Food Agility CRC

BEST PRACTICE DATA POLICY

foodagility.com
@foodagility
hello@foodagility.com

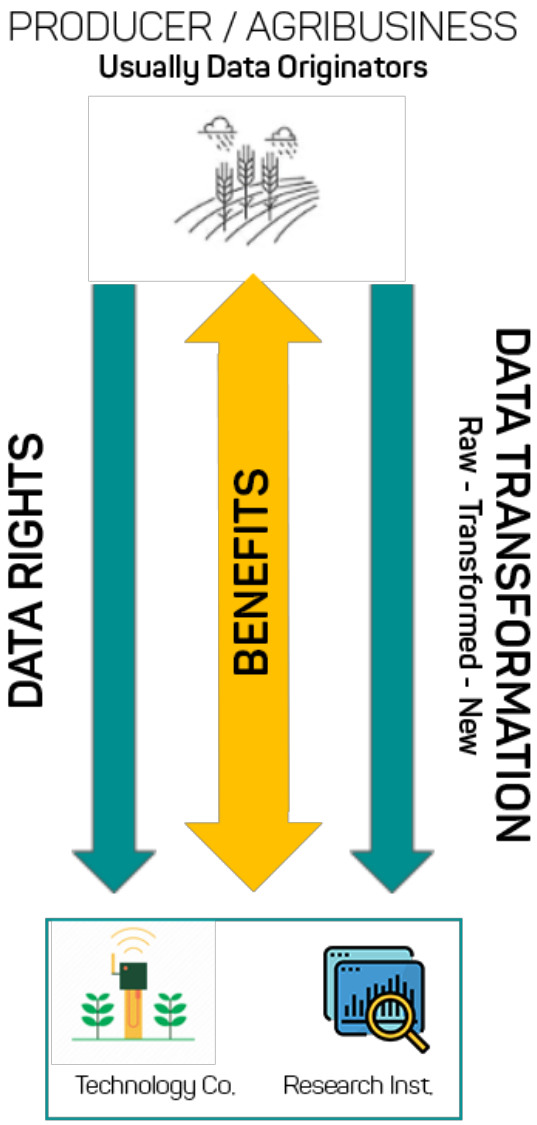
foodagility

Food Agility Best Practice Data Policy 'at a glance'

VERSION 3.0
LAST UPDATED
17.06.20
BY CAMILLA ROBERTS

Food Agility CRC is committed to leading the way in how data is managed in agrifood research to encourage data sharing, propel innovation and protect the interests of those that generate data. This Best Practice Data Policy sets out a framework to manage data in Food Agility CRC and the projects we invest in. We recognise that its application will be a transition and our partners will require support to implement some aspects. We look forward to working with our partners to evolve the Policy over time so together, we are world leaders in data management for collaborative research.

This is a brief overview of the Data Policy. See subsequent pages for the detailed policy and case studies.



DATA ORIGINATORS & RAW DATA

Producers/agribusinesses provide raw data to improve sustainable returns and research & technology development. They have the right to control it, benefit from it and have their confidentiality protected. The Data Originator can ask that raw data (but not Transformed or New data – definitions below) is moved or deleted at any time.

INCENTIVISING RESEARCH

We encourage & incentivise Data Originators to allow extended use of their raw data for research. See Section 7.2. Research use.

NEW DATA

As data is transformed, aggregated, analysed, modelled etc, it becomes much more valuable and may become New Data. That's why this new copyright tends to be owned by the organisation that created it. See Section 6. Summary of Data Types & Rights.

DATA RIGHTS TO NEW & TRANSFOREMD DATA

This Policy proposes but does not mandate the Technology Partner generally as the copyright owner of New Data as they have the greatest incentive to scale insights for industry benefit, however only with all parties' consent and confidentiality protected.

LICENCES VS OWNERSHIP

The owner can grant licences and rights to other parties. For example, the owner grants research rights to New & Transformed Data to the research institution for ongoing research and exclusive rights for the producer / agribusiness to use for business decisions. Confidentiality is still upheld. All parties would still need to agree this is the best balance of rights prior to project commencement.

DATA SERVICE PROVIDERS

The organisations that transform data or create new data from raw data, have a responsibility to ensure its accuracy and to protect it. They must get explicit written permission from Data Originators about how raw data is accessed, stored and used.

WHAT THIS POLICY MEANS FOR ME

AgriFood Supply Chain Partners

Farmers and other businesses in the agrifood supply chain are usually (but not always) the Data Originators in Food Agility projects. They provide their raw data to help improve sustainable returns and research and technology that will benefit their industry. We recognise that providing business information to third parties can feel risky. **This policy seeks to ensure farmers, agribusinesses and other Data Originators have control over their raw data and that their privacy and confidential information is protected.**

My rights as a Data Originator are to:

- control and benefit from the raw data I originate
- provide written, explicit consent for how and who manages and uses the raw data I originate
- have my confidentiality protected
- have my raw data deleted or moved if I request it
- agree any benefits I might get from new or Transformed Data relying in-part on my raw data
- be notified if there is a data breach
- know where the data is stored

My responsibilities as a Data Originator are to:

- ensure the data I provide is as accurate to their level of knowledge

Technology Partners

Our Technology Partners are usually Data Service Providers in Food Agility projects. Data Service Providers spend up to 60% of their time transforming data (Kelleher & Tierney, 2018) and often longer in agriculture because of the low digitisation of the industry. This processing may include standardising, formatting, aggregating with other data sets and outputs from models (see Transformed Data and New Data definitions). **This Policy sets clear expectations for how data should be managed, so technology partners can use data to develop new tools and services for the agrifood sector and can assure Data Originators that their data is well managed.**

My rights as a Data Service Provider (commercial) are to:

- (in typical cases) own the copyright of the data I transform (transformed data)
- (in typical cases) own the copyright of the data I create (new data)

My responsibilities as a Data Service Provider (commercial) are to:

- get explicit written consent from the Data Originator about access, use, storage and security
- agree with the Data Originator on their access to transformed /New Data
- agree with the Data Originator on confidentiality requirements
- agree with the Data Originator on fair benefits from other commercial uses of new/Transformed Data reliant in-part on their raw data
- delete or move raw data if requested by the Data Originator
- notify the Data Originator if there is a data breach
- notify the Data Originator where the data is stored
- be clear about limitations of data others rely on to make decisions (e.g. farmers using an app).

