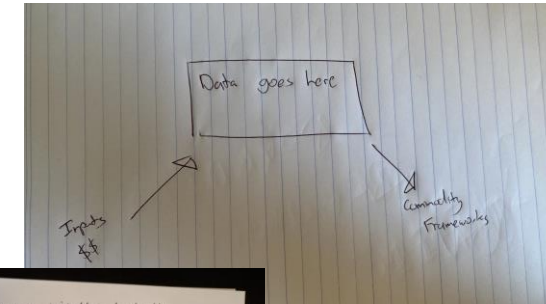
In November 2023, a mini-workshop was held as part of the November AASF Community of Practice meeting. During the workshop, participants were organised in small groups and asked to describe their vision for the data ecosystem, how the data ecosystem might support them, what the data ecosystem would need to include for them to engage in it, and who else should contribute to the data ecosystem. This information was recorded and used in the analysis phase.

Scientific literature and previous report

The nature and design of data ecosystems are a relative active area of academic research within a number of domains. A key input into the discovery phase has been recent research into the governance of data ecosystem and the pros and cons of different models currently being used in different sectors of society around the world.

Other

The project team has taken guidance and advice from other members of the AASF project team whose experience and knowledge, in particular of the Australian agriculture sector have been invaluable



Limitations

As with all research it is acknowledged that there are limitations to the approaches taken, including:

1. The project team only interviewed those willing and available to participate in an interview between October 2023 and December 2023. This was a very busy time period for many in the agriculture and sustainability domains – with UNCOP and other international and national events taking place. Ultimately, more invitations for interview were sent than interviews undertaken. However it is noted that over 85% of those approached agreed to participate, which is an exceptional return rate.
2. Invitations were only sent to existing members of the AASF Community of Practice and/or those introduced by the client or other participants. Results of interviews are therefore biased towards the interests of those who were already engaged, or had expressed an interest in, the work of the AASF.
3. Interviews were limited in number (33) although, in discussion with the client, we are comfortable the cohort was representative of the agriculture sector and no major groups were missed. In all cases, interviewees were very engaged in the conversation and more information could have been gathered had time permitted.
4. The team is aware that various perspectives have not been sought and included in this stage of the research. For example, the perspectives of Traditional Owners, Standards Groups, International Peers, and others are not yet included.
5. Interviews were exploratory in nature and not intended to be comprehensive nor to be used for quantitative scientific analysis. The purpose of this research was not to provide definitive insights into cohort behaviours and patterns.

2.0 Findings

Overview

This findings section is divided in three components, each of which provides analysis and insights into the information provided from the literature, interview and workshop participants. The purpose of this section is twofold: firstly, to provide insight into the various thoughts, behaviours and processes currently at play in the Australian agricultural data ecosystems; and, to provide an evidence base for the recommendations contained in the final part of the report. The three components of this section of the report are:

2.1 What is a data ecosystem?

This component outlines various definitions of what a data ecosystem can be and the value it can provide. The section describes in detail the:

- *Benefits of data ecosystems*
- *Limitations of data ecosystems*
- *Technical Components*
- *Actors*
- *Examples of Data Ecosystems*

2.2 How do agriculture data ecosystems function in Australia?

This section presents the analysis of the information gathered from literature review, mini workshop and, most importantly, stakeholder interviews. This information is presented in two ways:

- *First, a stakeholder landscape analysis, describing the various cohorts of participants in the Australian agricultural sustainability landscape and the relationships between them.*
- *Secondly, a set of insights observed with respect to the Lines of Enquiry described in Section 1. The insights give an understanding of the various motivations, behaviours, incentives and requirements that stakeholders in the AASF Data Ecosystem might have. For each subject a high-level or key insight is described followed by sub-insights related to the subject. Quotes are provided from interview participants (de-identified).*

2.3 What variables need to be accounted for in establishing an AASF Data Ecosystem?

This component outlines the additional variables which need to be considered when designing a data ecosystem for AASF. These include:

- *Stakeholder cohorts and data supply chains*
- *Use cases*
- *Other considerations*

2.1

What is a data ecosystem?

A data ecosystem is an interconnected, dynamic system comprising technical components and actors that collectively engage in the production, management, exchange, and consumption of data. They function through the seamless interaction between data owners, users, and producers within a framework designed to facilitate data availability, reliability, and integrity for specific or broad purposes.

Unlike isolated data management practices, a data ecosystem emphasises collaboration, sharing, and governance, enabling data to flow efficiently between different entities for mutual benefit.

Benefits

The primary benefits of data ecosystems over isolated data systems are:

- *Economies of scale* – Through the application of standards and common supporting infrastructure, data ecosystems can reduce the cost of:
 - data collection – through the reduction of duplicative data collection and increase in data reuse/repurposing
 - data discovery – through shared mechanisms to register and search for existing data collections
 - data access – through use of common, less restrictive, data licences and application of standards-based data access protocols; and
 - data use – through the application of standard vocabularies and conceptual models to reduce data processing costs and the risk of data misinterpretation
- *Enhance collaboration* – through bringing those within the ecosystem together to address problems that are more difficult to solve individually. This may include collaboration on data collection activities of common interest or the development of common methods or tools.
- *Enable new opportunities* – This can include the generation of new insights through the ability to connect previously disparate data sets together, through to monetising various aspects of the ecosystem to return value to contributors.

Limitations

While data ecosystems offer significant advantages and opportunities, they also face limitations, including:

- *Data Quality and Consistency* – Ensuring high-quality, accurate data across diverse sources remains a challenge.
- *Privacy and Security Concerns* – Balancing data accessibility with the need to protect sensitive information and comply with privacy regulations.
- *Technical Complexity* – Integrating disparate systems and maintaining interoperability demands substantial technical expertise and resources.
- *Governance and Trust* – Establishing governance structures that ensure fair data sharing and use, while building trust among ecosystem participants.