Research Partners

Researchers are also often the Data Service Providers in Food Agility projects. Researchers need time to analyse data to determine what is useful for the intended purpose. Data Originators can struggle to share data for this exploratory phase without a clear ROI. Researchers may also need extended access to data so they can build on their research beyond the project. For these reasons research uses are distinct from commercial uses. Food Agility encourages partners to consider sharing all or part of anonymised data sets for future research (subject to confidentiality). This Policy aims to set clear expectations for how data should be managed so research partners can use data to uncover new knowledge for the benefit of the agrifood sector. It also incentivises Data Originators to allow extended use of their data for research purposes.

My rights as a Data Service Provider (research) are to:

- (in typical cases) to be granted a research licence to use transformed or new data:
  - beyond project completion for research purposes and public good
  - to publish theses or journal articles

My responsibilities as a Data Service Provider (research) are to:

- get explicit written consent from the Data Originator about access, use, storage and security
- agree with the Data Originator on their access to transformed /New Data
- agree with the Data Originator on confidentiality requirements
- agree with the Data Originator on fair benefits from other commercial uses of new/Transformed Data reliant in-part on their raw data
- delete or move raw data if requested by the Data Originator
- notify the Data Originator if there is a data breach
- notify the Data Originator where the data is stored

be clear about limitations of data others rely on to make decisions (e.g. farmers using an app).

CASE STUDIES

These are hypothetical situations, playing out how the policy might be applied in practice. They are not the only way the policy can be applied. Every scenario is specific to the project, so partners are encouraged to work out the most mutually beneficial approach, within the parameters of the policy.

Case study 1: Yield prediction

We Love Lettuce is a horticultural company that wants to more accurately predict the quality and volume of their yield so they can optimise their planting and sales decisions. *We Love Lettuce* asks the *University of Data Decisions* to develop a model to predict yield under certain conditions. It also asks a technology company, *Leafy Tech* to deliver this in a smart phone App to their growers. *We Love Lettuce* sends growing information (planting date, inputs, field) and harvest information (harvest date, weight, quality grading, field) to researchers at the university. Researchers clean the data and transform it ready for analysis. They identify the most influential characteristics on yield performance by augmenting farm data with weather information from the Bureau of Meteorology. *Leafy Tech adapts the model to scale and deliver real-time results through* a smartphone app that draws in farm data and BOM data so farmers can make real-time decisions about planting and sales.

We Love Lettuce is the Data Originator. Their contract with the university and technology company should clearly specify how the raw data will be used and who will access it. They have the right to control how the raw data is managed and to benefit from it. They can also ask for the raw data to be deleted or moved if they want. This should be clearly explained in the contract with the university and technology company.

University of Data Decisions and *Leafy Tech* are the Data Service Providers. They have a responsibility to securely store the data to protect *We Love Lettuce's* confidential data. Any of the parties may have a claim to the Transformed data set. The Data Originator could own the dataset and licence to the Data Service Providers or vice versa. The parties determine that in this instance *Leafy Tech* will own the copyright to the Transformed Data because it underpins the model their developing which can be adapted to benefit the whole industry whilst still maintaining We Love Lettuce's competitive advantage and confidentiality.

Leafy Tech grants a research licence to the *University of Data Decisions* so the researchers can use the Transformed Data in future projects. *Leafy Tech* grants an exclusive commercial licence to *We Love Lettuce* so that it can use the Transformed Data to make business decisions. *Leafy Tech* cannot share the Transformed Data with other organisations without the expressed consent of *We Love Lettuce*. *Leafy Tech* must delete raw data if requested by *We Love Lettuce*, but is not required to delete Transformed Data. The agreement should specify who else *We Love Lettuce* can share the New Data with, for example a partner or contractor.

Case study 2: Pooling farm data to value Natural Capital

A group of Queensland cattle farmers want to improve the way they manage their natural capital and make business decisions for long-term sustainability. They engage in a project with *Queensland State University (QSU)* and technology company, *Natural Data* to create a model-driven tool that will make recommendations for better management of natural capital.

QSU and *Natural Data* collect data from the farmers about their finances, production and natural capital, some of which will come directly from the farmers' accounting platforms. They also collect data from a land quality assessment platform (which has ingested and transformed public data) and infield tests performed by the researchers.

Analysis by researchers identifies the key attributes that contribute to good management of natural capital. The combined datasets are run through the model in a farm decision-making tool created by *Natural Data* which makes practice recommendations that are likely to improve the natural capital outcomes.

The Queensland cattle farmers are the Data Originators of their finances, production and natural capital data. The land quality assessment platform, QSU and Natural Data are Data Service Providers. The commercial Data Service Providers own the Transformed Data in their platforms. All parties including the Data Originators have a clear agreement to each of the data sets in relation to their rights (access, storage, confidentiality, use purpose etc.) which is summarised in the Food Agility Project Agreement.

The farmers have also agreed to share the Raw Data in an aggregated, anonymised form in perpetuity for research purposes so *QSU* researchers can continue to build on their research for the benefit of the whole industry.

Natural Data owns the rights to the New Data created, but grants a research license so the researchers can continue to use the aggregated, anonymised data in future projects and for publishing theses and journal articles (so long as confidentiality and trade secrets are protected and the Data Originators agree). *Natural Data* also grants exclusive rights to the New Data to the farmers.

Food Agility rewards this project with further benefits (to be determined) because participants are allowing their data to be used beyond the project to benefit the whole industry.

Case study 3: Leafy Tech New Data

The project involving *Leafy Tech, We Love Lettuce* and *University of Data Decisions* is completed. It created an algorithm to accurately predict yield for *We Love Lettuce*. *Leafy Tech* wants to commercialise the algorithm by making it more applicable to the broader lettuce industry. Currently the algorithm is trained on *We Love Lettuce's* unique conditions and so is not yet generalisable for all lettuce crops. Maintaining confidentiality of *We Love Lettuce's* raw data and the algorithm outputs related to their farm, *Leafy Tech* works with other lettuce companies on their separate data. It aggregates it with 5 other growers and reengineers the algorithm so it's more generalisable to all lettuce crops. Its prediction is not as precise as the custom-built algorithm for *We Love Lettuce* but it's still much better than what current lettuce farmers are using. *Leafy Tech* commences another research project to achieve this with Food Agility.

Hearing of this new aggregated lettuce grower data set, the State Government has asked to licence the aggregated data set from *Leafy Tech* and use it to recommend improvements to growing practice for the industry.

The lettuce farmers are the Data Originators. *Leafy Tech* is the Data Service Provider and owns the New Data (the aggregated farm data and algorithm outputs). As the State Government opportunity is a new purpose for the data it should be transparently discussed with all Data Originators (lettuce farmers including *We Love Lettuce*) explaining the potential risks and benefits. Benefits should be considered based on the Data Originators' individual investments, in proportion to the full value of the data set which is made up of other originator's data and many man hours from *Leafy Tech*. *Leafy Tech* must continue to adhere to individual confidentiality requirements or benefit arrangements agreed with originators should their data be commercialised.

foodagility

Data Policy

1. Entity mission

Food Agility CRC's mission is to lead the digital revolution for a sustainable food future. Critical to this is trust and collaboration so problems and data are unlocked for researchers, technology providers and supply chain participants, to collaborate on solving them faster.

Food Agility's role is to provide better, faster, longer access to data in a way that best manages the concerns of all parties. Central to achieving this is establishing trust and core to establishing trust is transparency, accountability and education. One vital role that Food Agility plays in the industry is to lead by example through the adoption of best practice and by communicating with and educating partners along the way to show how the benefits of responsible and controlled sharing of data outweigh the risks.

2. Purpose and scope

This policy applies to all directors, office holders, employees and those contracted to Food Agility CRC Limited, as well as project participants and the data inputs and outputs of projects.

The purpose of this policy is to state the principles relating to data that Food Agility will uphold, and that we will do all we can to ensure these principles are upheld by directors, office holders, consultants and employees.

Food Agility's goal is to unlock responsible and controlled data sharing for faster, decision-quality insights. The Policy addresses Data Originators and other participants in the supply chain through to users of data (e.g. technology companies and research institutions) and insights derived from data.

The policy covers both the management of data in projects and strategic approaches to provide enduring agriculture data access for stronger industry and research collaborations.

We recognise that principles mean little unless assured through implementation processes and methodologies. We will do what we reasonably can to ensure that these principles are reliably and verifiably implemented.

It is important to note at the outset that most data is not 'owned' in the traditional sense of ownership of physical property. This is because data is generally not recognised as 'property' in law. The terms 'data owner' and 'data ownership' should therefore be used with caution, as they do not have clear legal meaning. It is therefore particularly important for entities handling and sharing data to be clear with each other (for example, by stating in a data sharing agreement or other written contract) which entity has what rights and obligations as to what data, and who is entitled to make decisions as to control of that data (such as decisions about how the data may be used or shared and under what conditions).

3. Review of the Policy

This Policy is intended to be a living document and will be subject to an annual review cycle to ensure it remains relevant. Food Agility welcomes feedback on how future editions of the Policy can be amended and improved.

4. Obligations

We expect (and this policy supports) that research institution partners are responsible for their researchers upholding the Australian Code for the Responsible Conduct of Research[1] and implementing the principles and best practice standards of good research data management and primary materials in accordance with the Australian Research Data Commons and the Research Data Alliance. Further, Food Agility is aligned to the leading data principles and farm data codes and regulations, including; the National Farmer's Federation's Australian Farm Data Code, Data to Decisions' CRC Big Data Principles[2], the FAIR Technical Data Principles, American Farm Bureau Federation Farm Data Principles, The European General Data Protection Regulation (GDPR) and EU Farm Code, Australian Privacy Principles, and the Consumer Data Right (CDR). While many of these are not legally mandatory in Australia, together they represent global best practice in the handling of agriculture data.

5. Roles & responsibilities

The Chief Operating Officer is responsible for meeting legal requirements and handling data disputes. The Chief of Ventures is responsible for the Data Policy and Strategy including how they are operationalised. Consistent with best practice, Food Agility will appoint a Data Protection Officer, who monitors and assures Data Originator's rights are respected as per their agreements. The Innovation Managers are responsible for ensuring the data and Project Agreement are compliant with the Policy for the projects and partners they are responsible for.

6. Definitions

Data All forms of information that are transferred between the Data Originator, data provider and data users or third parties during the course of business operation.

- Anonymised data is data that has been rendered anonymous, and is thus no longer personal, by irreversibly stripping it of any identifiable information. This makes it impossible to gain insights into a discreet individual, even by the party that is responsible for the anonymisation. Privacy laws do not apply to anonymised data since it is not personal. Anonymisation has to be sufficient as to ensure that cross-referenced data could not reidentify the data, for example any geographically based data could identify a farm so location data would also need to be anonymised.