‘the complex environment of co-dependent networks and actors that contribute to data collection, transfer and use’¹

- ‘A successful ecosystem balances two priorities:*
- *Building economies of scale ...*
 - *Cultivating a collaboration network ...’²*

1. Oliveira, Marcelo Iury S.; Lóscio, Bernadette Farias (2018-05-30). "What is a data ecosystem?". *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age*. New York, NY, USA: Association for Computing Machinery. pp. 1–9. doi:10.1145/3209281.3209335

2. Abdulla, Ahmed (March 8, 2021). "Data ecosystems made simple". *McKinsey Digital*

Components of a Data Ecosystem

Typically, a data ecosystem will consist of a range of technical and social components. These may include some, or all, of the following:

Technical Components

- *Data Infrastructure*: This includes physical and virtual data storage, databases, registries, cloud services, and networking facilities that enable the storage, discovery, retrieval, and processing of data.
- *Data Management Tools*: Software applications and platforms that support the organisation, quality control, analysis, and visualisation of data.
- *Data Security Management*: Technologies and protocols that ensure data privacy, permissioning, access, and sovereignty
- *Integration and Interoperability Mechanisms*: Technologies and protocols that ensure data can be shared and used across different systems, platforms, and organizations.

Actors

- *Data Owners*: Individuals or organisations that hold rights to data sets. These are responsible for the governance and often the privacy and security of the data.
- *Data Users*: Entities that consume data for analysis, decision-making, or as input for further processing. Users can range from businesses leveraging data for strategic insights to researchers analysing data for scientific purposes.
- *Data Producers*: Actors involved in generating data, which can include sensors and IoT devices, individuals participating in data collection, or organizations conducting research.
- *Data Custodians*: Organizations or individuals responsible for the safe custody, transport, and storage of data. They ensure data integrity, availability, and protection.
- *Data Scientists and Analysts*: Specialists who analyse data to derive insights, trends, and patterns, contributing to decision-making processes across sectors.
- *Policy Makers and Regulators*: Authorities that establish data governance frameworks, privacy standards, and usage policies to ensure ethical data handling and protection of individual rights.
- *Data Advocates and Consumer Rights Groups*: Entities that champion the ethical use of data, transparency in data practices, and protection of personal data rights.
- *Technology Providers*: Companies and organizations that develop and offer technologies and platforms facilitating data collection, analysis, and sharing.
- *Researchers and Academics*: Individuals or groups engaged in scholarly research that relies on data to explore theories, validate findings, and contribute to knowledge across disciplines.
- *Community and Non-profit Organizations*: Grassroots and non-commercial entities that use data to support social causes, community development, and advocacy campaigns.
- *Data Brokers*: Companies that collect, aggregate, and sell data to third parties, often playing a controversial role in the data ecosystem regarding privacy and ethics.
- *Standards Bodies and Consortiums*: Organizations that develop data standards and protocols to ensure interoperability and compatibility across systems and platforms.
- *Ethics Committees and Boards*: Groups that review and provide guidance on ethical considerations in data projects, ensuring that data use respects societal norms and values.

Note: No data ecosystem is likely to include all these components and actors. Rather, it will include those components necessary to achieve its vision. A key activity in designing a data ecosystem therefore is to determine which of these components and actors will be necessary.

Data Ecosystem Examples

Best practice data ecosystems exemplify effective integration of technical and social components, balanced governance, and a clear focus on achieving specific objectives while ensuring data privacy, security, and ethical use.

My Health Record

Operated by the Australian Digital Health Agency (digitalhealth.gov.au) the My Health Record system enables the permissioned sharing of individuals' summarised health data such as immunisations, pathology, imaging reports, prescriptions and much more between health professionals treating that individual. Its purpose is to improve patient outcomes no matter where they present for treatment. Key components of the system include: standards for data sharing and access, strong data security and chains of trust, identity management for patients and health providers, and strong governance.

The My Health Record is enabled by federal legislation (My Health Records Act 2012)

Smart Cities Initiatives

Around the world there are a number of urban data ecosystems that collect and analyse data from various sources to improve planning, sustainability and resident services. Information collected can include data from real time sensors, public services, and citizen feedback. Key aspects of these initiatives with respect to data is a strong focus on individual privacy and security.^{1,2}

International Air Transport

The International Air Transport Association (IATA) has over 320 members. Among a range of advocacy and other activities, IATA develops and maintains standards for sharing data about air transport related activities thus enabling interoperability across the entire international airline industry. These standards include industry agreed vocabularies, ontologies, APIs, identity management and more. Governance of these standards falls

Open Data Platforms

As the name suggests, the purpose of open data platforms is to provide open (free) access to a range of data sets. These programs are generally supported by governments (e.g. European Data Portal - <https://data.europa.eu/en>). Within the ecosystem, the key focus areas are data discovery, appropriate data licencing and in some cases interoperability standards.

1. Box, P., Lee, A., Smith, G., Mackenzie, A., Sanderson, T., Reeson, A., ... & Fleet, R. (2020). Data platforms for smart cities: a landscape scan and recommendations for smart city practice.
2. Lee, A., Mackenzie, A., Smith, G. J., & Box, P. (2020). Mapping platform urbanism: Charting the nuance of the platform pivot. *Urban Planning*, 5(1), 116-128.

2.2

How do agriculture data ecosystems function in Australia?

Overview

This section presents the analysis of the information gathered from literature review, mini workshop and, most importantly, stakeholder interviews. This information is presented in two ways:

First, a stakeholder landscape analysis, describing the various cohorts of participants in the Australian agricultural sustainability landscape and the relationships between them.