[1] https://www.nhmrc.gov.au/about-us/publications/australian-code-responsible-conduct-research-2007

[2] https://p2d-bdra.web.app/pages/principles/principles.html

- Pseudonymised data is data that has been stripped of direct personal identifiers and is handled subject to controls and safeguards that reliably reduce the risk that any entity that has access to that data will be able to identify the data subject to the level where risk is reliably and verifiably assured as remote. Pseudonymised data is therefore not fully anonymised and must be handled in accordance with controls and safeguards against reidentification of the data subject. Outputs from that pseudonymised data environment must also be assessed for whether those outputs will be able to reidentify the data subject by reference to those outputs (and also taking into account other information that may be reasonably available to that entity). If the risk of reidentification of a data subject from those outputs is unknown, or should be assessed as greater than remote, the outputs should be regarded as identifying data and handled as such by the disclosing entity. Privacy laws do apply to pseudonymised data about identifiable individuals to the extent that it can become personal information about individuals.
- Raw data is data that is generated and collected without editing or any other form of processing.
- Transformed Data is data that has been transformed from its original raw state so that it is usable in analysis. It is distinct from New Data. Transformation processes include:
    - Duplication detection
    - Deidentifying (anonymising, pseudonymising)
    - Data enhancement (e.g. adding country code to a mobile number or postcode to an address)
    - Creating consistent, logical variable formats (e.g. date, a character, a factor, a numeric, an integer or logical variable, e.g. Date formats consistent)
    - Summarising (e.g. by sum or average – rainfall volume grouping rather than every individual volume measure)
    - Recoding (e.g. data set may use a character variable but a factor variable is needed - numeric age given but age group needed)
    - Recoding the levels within a factor (e.g. rather than all varieties selecting the top 6 and the rest go in 'other')
    - Missing data
        - missing as not applicable (e.g. bulls cannot be pregnant)
        - missing as not available (e.g. no data as the plot had no yield due to stress)
    - Decisions related to missing data
        - leave as missing
        - fill in data through imputation
        - or at the model stage use a parameter model of another variety (e.g. mango model to adjust for avocadoes).
    - Standardise different systems (e.g. if one system records as bushels per acre and another in kg per hectare, and Kg is needed then bushels are converted to Kgs.
    - Variable matching (e.g. often the baseline has changed when data has been collected over a long time series and requires assimilating)
- New Data is data that is developed in processing and combining data to create results or New Datasets, for example the results produced by an algorithm would be new data.

Summary of Data Types & Rights

Data Type	IP Type	Rights
Raw Data & Cleaned Raw Data	Typically considered confidential information /trade secrets of the data originator.	Data originator grants rights to the data they originate and as laid out in this Policy.
Transformed Data	Dependent on agreement reached between the parties that balances the transformation time and skill investment with the overall investment and projected benefits to partners. There are multiple models to consider here including but not limited to: Data Originator owning and licencing to Data Service Providers Data Service Provider with exclusive licence to the Data Originator.	The first model is that the Data Originator still owns the data and licences the Transformed Data to the Data Service Provider. There are 6 potential approaches for this under the Creative Commons licences where copyright exists. The second model, could see the Data Service Provider owning the Transformed Data granting rights to the transformed data. This is to be negotiated and agreed with the Data Originator and may include that the Data Originator has exclusive rights to the Service Provider's Transformed Data. Confidentiality restrictions still apply but the data originator cannot port or terminate transformed data only the raw data.
New Data	Typically considered copyright of the party that created the new data.	Service Provider that created the new data grants rights to the new data. This is to be negotiated and agreed with the data originator and may include that the data originator has exclusive rights to the Service Provider's new data. Confidentiality restrictions still apply but the data originator cannot port or terminate new data only the raw data. Where the Data Originator (individual or business) cannot be reidentified confidentiality restrictions may no longer need to apply.

- Metadata is data about a particular information asset. Specifically, the contextual information about an information asset upon which the asset was established and will be managed on an ongoing basis. Metadata may include information about rights, the applicable constraints, and performance measures that will be or are being applied to the information asset. As contextual information, metadata assists in ensuring the authenticity, reliability, usability, integrity, and accessibility of digital records over time. Broader definitions of metadata include three key concepts. Namely the contextual metadata, plus the metadata

schemes (such as classifying values used) and the metadata schema – all of which are present within the abstract model in their component parts.

- Identifying data may be either personal information about an individual or information about a farm or farm business that is identifiable.
- Personal data is information about an individual human who is identified or reasonably identifiable (Privacy Act 1988 (Cth), section 6(1)). Personal identifiers include name, identification number, location data, telephone number, or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural, or social identity of an individual human. Farm data and GIS location at the distinct property level is identifying data but may not be personal information about an individual human. Linking datasets raises risks of seemingly non personal data becoming personal data. Personal data also includes Food Agility employee data.
- Publicly available data is data that can be freely used, reused, and redistributed by anyone with no local, national, or international legal restrictions on access or use.
- Primary data is raw data transformed into values that are identifiable by people (primary processing). For example, field data (e.g. soil data, water data, field data).
- Aggregated data is a combined dataset made up of a few or a wide range of sources (e.g. sensors, systems, farmers or data platform or other data sets). The aggregation of data can provide additional value when combined. If revealing information is stripped away, aggregating creates fully anonymised data. If aggregation still leaves the risk that some identifying data remains, the data should be treated as pseudonymised data. For example, individual data sets may not contain identifying data, but when combined with data such as geolocation data (e.g. field coordinates linked with registered farm business addresses and satellite data) could identify yield volume and expected revenue of a farm business.
- Agricultural data is data related to agricultural production including farm data and all types of data generated within the farming processes (refer to Annex).
- Big data is vast volumes of highly diverse data that can be captured, analysed and used for decision-making.
- Consumer Data Rights Data is information that is within a class of information specified in the designation instrument for an industry sector which is brought within the CDR framework or derived from CDR data.
- Derived CDR data is data that has been wholly or partly derived from CDR data, or data derived from previously derived data. This means data derived from 'derived CDR data' is also 'derived CDR data'. 'Derived' takes its ordinary meaning.

Data roles

- Data Originator is a person or organisation who has rights to the underlying asset from which the raw data is generated, either through ownership, leasing or some other arrangement, and the asset would be considered a part of their personal or business operations. The Data Originator grants rights to data collectors and data providers to access and use their raw data

as agreed by all parties. They claim the exclusive right to the raw data and control its downstream use or reuse. For example, a farmer providing or allowing access to property data to a data collector (machinery or sensor company), data user, or Data Service Provider.