Secondly, a set of insights observed with respect to the Lines of Enquiry described in Section 1. The insights give an understanding of the various motivations, behaviours, incentives and requirements that stakeholders in the AASF Data Ecosystem might have. Insights are structured around:

- *Existing practices, frameworks & governance arrangements*
- *How stakeholders engage with AASF currently, and/or would like to use it*
- *How data supply chains currently function for agricultural industries and the institutional arrangements which inform their design*
- *Stakeholder behaviours, existing projects, legislation, rules (institutional arrangements, perceptions of standards, and perceptions of others and their roles)*
- *Opportunities, blockers and potential risks*
- *Perceived social, economic and environmental benefits*
- *Stakeholder concerns and requirements*

For each subject a high-level or key insight is described followed by sub-insights related to the subject. Quotes are provided from interview participants (de-identified).



Image of Miro board containing data generated during participant interviews and literature reviews. Analysis of this information resulted in the insights presented in this section of the report.

Stakeholder Cohorts

As noted in Section 1, the project team engaged with a broad range of stakeholders across multiple areas of the agricultural sustainability community. It was found that the community consists of a range of very different stakeholders with an equally broad range of needs, roles and responsibilities.

To better understand these different needs and determine if there are patterns of need with different types of stakeholders, the project team spent time exploring ways to classify stakeholders into cohorts. There are numerous ways this can be done. For example: whether the organisation is exposed to international markets or not; the organisation's role in agriculture; the scale of the organisation (small, medium, large). The team found two types of classification into which interviewees could be organised that demonstrated clear patterns of behaviour. These are:

Role in data supply chain

This cohort consists of:

- **Data producers** – Undertake activities from which they or others wish to capture data for sustainability purposes. The most obvious members of this cohort are farmers, but this cohort also includes processors, transporters and others
- **Tool providers** – Those that provide tools/technologies that may help with the measurement/collection/capture/sharing of sustainability related data. This cohort includes ag tech companies, researchers, ag services companies and others
- **Data users** – Those who need sustainability related data for their own purposes. This cohort contains a wide variety of organisations and has multiple purposes for the data. It includes retailers, FMCGs, financiers, RDCs, researchers, and many others. At present their primary need for sustainability related data is to support their own sustainability reporting requirements.

The team found that there are significant power imbalances between the cohorts in this categorisation. That is, data producers often find themselves at a disadvantage with respect to those needing data due to differences in scale and supply chain dependencies.

AASF Use Case

Another way the cohorts can be characterised is by the use cases they are interested in with respect to the AASF. These are:

- Those interested in **national scale reporting**, usually by a single commodity. For these organisations, the data needed reflects an entire commodity or region and so the AASF presents an opportunity to coordinate around collection of data sets of common interest. This cohort will include existing commodity sustainability frameworks, government, and others.
- Those interested in **supply chain/portfolio reporting**. Here the organisation is interested in the AASF to provide guidance on principles, criteria and indicators for their sustainability purposes, but data collection remains specific to the supply chain or portfolio. National scale data is less useful to this cohort apart from for benchmarking purposes. This cohort will include organisations that need to produce their own sustainability reports such as retailers, banks, large corporates, etc.

When considering this characterisation of stakeholders, a key difference is the nature of the data they need to access (national scale vs supply chain/portfolio). The ways in which the data ecosystem can support these use cases, whilst not mutually exclusive will differ. These are explored further in this section.

Existing practices, frameworks & governance arrangements

Different drivers are informing how organisations develop their data practices, frameworks and governance arrangements

There is no single driver that is influencing all agriculturally focussed organisations in their activities around sustainability. For some the impetus is coming from international clients and markets, for others domestic issues are paramount. Some are focussed purely on emissions while others see different challenges as a higher priority. This impacts the data organisations are collecting, the way they collect this data, where they collect it from, and any associated governance arrangements.

There is a significant difference in focus between those parts of the industry that are trade exposed and those which primarily operate domestically

Most trade exposed industries are very aware of the existing and emerging international sustainability frameworks and responding to one or more of these. Domestically focussed industries are less focussed directly on these frameworks and are more likely to be driven by customer needs or the desire to 'do the right thing'.

Those organisations that are exposed internationally are looking to a range of different standards (SBTi, TNFD, GRI, TCFD, etc).

This results in a range of different approaches, indicators, and languages being used that are not necessarily interoperable. This, in turn causes greater work for those that are required to provide data to support multiple frameworks.

An important driver is the availability (or not) of data to support reporting

The availability of data to support reporting against key metrics is a key driver although this driver can take different pathways. For some, data availability will be a key determinant in what indicators are used for reporting. For others, while indicator choice may be independent of the availability of data, once these indicators are agreed, significant effort is undertaken to find and/or develop needed data.

There seems to be a disconnect between government and industry

There does not appear to be any regular ongoing dialogue between government and industry about data collection. Rather, where discussions occur, they appear to be ad-hoc and, in some cases, unresolved. The result is lack of understanding from industry on what data government collects and why, and an equivalent lack of understanding from government of what industry needs.

Many are looking for exemplar approaches that they might follow.

There are many in the industry that are looking for guidance on what data they need to collect, how they should collect it, and how it might be used. In the absence of industry standards, they are looking to research, other industries, peers and even internationally. This will potentially result in numerous approaches being adopted.

Most existing sustainability frameworks focus on a single commodity with minimal ability to account for multi-commodity farms

At present, most reporting is done on a per commodity or a per supply chain basis. Only a very small number of data collection activities consider that many farms produce multiple commodities and/or sell into more than one supply chain. Not only does this represent a burden to data producers, there is a risk of double accounting.

"for me, it's the three of SBTi, TNFD and GRI that inform my work here..."

"we've made a number of commitments, and SBTi is one we've committed to and we've set a scope 3 goal of ensuring our suppliers have a baseline and target for emission reduction by 2027...."

"when we start looking at biodiversity for example, that's a whole other issue...traceability, animal welfare, carbon, biodiversity, are all part of our strategy pillars....I think we have the first two covered, we're working on carbon and for biodiversity I don't think anyone can provide us with an answer on that one yet...and we're still working out as a company what we need to report on and what metrics we'll have and what data we'll need..."