- Data Collector is a person or organisation who has created/collected this data either by technical means (e.g. agricultural machinery, electronic data processing programs), by themselves, or who has commissioned data providers for this purpose. The data collector has the responsibility to ensure Transformed Dataaccuracy, reliability, security, and availability. NB: In the event that one entity or person has more than one data role then the policy applies to all the roles they fulfil.
- Data Service Provider is a natural or legal person that under an agreement delivers a data service.
- Data Provider is a natural or legal person that under an agreement delivers data to data users.
- Data User is a natural or legal person that receives data and uses it. The data users are most likely to be a Data Service Provider but may also be some other user of the data not providing a service to the Data Originator.
- Data Holder as defined by the CDR is the role that neither owns nor shares the data but is authorised to hold the data.
- Data Sharer as defined by the CDR is the role that neither holds nor owns the data but is authorised to share the data.
- Third party is a natural or legal person other than the Data Originator who receives data from the data user or data provider under an agreement.

Data Processes

- Data porting is where data is moved from one data user to another.
- Data sharing is the practice of making data available to data users or third parties.
- Data storage is the recording (storing) of information (data) in a storage medium. The Data Originator can store data in a primary location, in a data platform, or in cloud-based storage platforms. The location in which data is stored is referred to as the 'data storage location' or 'storage location'.
- Transformation See also Transformed Data above.
- Pseudonymisation is a procedure by which the most revealing fields within a data record are replaced by one or more artificial identifiers, or pseudonyms. The pseudonym allows the data to be traced back to its origins, which distinguishes pseudonymisation from anonymization. The purpose of pseudonymisation is to render the data record less identifiable and therefore lower the risks involved in its use. See also Pseudonymised Data above.

Data platform is software where applications are made available for data processing. Data platforms may be closed (just for members or open for Application Programming Interfaces – APIs), or may be open-source hardware platforms and software libraries.

Commercial in confidence is a classification that identifies information that, if disclosed, may result in damage to a party's commercial interests, intellectual property, or trade secrets. You must not disclose any information marked 'Commercial in Confidence' without permission from the party who supplied it.

Confidentiality means ensuring that information is accessible only to those authorised and is protected from unauthorised disclosure or intelligible interception.

Public funding means external research funding provided by a public agency in Australia or internationally. Public funding includes, but is not limited to, competitive grants from Federal or State Governments.

Research data management means all the processes and actions required to manage research data and primary materials throughout the research lifecycle for current and future research purposes and uses.

7. Policy

7.1. Attribution of underlying rights to derive data

The Data Originator controls data they originate. As a basic principle, when data is produced by an agrifood-chain operator due to their activity, or is commissioned by this operator, the operator is considered the Data Originator. The Data Originator has the right to determine who can access and use the data. This does not cover data/information generated by processing this data from multiple originators (aggregating), but the provision of data for such purposes should be part of an agreement. For instance, the rights regarding data produced on the farm or during farming operations are granted to the farmer and may be used extensively by them.

The nature and means of collecting different agricultural data leads to different levels of attribution of data rights. It is therefore crucial to set some key principles for agricultural data access and usage rights.

The parties (originator, collector, provider, user, holder, sharer, or third party) should establish a contract that clearly sets the data collection and sharing conditions according to the needs of the contracting parties. Details referring to data sharing must feature in a dedicated and exclusive section of the contract where possible.

The contract should acknowledge the right of all parties to protect sensitive information (e.g. Intellectual Property (IP)) via restrictions on further use or processing. Parties should not use, process or share data without the explicit, written consent of the Data Originator. The CDR considers current consent as given within 12 months. Food Agility will move towards this as soon as is practical, but it is not yet required by law. Data Originators have a right to know where their data is stored and if there has been an attempted data breach or breach. Universities have strict ethical approvals and

consent requirements depending on the type of data, and these additional requirements must be maintained through the dataset's life, even if it is aggregated into another dataset.

The Data Originator has the right to benefit from the data they originate. This Policy recognises the Data Originator's right, whether they are a farmer or another party, to benefit from and/or be compensated for the use of data created as part of their activity. It also recognises the need to grant the Data Originator the right to control the access to and use of data from their business and to benefit from sharing the data with any partner that wishes to use or aggregate their data with other data sets for new purposes. Therefore, the contract should clearly establish the benefits for the Data Originator. The originator could be compensated for the value created by an exchange of value as agreed by both parties. Of course, significant value may be added through transformation and combination of data and analysis to create useful insights. This added value should be recognised when considering the value of the raw data inputs as provided to the entity that adds to the value of the data inputs.

Terms shall be transparent, and contracts should use simple language in order to explain the content or be accompanied by an informal document that explains data-related aspects. Contractual agreements should specify:

- the most important terms and conditions
- the purpose of collecting, sharing, and processing
- the data rights and obligations that the parties have related to data, rules and processes for data sharing, security, and the legal framework in which the data is kept and in which back-ups are stored
- the software or the relevant application and information on the storage and use of agricultural data
- Verification mechanisms for the Data Originator
- Transparent mechanisms for adding new or future uses

7.2. Data access, control and portability

Explicit, informed, written consent must be granted by the Data Originator before collection, access, storage, and usage of the collected agricultural data occurs. The Data Originator must be informed in a clear and unambiguous manner if someone intends to collect and store their data. If parties are in agreement, the contract should specify the conditions under which the identification of the Data Originator may be possible. Otherwise, the data should be subject to pseudonymisation.[3] Universities have strict ethical approvals and consent requirements depending on the type of data, these additional requirements must be maintained through the dataset's life, even if it is aggregated into another dataset.