A very strong thread in our conversations is that data needs are universal, and no-one has an answer...sorry we don't have a silver bullet....all the people are of the same opinion, there is a lot of data out there and we need a process to handle it - it is about quality not quantity"

"we do look for gaps (in our data), we're big on looking for gaps, because we come across so many"

"the regenerative organic standard covers elements similar to AASF, not the conventional export standard"

"we are obligated to collect or process certain types of data due to international agreements and mandates..."

How stakeholders engage with AASF currently, and/or would like to use it

Different users will engage with and use the AASF and hence the AASF Data Ecosystem in different ways

There is a great deal of interest in the AASF and many are looking forward to being able to use it for their purposes. These purposes vary between direct use of the full framework for reporting purposes; adopting elements of it to include in other reporting frameworks; and guidance for development of their own reporting. The latter are likely to be company reporting along supply chains or investment portfolios.

The data ecosystem will need to support a range of use cases

National-scale, whole of agriculture reporting is just one use-case for the AASF. The data ecosystem will need to support this and other use cases to be sustainable. Other identified use cases include: single commodity reporting; regional reporting; supply-chain/portfolio reporting; benchmarking; and more. One of the more important use cases for the data ecosystem will be working as a community to address challenges that would be inappropriate or unachievable for any single organisation to tackle. This includes the development and management of standards as well as addressing key data gaps.

A high priority for the data ecosystem is a standard set of AASF metrics that can be adopted by users

Many stakeholders are looking to the AASF to provide a set of standard metrics/indicators that they can adopt for their own purposes. For many, this is the highest priority for the AASF Data Ecosystem as it will then be used to guide the development of their own sustainability framework. These standards are also an important enabler of interoperability across the ecosystem.

There is no consensus on the form of the ecosystem

When asked about what form the data ecosystem should take, stakeholders identified a range of options. These varied from a single data warehouse containing all data, to a set of interoperability standards, to a set of guidance on the management and sharing of sustainability data and many other options. The key message here is that the community is engaged and looking for solutions but there is no consensus on the form.

While there are a number of identified use-cases, further applications for the data ecosystem will emerge over time

While stakeholders already have their own priority use cases for the AASF Data Ecosystem, it is likely that new use cases for the ecosystem will emerge over time. To support these use cases, the ecosystem (both technology and institutional arrangements), will need to be flexible enough to be able to evolve.

There is a need to distinguish between the governance of the AASF and governance of the data ecosystem.

The AASF itself (as a framework and community) and the AASF Data Ecosystem need to be differentiated as they will likely have overlapping but not the same sets of stakeholders and use cases. Therefore, governance of the AASF Data Ecosystem needs to be considered both within the context of broader AASF activities, but also with its own specific and nuanced arrangements developed to support its purposes.

"I saw the potential for AASF and some of the categories....as being a useful point to incorporate into the taxonomy here...the problem with the AASF is that it's at principles level, not at the 'what does that mean' level, let alone the data to underpin that..."

"we are using 'metric' as an umbrella term to cover measures, indicators and indices"

"that's the question - who will be the custodian of these vocabularies which will be really critical to Australia as a nation - especially when there's no consistency...."

"AASF have the power to potentially lobby for datasets through the government to fund and do this and collect on our behalf and have a new data source for us to use in our reporting, and we could give data back"

"There's probably a clearer benefit for the data collection side of AASF than the actual framework"

"....we recognise there is diversification on farms across multiple commodities....if there is a way for some of that data to be centrally collected and divvied out to the individual commodities, then that is something we'd have an interest in"

"this will be lots of different vocabularies defined by different people for their own needs....how much does department of agriculture want to play in the governance of this vocabulary?"

"...for banks to report against the taxonomy- they need a definition, and they're all doing it differently....at the moment the international markets are saying use the EU definitions, but they miss a lot of the nuance of the Australian context"

"I'm not surprised the audience broadened - so even more important for us to determine audience priorities - otherwise will be pulled in many ways, diffusing the message..."

How data supply chains currently function

The current agriculture sustainability data ecosystem is anarchic in nature

The current agricultural sustainability data ecosystem in Australia lacks any form of coordination or organisation. As a result, individual organisations, whether they be commodity specific sustainability frameworks, supply chain participants or others with a need to access and use sustainability data are all 'doing their own thing' with respect to data collection and management. The result is significant cost across the ecosystem due to duplicative data collection, lack of consistency around what is being asked for, and an increasing burden on those being asked to provide data with a subsequent degradation in data quality.

Multiple approaches to data collection

Due to the lack of coordination, those collecting data are taking very different approaches to their data collection activities. This includes what they collect (different metrics), how they collect it (surveys, direct access) and how often it is collected (regular, irregular). In some cases, the same information is being collected multiple times in different ways from the same provider (for example multi-commodity farms selling into different supply chains may be asked to provide sustainability related data into each of these supply chains as well as to financial institutions) This is inefficient and places a significant burden on those being asked to provide the data. This is challenge is amplified by the fact that data providers are given little in the way of positive incentives to provide data.

Multiple approaches to data management and use

There are also no consistent approaches to data management across the industry. As a result, there are different and disconnected governance arrangements in place, restrictive and inefficient licencing practices and a range of different terminologies/vocabularies being used. This results in a lack of interoperability of data across the sector further increasing costs.

Ag tech sector lacks maturity when it comes to interoperability and reuse of data

Australia's agricultural technology (agtech) sector primarily consists of a large number of very small technology companies focussed on developing their space in a very crowded market. The agtechfinder website (an initiative of the Food Agility CRC) boasts that they list hundreds of ag tech companies. In many cases, the technologies these companies produce are often narrow in focus resulting in the need for producers to invest in multiple technologies and platforms to support their businesses. This is a known frustration for producers. There is also little coordination across the ag tech sector and, subsequently, little activity around supporting interoperability between these many technologies. In some cases, there are still actors within the ag tech sector of the opinion that data that they source from their clients is their proprietary property and that it has some inherent value that they plan to exploit.

Lack of coordination (leadership) for data and data related activities across the sector

There does not appear to be an obvious forum within the sector to lead the development and management of interoperability standards or standard metrics. There exist projects to develop cross sector capabilities such as standards. However, these project all end, and it is not clear which organisations will take carriage of the outputs of the projects

"the data we've collected has grown like topsy...it's like we've got stuff we sent up to space in the 70s....and if you could start again, you would" (ie need to design for technical evolution too)

[Farmers say]... "We don't want 25 apps on our phones that we have to use"

"we need a way to reduce the burden on producers to report. And for this to be easy and I can send to whoever I need to send it to"

"the review said that existing information published about agriculture isn't meeting needs...and there were issues there about reporting by non-farmers, not getting data out fast enough, not getting enough detail"

"my department doesn't have any primary data, it all comes from ecological consultants and jurisdictions"

"Generally, what we've found in the farming community is that they are very smart, very switched on, very innovative people. However, not very trusting"

"we are developing consistent language and metrics across sectors and land use types...to get natural capital measures consistent here and with international frameworks"

"we meet every quarter to do ongoing materiality scans, to make sure we're on top of stuff and nothing is missing...every six months we have a stakeholder representative group - a select group of people....everyone who has an impact...and we say what we're doing, what we're missing and ask what else they'd like from us so we can have an ongoing dialogue"

How stakeholders currently behave with data

Data sharing within the agricultural sustainability sector is undertaken on an ad-hoc basis

Currently, the primary approaches to sharing data within the agricultural sustainability sector is to employ a set of 'point-in-time' solutions. That is, solutions that support a specific need at a point in time rather than general solutions. This applies to both the data being shared and the methods by which it is shared. The result is a complicated array of, often, incompatible data and applications, addressing very specific needs, applications, and resultant costs for both data producers and users.

There are many initiatives within agriculture seeking to solve the problem of data sharing along supply chains

There are activities being undertaken to address the challenge of sharing sustainability data between organisations. Some of these are within individual supply chains (cf. Tassie to Townsville project) whereby sustainability related credentials are being passed down a supply chain to support sustainability reporting for one organisation. Others have been initiated by government and are being led by the research sector (Australian Agricultural Traceability Standards Working Group). There are also technology companies providing services to enable connecting of data to users. An observation of some of these projects is that they are often unique to individual supply chains, and few seem to consider the burden on data producers/owners, particularly if they produce multiple commodities. Very few of these projects seem to consider a value return to data producers¹.

Data collection methods vary, some of increasingly lower quality and value

A data collection method commonly used by government and various industry groups is the survey. The project found that many organisations conduct surveys of their stakeholders and, in some cases, more than one survey per year. However, some interviewees reported that the quality of the responses to these surveys is declining. For one interviewee, the quality has declined to a point that results were unpublishable and hence they have discontinued their annual farm survey program. Reasons given for this decline primarily relate to those being surveyed being 'over surveyed'.

There are many assumptions about the roles of key stakeholders within the sector

A common misunderstanding from a number of those interviewed is with respect to the drivers for and processes by which government makes decisions about the data it collects. For example, many stakeholders believe that government data collection activities can be influenced by industry simply asking for change. Similarly, there are misconceptions of the role and capabilities of certain government research organisations. These misunderstandings lead to assumptions being made about how a future data ecosystem might operate and the roles particular cohorts will take. An early activity in implementing the data ecosystem will be to address these myths and misconceptions.

Numerous governance patterns emerging within parts of the industry or along supply chains

As noted above, there are numerous activities, often along individual supply chains or within single commodities, seeking to address data sharing issues and needs. With each of these activities, different, often inwardly focussed, governance processes and institutional arrangements are being implemented. This can lead to challenges such as restrictive data licencing conditions preventing future use of data.

"we're only at the very beginning of this, [our talks] will be directly with the producers...and we won't promote one type of on farm software..."

"it's quite busy space, we want to know where should we put our time and effort in, to be honest"

"if I'm a farmer and I'm going to use some software to calculate emissions, how do I know it is going to do that the right way? There can be lots of penalties for reporting things to the EU for example who are saying they will take a 4% penalty of your revenue if your report incorrectly"

"that is something we need a partnership on...because the certification bodies all have different programs in terms of how they enter, store and oversee the data...so it makes it difficult for us to make recommendations on the data types we might need to have and how"

"I am concerned about misrepresentation of the data - if we are the ones who control and understand the data and we're putting it out there we have not full but greater control of the messaging that goes with it"

"...is the bank going to look at the baseline and say it's not very good and they'll charge me more? from a producer point of view, is there perverse outcomes to consider?"

1. The "Tassie to Townsville" project is an outlier here. In that project, producers are paid a premium for their produce if they supply data.

Perceived social, economic and environmental benefits of a data ecosystem

In general, stakeholders can see a range of benefits coming from the AASF Data Ecosystem

Most stakeholders are generally enthusiastic about improving the ability to more easily supply/access data related to agricultural sustainability. They also agreed that working as a community was the best way to achieve this outcome. In most cases, they can see benefits for themselves and acknowledge that there will be benefits for others as well. All are keen to remain involved in the development of the ecosystem in some way.

A key benefit will be the ability to provide consistency, clarity and ultimately efficiencies around data collection and sharing

The primary benefit that most stakeholders see is the opportunity for the community to agree on what data will be collected, who will collect it, when it will be collected, who will have access to it and more. They see opportunities to ensure that data is collected once, reducing the burden on data producers but also creating efficiencies in data collection processes.

The data ecosystem will also provide a mechanism through which the community can identify and address gaps in national data sets and infrastructure

A number of stakeholders, in particular those interested in commodity level reporting, believe the data ecosystem will provide the opportunity to come together to address issues of broad interest. This may be in the form of providing a united voice to engage with and influence existing data collection activities and agencies. It may also be in the form of co-investing in data sets of interest. The data ecosystem may also provide shared data infrastructure to support the various use cases.