The Data Originator must grant permission for data to be used and shared with third parties, including circumstances in which decisions are made based on the data. Information should only be given to third parties as aggregated, pseudonymised or anonymised data (including any geographically based data), unless it is required to deliver the requested service and/or the

[3] According to Art 4 of the Regulation (EU) 2016/679 (General Data Protection Regulation)

conditions specified in the contract. Unless expressly stated in the contract that identifying data will be made available to third parties, the data user must take all reasonable precautions to ensure that the risk that any data (including insights and reports from data) may reidentify a farm, business or individual human is remote. If it is expressly stated in the contract that identifying data will be made available to third parties, those parties should be listed, the purpose of their receipt of that data specified, and they should be required to implement controls and safeguards against further re-identification by others.

Data should only be collected and used for the specific purpose agreed in the contract.

The datasets should only be kept for as long as is strictly necessary for the relevant analyses to be carried out, unless additional permissions and conditions have been granted for research use.

Access to data should be strictly and verifiably controlled.

Any transfer or change to the data (e.g. input, modification, removal), should be fully traceable (accompanied by metadata about the author and modification).

The Data Originator has the right to exclude others from using their data and prevent modification of their data should they choose.

The Data Originator has the right to port raw unit data. Data Originators should be granted appropriate and easy access and be able to retrieve their raw unit data. Aggregate data that is based on the data of more than one Data Originator, that has been anonymised and is no longer specifically identifiable, cannot be ported by the Data Originator. This is aligned with the CDR.

Data collectors, users, service providers, and holders should be responsible for making input data easily available, accessible, and readable where technically feasible. If not technically feasible, the data provider should provide reasonable justification. The Data Originator shall have the right to receive the data concerning their operation as specified in the contract, in a structured, frequently used and machine-readable format.

Unless otherwise agreed in the contract, the Data Originator has the right to easily port raw unit data to another data collector, user, holder, or service provider. If agreed between the parties, the Data Originator shall have the right to have the data transmitted directly from one data user to another, where technically feasible and at a reasonable cost.

Data Originators should be in no way restricted should they wish to use their data in other systems/platforms/data storage facilities, unless stated in the contract.

The data user shall disclose the means (e.g. if and how) through which a Data Originator may view, correct, retrieve, or extract data and the means by which data would be ported to another service including the data interchange standards and formats supported.

This should be done without compromising restricted access to machine data or sensitive data (relevant to the correct functioning of the machinery) and should be clearly specified in the contract between farmers/contractors and device manufacturers.

The Data Originator has the right to terminate and have their raw unit data destroyed. Data Originators are allowed to discontinue a service or halt the collection of raw data at any time subject to appropriate ongoing obligations.

Procedures for termination of services should be clearly defined in the contract.

Research uses of agricultural data are in the public interest, so it is important to note the distinction between Commercial use and Research use (non-commercial). As a research centre, Food Agility encourages flexibility in the parameters of the data collected and encourages longer storage of data and, where confidentiality is not breached, wider access to researchers on a permission basis. This is aligned with the Australian Code for the Responsible Conduct of Research[4]. There are more unknowns in discovery stages of research and so greater flexibility is encouraged but not mandatory for research use. It is Food Agility's responsibility to educate its partners to understand how to safely and securely share data that results in public benefit through research.

Data Originators have choice in relation to the parameters of data access for research use. As a research centre Food Agility encourages the first option for each of the three categories below, but it is not mandatory. The Data Originator's choices will be recorded in the contract.

1. Time period: in perpetuity (preferred in research for replicability), for 5 years or for a period the Data Originator is comfortable.
2. Access controls: open access to researchers, permission-based access where Data Originators grant permission as requested, or restricted access to the particular project for which the data was originally intended
3. Storage location: in shared infrastructure that Food Agility manages, in the infrastructure of the university of the research lead on the project, or in a commercial access space.

Data for research use will be made available for access and re-use by other researchers subject to any contractual, ethical, privacy or confidentiality matters. Metadata will be made available to other researchers via open access repositories so that existing data is findable. Where appropriate, and permissions have been sought and given, research data may be made available under open access licences or by negotiated or controlled access. To protect Data Originators, research Data Service Providers should pseudonymise or anonymise the data.

Incentives will be provided to encourage collaborative behaviours that will enhance the speed of useable insights to partners and the Australian agricultural industry as a whole. For example, data that can be shared more broadly and for longer periods so different researchers or technology partners can build on the body of work. Food Agility will not require data (whether Raw, Transformed or New) be held centrally but will encourage and incentivise a system that enhances research and commercialisation outcomes from data for Australian agriculture.

[4] https://www.nhmrc.gov.au/about-us/publications/australian-code-responsible-conduct-research-2007

7.3. Data protection and transparency

Data users, holders, collectors, and service providers must protect the data. Data Service Providers who control the database must have a protocol on data protection safeguards for individual originators that prevents unauthorised sharing with third parties. Furthermore, personal data in databases must be both stored under a pseudonym and encrypted or protected with similar methods. This is to render the data less identifiable and mitigate risks in the course of normal operations and in the event of a data breach. Data Originators will be provided with the Data Service Provider's contact for support or complaints.

Data Originators must be requested to consent before data sharing with new parties. If data is to be sold or shared with a third party that is not initially mentioned in the contract, or for a purpose not originally referenced when consent was obtained, the Data Originator must be able to agree on or refuse this without financial or other repercussions. Careful attention must be paid to sensitive data that requires university ethics reviews and mandates additional consents or requirements for data to be used. The Data Service Provider should only sell or disclose data to a third party if they have secured the same terms and conditions as specified in the contract between the Data Service Provider and originator or the Data Originator approves the third party's conditions.

Data Originators must be given the option to opt out of the contract and terminate or suspend the collection and usage of their data, provided that the contractual obligations have been met. This must be clearly stated in the contract and the Data Originators should be informed of the consequences of these decisions. Either this should be done upon their first request and is of immediate effect or it should be done after a previously defined notice period of a reasonable duration. This clause must grant the Data Originator permanent access to their data during the notice period.

Data Originators must be given choice. If several different services are on offer, Data Originators must be able to opt for none, one, or some. All services must be explained to have secured explicit consent. Where university guidelines state ethical approvals are required, these too must be followed.