The data ecosystem will enable greater engagement across the industry to address myths and misconceptions, and collaborate to solve problems

An opportunity that many stakeholders acknowledged is the value of bringing the community together into a forum where they can discuss issues, learn from each other and work together to solve challenges. There are many forums where members of the agricultural sector can come together. However, there does not appear to be any forum in which discussions about agricultural data collection activities can be held.

The data ecosystem may also provide the opportunity for stakeholders to benchmark themselves against their peers

With consistency in the data being collected across the agriculture sector, it may become possible for organisations to benchmark themselves against the industry averages or each other. This, in turn, has the potential to drive improvements in sustainability outcomes across the industry as organisations compete with each other.

Ultimately, the data ecosystem may also enable the community to find ways to return value to data producers.

While no solutions were suggested, a number of stakeholders identified the desire to be able to return value to data producers in some way.

"you need to come back to what are you trying to demonstrate and for whom, and then work out what data can tell that"

"you could end up with data and metrics that are in search of an audience....I think there is a top down story here and a bottom up story here...and we're hoping people don't lose the will to live in the middle section"

"the challenge is that everyone is coming from a slightly different perspective - different industries, international markets, environmental lense, social lense etc....that is where it is hard, because it is hard to make everyone happy"

AASF ecosystem could be "a database online of compiled data representative of ag industries on standardised metrics and data which can be filtered by commodity"

"the more that the data ecosystem produces information that people can use, rather than having to create that information themselves, the better it will be used"

"our members are driven to participate because they see the need...Australia is seen as an increasingly risky investment opportunity...this will remain unless we transform our industries....Australia will be left with their pants down if we don't transition..."

"show some levels of truth on broad areas....and a dashboard can show you the various levels of truth as compared in areas and internationally"

Opportunities, blockers and potential risks to consider for a data ecosystem

The greatest opportunity of, and the greatest risk to, the data ecosystem is *trust*

There are many areas of distrust when it comes to data and data sharing within Australia's agricultural sector. The AASF data ecosystem can be a vehicle for building trust within the sector as stakeholders will need to work together to achieve the collective benefits on offer. However, this trust could easily be compromised if the implementation of the data ecosystem does not meet stakeholder expectations. These expectations include transparency, equality, inclusion and that the value of the system needs to be enduring.

"there is recognition that having auditors and consultants who can help farmers get the data is really important"

"There's a danger that the ASF becomes just another layer in something that needs to be broader"

Getting leadership of the data ecosystem right is essential

There is currently no consensus within the stakeholders interviewed as to how the data ecosystem should be governed once established. Stakeholders did state that the key characteristics of any leading organisation would be: trusted, reputable, respected, independent, apolitical and understand the problem.

A number of options for this leading organisation were suggested. In some cases, these suggestions were based on misunderstandings of the role and capabilities of that organisation in the current landscape.

Many stakeholders were forthcoming with who they thought should not lead, further emphasising the trust issues in the sector based on previous experiences and misunderstandings.

It will be essential that the AASF community works together to develop an acceptable leadership model.

A data ecosystem must specifically address, and have ongoing solutions for, privacy/security and the prevention of misuse

The privacy of individuals and security of their data is paramount. Regardless of the use cases and functions supported by the data ecosystem, management of privacy and security of data shared through the system must be a first-class issue.

"you want as much interoperability as possible, to reduce friction and costs with international markets"

"our legislation really restricts how people can access and use our data"

The data ecosystem community needs to provide strong guidance to the ag tech sector on its needs and expectations

While there may be initial reticence, technology companies generally look at the communities they support for guidance on what that community's needs are. To this end, the data ecosystem community needs to work with the ag tech sector to ensure that community needs with respect to the sharing of sustainability data along supply chains, are able to be supported by their product offerings.

"ultimately my goal is to move towards more satellite data - for soil carbon, ground cover, biodiversity cover, I think there is opportunity to have these remotely monitored..."

"farmers in every commodity are collecting so much data...and we have identified it would be great if they could put their data into a central pool and use that at an industry aggregated level...but it is really hard because there are some suppliers who are holding onto that data tightly"

"we have an issue with confidentiality....this is a key issue with our farm survey data which is a key thing that often people would like to have access to....we need methods of desensitising that information basically...."

"we're looking for data that is free or cheap, robust and good enough to tell a story over time...and we're looking to grow the quality of this over time"

"The system needs to work for everyone and including along the value chain especially customers (brand owners, retailers, consumers) and enablers and not put the burden on a particular group"

Summary of agriculture data ecosystem functions in Australia

Different drivers are informing how organisations develop their data practices, frameworks and governance arrangements

There is no single driver that is influencing all agriculturally focussed organisations in their activities around sustainability. For some the impetus is coming from international clients and markets, for others domestic issues are paramount. Some are focussed purely on emissions while others see different challenges as a higher priority. This impacts the data organisations are collecting, the way they collect this data, where they collect it from, and any associated governance arrangements.

Different users will engage with and use the AASF and hence the AASF Data Ecosystem in different ways

There is a great deal of interest in the AASF and many are looking forward to being able to use it for their purposes. These purposes vary between direct use of the full framework for reporting purposes; adopting elements of it to include in other reporting frameworks; and guidance for development of their own reporting. The latter are likely to be company reporting along supply chains or investment portfolios.

The current agriculture sustainability data ecosystem is anarchic in nature

The current agricultural sustainability data ecosystem in Australia lacks any form of coordination or organisation. As a result, individual organisations, whether they be commodity specific sustainability frameworks, supply chain participants or others with a need to access and use sustainability data are all 'doing their own thing' with respect to data collection and management. The result is significant cost across the ecosystem due to duplicative data collection, lack of consistency around what is being asked for, and an increasing burden on those being asked to provide data with a subsequent degradation in data quality.