7.4. Privacy and security

Data user's security and confidentiality responsibilities must be clearly defined in the contract. The data user should keep track of the data as much as possible throughout the value chain and share the gathered information with the Data Originator. Collectors and users of data should not use this data for unlawful purposes or take advantage of it to speculate for other such purposes.

As well as complying with other obligations in law, the Data Service Provider should not be permitted to use, or assist others to use, a Data Originator's data or insights or reports derived from a Data Originator's data to target treatment of that business or person to the detriment of that Data Originator as compared to other farms, farm businesses, or individuals.

If a data service provide reasonably anticipates that a recipient of a Data Originator's data, or insights and/or reports derived from that data, may be used by the recipient to the detriment of the

originator as a business or individual, the Data Service Provider should either not supply that recipient or very clearly state the risk upfront in a written contract between the Data Originator and the recipient.

Business sensitive data such as trade secrets and other identifying information about a farm or farm business should be securely maintained by a Data Service Provider applying at least equivalent standards to those required to safeguard personal information about individuals from unauthorised use or disclosure.

The Data Protection Officer of a Data Service Provider is responsible for monitoring and assuring Data Originator that their rights are respected as per their agreements.

Data Originators have the option to destroy raw unit data. There must be the option to remove, destroy (e.g. the right to be forgotten), or return all raw unit data upon the Data Originator's request within an agreed and reasonable time period such as 30 days. If hacking, seizure, confiscation, acquisition, insolvency, or settlement proceedings are detected, the Data Originator should immediately be informed by the Data Service Provider about the non-personal data being compromised and the measures taken. For personal data the Privacy Act 1988 (Cth) requirements need be followed.

7.5. Liability and intellectual property rights

The terms of liability shall be clearly laid out in the contract.

The Data Originator should state the known accuracy of their data (and/or completeness) and any known defects or limitations. They should not be liable for damage arising from and/or connected with the generation, receipt and/or use of this data by machines, devices, data users, and/or third parties.

Protecting trade secrets, intellectual property rights, and protecting against tampering are the main reasons why information is not shared and why even business partners in joint projects are not permitted to receive data. One main issue is being able to guarantee that these two interests, expressed as licensing conditions in the contracts, are respected. Protecting the intellectual property rights of different stakeholders in the value chain is fundamental.

7.6. Regulatory Framework

7.6.1. Regulation

Data is governed by a number of regulations:

- Privacy Act 1988 (Cth), including mandatory breach notifications
- Competition and Consumer Laws, for example, Unfair Terms laws within the Australian Consumer Law (ACL) and anti-competitive provisions of the Competition and Consumer Act 2010 (CCA)
- Regulations around Australian Government information security (for example national datasets on Australia's water reserves)

- State-based laws (for example Queensland Information Privacy Act, 2009)
- Environmental and Biosecurity laws
- Workplace health and safety laws

The Privacy Act 1988 (Cth) is the federal legal privacy framework in Australia is under the supervision of the Office of the Australian Information Commissioner (OAIC). The Privacy Act 1988 (Cth) is the main privacy law throughout Australia. The Act sets out the Australian Privacy Principles (APPs) and only applies to personal information. The Act applies in full to Food Agility. There are no exemptions for charities with an annual turnover above $3,000,000.

A mandatory breach notification regime[5] requires notification of any individuals affected by a data breach that is likely to result in serious harm. A data breach occurs when personal information is lost or subjected to unauthorised access or disclosure. This may include when:

- a device containing customers' personal information is lost or stolen;
- a database containing personal information is hacked; or
- personal information is mistakenly provided to the wrong person.

Where the data breach is likely to result in serious harm, the organisation must notify the individuals at risk of the serious harm and the OAIC as soon as is practicable. Exemptions apply when an entity can determine with a high degree of confidence that it has taken action to remediate harm, such that a reasonable person would conclude that the loss, access, or disclosure is not likely to result in serious harm to the individuals.

Law of Contract governs non-personal information. Food Agility will make clear in its contracts important elements such as scope and limits on use of data, confidentiality and security of data, duration of the contract and rights in relation to data on termination of the contract.

CDR

The CDR is currently only regulated in the Banking and Finance sector and is not a requirement in agriculture. As it will likely become a requirement in the medium term, Food Agility's Policy is aligned with the CDR as a best practice standard where feasible.

7.6.2. Compliance

Access Management Systems must be used for data in Food Agility projects in order to automatically record access, permissions, and breaches. These are routinely audited. Technological protection mechanisms must also be used (passwords, encryption, secure cloud, firewalls, site access controls, and 2-step authentication).

Privacy Management Framework and early risk assessment is used to minimise risk, especially in relation to confidentiality clauses and personal data. Food Agility uses the Privacy Management

[5] Office of the Australian Information Commissioner, Notifiable Data Breaches Scheme [website] www.oaic.gov.au/privacy-law/privacy-act/notifiable-data-breaches-scheme, (accessed 4 April 2018).

Framework[6] and the Five Safes Framework to determine privacy risks and records this in a Data Confidentiality Traffic Light System in the Risk Register.

Technical standards such as ISO 27001, NIST 800-27 Rev A and the FAIR Technical Principles are adhered to where practical. Compliance with the law is mandatory but compliance to technical standards are only of value where benefits outweigh the costs of compliance. For example, some standards (such as ISO27001) are overly onerous for small businesses. The Privacy Act 1988 (Cth) makes exemptions for small businesses and Food Agility advocates the same. Where the benefits of compliance do not outweigh the costs of compliance, the principles such as the FAIR Technical Principles can substituted.

The Data contact person at Food Agility for inquiries or complaints will be the Chief Operating Officer.

Other Protection Mechanisms include policies, procedures, and staff and partner education and training. The Food Agility Privacy Policy https://www.foodagility.com/content/privacy-policy.

7.6.3. Legal principles for a balanced contract

Food Agility has adopted the EU Code of Conduct on Agricultural Data Sharing's 'legal principles for a balanced contract' in the formation of its data agreements and encourages partners to independently adopt these principles also.