Data sharing within the agricultural sustainability sector is undertaken on an ad-hoc basis

Currently, the primary approaches to sharing data within the agricultural sustainability sector is to employ a set of 'point-in-time' solutions. That is, solutions that support a specific need at a point in time rather than general solutions. This applies to both the data being shared and the methods by which it is shared. The result is a complicated array of, often, incompatible data and applications, addressing very specific needs, applications, and resultant costs for both data producers and users.

In general, stakeholders can see a range of benefits coming from the AASF Data Ecosystem

Most stakeholders are generally enthusiastic about improving the ability to more easily supply/access data related to agricultural sustainability. They also agreed that working as a community was the best way to achieve this outcome. In most cases, they can see benefits for themselves and acknowledge that there will be benefits for others as well. All are keen to remain involved in the development of the ecosystem in some way.

The greatest opportunity of, and the greatest risk to, the data ecosystem is *trust*

There are many areas of distrust within the agriculture sector. The AASF data ecosystem can be a vehicle for building trust within the sector as stakeholders will need to work together to achieve the collective benefits on offer. However, this trust could easily be compromised if the implementation of the data ecosystem does not meet stakeholder expectations. These expectations include transparency, equality, inclusion and that the value of the system needs to be enduring.

2.3

What variables
need to be accounted for
in establishing an
AASF Data Ecosystem?

Stakeholder Cohorts and Data Supply Chains

It has been noted previously that there are many different stakeholders with a potential interest in the AASF Data Ecosystem. The discovery phase has tried to engage with as many of these cohorts as possible, although it is unlikely that all voices have been heard. Even with the stakeholders engaged, the variety of expectations and needs is significant. These include:

- *Organisations wishing to tell the Australian agricultural sustainability story to secure market access (Government, others)*
- *Organisations wishing to tell the sustainability story of a particular agricultural commodity (or commodities) (Government, industry groups)*
- *Organisations needing to report on the sustainability of their portfolio of clients to secure capital (finance)*
- *Organisations needing to report on the sustainability of their supply chains for regulatory purposes within the markets they operate*
- *Organisations that need to supply sustainability information along supply chains.*
- *Organisations that collect sustainability related data at a local, regional or national scale*
- *Organisations that provide digital productivity tools into the agriculture sector*
- *... and many more*

A challenge for those designing the data ecosystem is to agree on which of these stakeholders and supply chains the data ecosystem will support initially and how this support will be provided. Essentially the challenge is to balance the competing challenges of achieving a scale that drives efficiencies with increasing complexity.

Use Cases

Related to the number and variety of stakeholders in the data ecosystem is the challenge of determining the high priority use cases the ecosystem will support, as this will greatly influence the required functionalities and structures.

The discovery process identified the following high level potential use cases:

- *Providing the data needed to support the development of whole of Australia agriculture sustainability reports for use in international trade negotiations*
- *Providing data that could be used to support existing and emerging commodity scale national sustainability reporting*
- *Providing standards and guidance around data that could be used in the development of supply chain/portfolio sustainability reporting*
- *Enabling access to data needed to support individual supply chain/portfolio sustainability reporting*
- *Enabling individual organisations to benchmark themselves against industry standards or their competitors*
- *Connecting members of the Australian agricultural sustainability network to enable discovery of data and tools*

These are just some of the potential use cases that the AASF Data Ecosystem might support. More will become apparent with increased stakeholder engagement in the AASF, and with increased maturity of the data ecosystem itself.

The task when prioritising these use cases is to find a balance between the number of stakeholders benefitting because their use is being supported, and the potential complexity (technical and institutional) of uses cases as the number of uses increases.

Other considerations

The design of the AASF data ecosystem will also need to consider:

- **Data Licencing** – in past years, data licencing was the bane of anyone trying to access data from others due to many organisations employing, restrictive and, often, incompatible, licences on their data assets. Reasons for this include to protect intellectual property, maintain privacy, prevent misuse of the data, or prevent people profiting from data use (to name just a few). Fortunately, many organisations these days (particularly government) have adopted far more flexible, less restrictive licences (e.g. CC By 4.0). However, data licencing is a subject that the data ecosystem should explicitly have policies for to ensure that data licences used within the ecosystem are compatible
- **Governance models** - A key consideration when designing a data ecosystem is the governance model to be employed. Micheli et al (2020)¹ explored Patterns currently in use including:
 - *Data Co-ops* – Cooperative models where data is collectively owned and managed by members, emphasizing democratic control and shared benefits.
 - *Data Trusts* – Legal structures that manage data as a trust asset for the benefit of designated beneficiaries, providing a framework for ethical and accountable data sharing.
 - *Data Sharing Pools* – Collaborative platforms where participants agree to share data under predefined conditions, facilitating access to a wider data set.
 - *Personal Data Sovereignty* – Principles and practices that empower individuals to control their personal data, including how it is used and shared.
- **Permissioning, Privacy, Security** – It is possible that data accessible to data ecosystem participants will be private or confidential in nature. It is essential that the ecosystem is able to ensure this information is secure, that users of data have permission to do so and that the privacy of individuals and organisations is maintained. Failure to address these issues risks both legal sanction for operators and participants in the ecosystem should breaches occur, as well as breaking trust in the system
- **Hosting** – it may be that technical infrastructure is required to support the functioning of the data ecosystem. A key consideration when implementing this infrastructure is the decision on how this infrastructure will be hosted. That is, which organisation will be responsible for maintaining this infrastructure and what approach will they take to this task. This decision needs to be made relatively early in the planning for the infrastructure as the unit responsible for maintenance need to be involved in the design process. Failure to decide on the host and include them in the design process will lead to extensive delays in deployment and may result in failure of the activity.

1. Micheli, M., Ponti, M., Craglia, M., & Berti Suman, A. (2020). Emerging models of data governance in the age of datafication. Big Data & Society, 7(2). <https://doi.org/10.1177/2053951720948087>

3.1 Our Recommendations

The AASF Community needs to agree on the *vision* for the AASF Data Ecosystem

The purpose here is to ensure the AASF community make a clear statement about what they hope to achieve through the implementation of an AASF data ecosystem. This statement can then be used to guide prioritisation of use cases, development of objectives for the ecosystem, decisions on institutional arrangements, and the design of technical and/or institutional structures.