7.7. Capacity Building

7.7.1. Communication

Communication is essential for establishing transparency and therefore the trust needed for data sharing. In agriculture, digital literacy is one of the lowest of any occupation category and there is a high level of distrust of data sharing. So, transparency needs to be tackled on two fronts:

1. Making information available
2. Helping to make sure the information has been understood.

Therefore, an ongoing primary pillar of communication for Food Agility will be communication in the form of thought leadership and education of data, such that it builds trust and encourages data sharing.

7.7.2. Education & Training

Staff Training will be provided so that staff are able to informally educate partners to facilitate smoother and faster data sharing.

Food Agility Network Education, particularly of more vulnerable groups such as farmers, is essential to ensure trust, informed consent and expectation management between all parties. Food Agility will collaborate with grower groups to develop programs which help to create educated data partners

[6] OAIC's Guide to Big Data and the Australian Privacy Principles (APP) outlines key privacy requirements and encourages the implementation of the Privacy Management Framework to protect Personal Data in the use of Big Data.

who understand their rights and responsibilities. Contract explanation sessions will be held so the data terms and choices are understood.

8. Annexes

8.1. Types of agrifood data

Agriculture Data

- Farm data – data referring to farms and farm operations, including farm management
    - Agronomic data – related to plant production (e.g. yield planning, soil data, input data)
    - Compliance data – data required for control and enforcement in relation to competent authorities
    - Livestock data – related to herd (e.g. age, sex, performance indicators such as live weight animal welfare, and health indicators)
    - Climate, transaction and other environmental data
- Machine data – used for machine operations (e.g. data flowing between system controllers and machine sensors), often encrypted and not made available to prevent 'reverse engineering' or modifications on the on-board system communication which could result in the malfunctioning of controls in place to protect the operator and the machine.
- Service data – data used for vehicle maintenance and repair.
- Agri-supply data (input) – related to the nature, composition and use of inputs such as fertilisers, feedstuffs, plant protection products etc.
- Agri-service provider data – data originating from an agricultural services provider operating to benefit a client (e.g. farmers). Of sole interest to the management of the service-providing company (e.g. working time of employee, machine performance) and not related to farm operations.

9. Change Register

Version	Endorsed by (date)	Amendment
1	Board 10 March 2020	Approved for external consultation with Directors' feedback addressed.
2.	CIPCo 6 May 2020	Updated with internal and external consultation feedback and recommended by the CIPCo for Board approval in June 2, 2020 Board meeting.
3.	Board 2 June 2020	Updated to include all the CIPCo and additional external consultation feedback.

10. Reference Documents

Accelerating Precision Agriculture to Decision Agriculture Report, CRDC, 2017, https://www.crdc.com.au/sites/default/files/CRD18001-001%20CRDC%20P2D%20Report%20low%20res.pdf
NB: Although the number 1 recommendation from this report was for the development of the Data Management Policy for Australian Digital Architecture, it has not yet been developed. Neither have the Voluntary Data Management Code of Practice and Data Management Certification or Accreditation Scheme also referenced in the report.

Accelerating Precision to Decision Agriculture Report, The Legal Dimensions of Digital Agriculture in Australia, L Wiseman and J Sanderson 2017.

Australian Farm Data Code, National Farmer's Federation, February 2020. https://nff.org.au/programs/australian-farm-data-code/

Final Draft Agriculture Data Rules, Best Management Practice, Wiseman, L. and Sanderson, J. (2019). Griffith University, USC Australia and Cotton Research and Development Corporation, Australia, https://www.crdc.com.au/sites/default/files/Final%20Draft%20Agricultural%20Data%20Rules.pdf

A National Vision for Digital Agriculture, Mr Mick Keogh, Commissioner Australian Competition & Consumer Commission, 16 September 2019.

Australian Privacy Principles (APP Entities), Office of the Australian Information Commissioner, Australian Federal Government, https://www.oaic.gov.au/privacy/australian-privacy-principles/

Data to Decisions CRC Big Data Principles, Data to Decisions CRC, https://www.crdc.com.au/sites/default/files/CRD18001-001%20CRDC%20P2D%20Report%20low%20res.pdf

Data Supplier Flow Chart, ARDC - Australian Research Data Commons, https://ardc.edu.au/

EU Code of Conduct on Agricultural Data Sharing, https://cema-agri.org/publications/19-brochures-publications/37-eu-code-of-conduct-on-agricultural-data-sharing

FAIR Data Principles, FORCE11, https://www.force11.org/group/fairgroup/fairprinciples
https://www.crdc.com.au/sites/default/files/CRD18001-001%20CRDC%20P2D%20Report%20low%20res.pdf

Data Science, Kelleher, J. D., & Tierney, B. (2018). Cambridge, UNITED STATES: MIT Press.

New Zealand Farm iv Data Code, New Zealand Farm Data Accreditation Limited, http://www.farmdatacode.org.nz

Privacy Management Framework, Office of the Australian Information Commissioner (OAIC), Australian Government, http://content.webarchive.nla.gov.au/gov/wayback/20190509025456/https://www.oaic.gov.au/resources/agencies-and-organisations/guides/privacy-management-framework.pdf

Privacy and Security Principles for Farm Data, American Farm Bureau Federation, https://www.fb.org/issues/technology/data-privacy/privacy-and-security-principles-for-farm-data

Privacy and Security Principles for Farm Data, American Farm Bureau Federation, https://www.fb.org/issues/technology/data-privacy/privacy-and-security-principles-for-farm-data

QUT Data Policy, QUT.

The Privacy Act 1988 (Cth), Australian Federal Government, https://www.legislation.gov.au/Series/C2004A03712

Research Data Alliance (RDA) Endorsed Recommendations, RDA: https://www.rd-alliance.org/recommendations-and-outputs/all-recommendations-and-outputs

United States Ag Data Transparency Evaluator Core Principles, https://www.agdatatransparent.com/principles