The discovery phase has found that different stakeholders have different visions for the data ecosystem. These range from a suite of common data sets related to AASF sustainability criteria that can be used to support national scale sustainability reporting to a sophisticated data sharing capability that supports the sustainability reporting needs of the entire Australian agriculture sector. Some envision significant investments in technology while others merely seek guidance and standards.

The AASF Data Ecosystem could be some or all of these things. What is important is that the community agree on what it will be.

The AASF Community needs to agree on the *goals* for the AASF Data Ecosystem

A number of goals for the data ecosystem were identified during the discovery phase. Some of these are:

- *Reducing the burden on data providers*
- *Increasing efficiencies around data handling*
- *Growing trust between stakeholders*
- *Ensuring the longevity of data (and analysis) for use in the future*
- *Ensuring consistency and commonality of data uses across agricultural sustainability sector*

As for the vision, the AASF Data Ecosystem could adopt some, or all, of these goals. It is important that the community be clear on what they are for (ie intention).

The purpose here is to identify priority stakeholder requirements to be addressed in achieving the vision. Whereas the vision states what the community wishes to achieve, goals provide guidance on how the community wishes to achieve this vision.

The AASF Community needs to agree on the *principles* of the AASF Data Ecosystem

A key issue that has been raised during the Discovery phase is that of *trust*. While building trust involves many factors, a set of agreed principles (conventions, rules) about the way the data ecosystem will be implemented is important. Principles give direction to all stakeholders on what is expected in terms of behaviours of both people and technical components within the data ecosystem.

The Discovery Phase did not specifically seek stakeholder feedback on principles for the AASF Data Ecosystem. The list below has been drawn from previous work undertaken by the authors¹ and best practice from other data sharing initiatives and communities:

1. **Data within the ecosystem is FAIR** – the FAIR data principles², developed within the research sector, require that data be Findable, Accessible, Interoperable and Reusable. This should not be confused with Open. Many communities have adopted the FAIR principles to support data and knowledge integration and promote sharing and reuse of data.
2. **Implementation Neutrality** – requires that the ecosystem not assume particular technologies for implementation. This allows for evolution to occur.
3. **Timeliness, Transparency, Traceability** – requires that data access must happen in a timely fashion, all information must be traceable to the point of truth and any processing undertaken must be transparent.
4. **Scalability** – ensure the ecosystem have longevity and wide accessibility.
5. **Simplicity** – all who interact with the ecosystem should be able to understand it to the level they need to understand it.
6. **“Who pays the cost?”** – efforts in simplification should be targeted at those areas where the most gain is to be made
7. **No Private Contracts** – there should be no component of the AASF Data ecosystem that depends upon unpublished private agreements which contradict or compromise the published standard.

1. Atkinson, R.; Cox, S.; Fitch, P.; Lemon, D.; O'Hagan, R.; Walker, G. The Water Resources Observation Network (WRON). Reference model - version 0.1 : Water for a Healthy Country. National Research Flagship. CSIRO LW/ICT Centre; 2007-03-31. [procite:487de1d3-5c7d-4725-8a47-65532b30a6d4](https://doi.org/10.1070/procite:487de1d3-5c7d-4725-8a47-65532b30a6d4).

2. <https://force11.org/info/the-fair-data-principles/>

The AASF Community needs to agree on the *high priority use cases* for the AASF Data Ecosystem

The discovery phase has uncovered several potential use cases that the AASF Data Ecosystem might support. These include, but are not limited, to:

- *Producing national/regional reports aligned with the AASF framework*
- *Supporting existing and future commodity specific sustainability reports*
- *Supporting supply chain/portfolio level sustainability reports*
- *Supporting sustainability benchmarking within the ag sector*
- *Supporting the development, maintenance and promotion of sustainability data standards for the ag sector*
- *Providing opportunities for the discovery of tools and data*

It is important that the community agree on which use cases the data ecosystem will support and prioritise those that are most important to achieve a sustainable data ecosystem. This is important as it will provide:

- *clear guidance to stakeholders on what the data ecosystem will and will not initially support*
- *direction to implementors on what functions are required and which stakeholders require these functions*

Note that use cases can evolve over time and agreement on use cases now does not mean new use cases cannot/will not be included in the future. The data ecosystem will need to evolve as technology and needs change. This implies then that one of the roles of the leadership of the data ecosystem is to manage the use cases and have in place an explicit program to evolve these as necessary.

How Might We...

The Discovery Phase of the AASF Data Ecosystem Design has found that, despite some concerns, the community is enthusiastic about the opportunity a data sharing ecosystem brings to solving many challenges associated with discovery of, access to and use of data related to agricultural sustainability. Multiple use cases for the data ecosystem have been discovered as have multiple stakeholders in the ecosystem. Thus, the challenge for the community now is not, whether or not, the data ecosystem is needed. Rather it is to answer the following questions:

How Might We...

- *design a robust data ecosystem that delivers enduring value for a range of stakeholders?*
- *develop components of the AASF data ecosystem so that stakeholders have equitable access and appropriate decision-making rights?*
- *implement a data ecosystem so that it can be sustained (have sustained resourcing) into the long term?*
- *continue to evolve the data ecosystem as requirements and technical maturity change?*
- *appoint leaders and designate roles, responsibilities and accountabilities within the data ecosystem which link to (or communicate with) the AASF governance structures?*

**As Australia's national science agency and innovation catalyst,
CSIRO is solving the greatest challenges through innovative science and technology.**

CSIRO. Unlocking a better future for everyone.

Contact us

1300 363 400
+61 3 9545 2176
csiro.au/contact

For further information

Environment Business Unit:

David Lemon
+61 2 6246 5724
David.Lemon@csiro.au

Laura Kostanski
+61 425 711 094
Laura.Kostanski@csiro.au
