Data Ethics and Stewardship White Paper

July 17, 2025

AgGateway is a global, non-profit organization that empowers agriculture stakeholders to gather, collaborate, create, and implement digital agriculture solutions that facilitate connectivity and seamless data flow. By targeting and resolving connectivity challenges in the value chain, we enable industry innovation by allowing ag manufacturers to focus on product development, and to embrace a collaborative approach to creating connections that build value.

AgGateway's more than 200 member companies include ag retailers, distributors, agriculture input manufacturers (seed, crop nutrition, crop protection, etc.), equipment manufacturers, grain and feed companies, precision ag providers, specialty chemical manufacturers, and software and data service providers. Our associate members include leading industry trade associations, international standards groups, state agencies, and members of academia focused on data exchange issues. More information is available on our website at http://www.aggateway.org .

1. Executive Summary

1.1 Overview

The rapid advancement of data-driven technologies in agriculture offers immense potential to enhance productivity, sustainability, and economic resilience. However, this increasing reliance on data also brings ethical, legal, and stewardship challenges that must be addressed to ensure responsible use and equitable benefits for all stakeholders. This document explores key principles of data ethics and stewardship within agriculture, emphasizing the importance of privacy, ownership, transparency, fairness, and responsible governance in agricultural data management.

1.2 Key themes

- **Data ethics by design**: Ethical principles should be integrated proactively into data agreements and data management systems to ensure compliance with regulations and maintain data originator trust.
- **Understanding agricultural data**: Agricultural data can be divided into several categories, including farm management, agronomic, machine, weather, and livestock data, all of which may have multiple stakeholders with different needs and expectations.
- **Standardization & interoperability**: Developing common standards for data formats, exchange protocols, and semantics is essential for efficient and effective data sharing.

- **Ethical considerations**: Privacy, ownership, control, transparency, fairness, and accountability are critical concerns when handling agricultural data.
- **Stewardship challenges**: Organizations face technical, legal, and ethical challenges related to data integration, security, regulatory compliance, and long-term sustainability.
- **Legal & regulatory compliance**: Managing agricultural data requires navigating evolving and conflicting laws related to privacy, ownership, and intellectual property.
- **Best practices & frameworks**: Several global sets of principles [examples in Appendix C] provide guidelines for ethical data governance and stewardship.

1.3 Best practices in data stewardship

To navigate the complex landscape of agricultural data management, organizations should adhere to the following best practices:

- **Data transparency & ownership**
  - Clearly define data ownership rights to ensure data originators retain appropriate control over their data.
  - Provide full transparency regarding data collection, storage, and usage policies.
  - Obtain informed consent before collecting or sharing data.
  - Allow data originators to access, transfer, and delete their data upon request.
- **Privacy, security, & ethical use**
  - Ensure data privacy and confidentiality by implementing secure storage and encryption.
  - Minimize data collection to only what is necessary for specific purposes.
  - Monitor and audit data movement to prevent unauthorized access or misuse.
  - Anonymize data effectively to prevent re-identification.
- **Fair & equitable access**
  - Promote inclusive data collection to avoid bias and ensure fair representation.
  - Mitigate algorithmic bias in decision-making processes.
  - Reduce the digital divide by communicating clearly and improving access to data literacy programs and digital infrastructure.
- **Data standardization & Interoperability**
  - Collaborate and adopt standardized data formats to enhance interoperability.
  - Use common terminology and definitions to facilitate effective communication.
- **Long-Term Data Stewardship & Governance**
  - Develop clear data-sharing agreements that define permissible use and all parties' obligations.
  - Ensure compliance with evolving regulations like GDPR and CCPA.
  - Implement robust accountability measures to track and report data use.

- Create sustainable data policies that account for system evolution and long-term storage.

1.4 Conclusion

The ethical and responsible management of agricultural data is essential to fostering trust, enhancing innovation, and ensuring equitable benefits for all stakeholders. In order to comply with these guidelines organizations shall proactively integrate ethical principles into data governance frameworks, address technical and legal challenges, and prioritize transparency, security, and fair access. By adhering to established best practices and ethical guidelines, the agricultural industry can harness the full potential of data while safeguarding data originator rights.

For further inquiries, visit http://www.aggateway.org.

2. Introduction

2.1 Background

Agriculture stands at the intersection of technological advancements, environmental sustainability, and societal well-being. The collection, analysis, and utilization of agricultural data have become integral to optimizing crop yields, managing resources efficiently, and ensuring food, fiber, and fuel security. However, with great data power comes great responsibility. **Data ethics**—the moral principles governing the handling and use of data—plays a pivotal role in shaping the future of agriculture. Equally important is **data stewardship**, which involves safeguarding data and using it for the benefit of data stakeholders throughout its lifecycle, from collection to dissemination. In this context, ethical considerations guide how we collect, share, and leverage agricultural data, ensuring that it benefits all stakeholders while minimizing risk of harm.

In this document, we use the term "data originator" to refer to any entity whose actions or property are being measured and the resulting data subsequently sent to another entity. Other entities that obtain or interact with data are referred to as "data exchange partners" or "data consumers." These and other important terms are defined in Appendix E.

2.2 Purpose

The scope and range of available agriculture data are rapidly expanding. These data are being generated, collected, and managed in many forms across agriculture value-chain segments and the potential for this mass of data in its variety, velocity, and volume [insert reference] to enhance our agriculture system is significant. AgGateway, as an industry eBusiness consortium, is in a unique position to help encourage the development of responsible data practices across the agriculture industry, specifically in:

- Defining common data categories or classifications.
- Developing standard and clear terminology for improving communication among data originators and data exchange partners.
- Facilitating a forum for the exchange of ideas and discussion of new developments in this continually evolving space.
- Maintaining collaborative relationships with industry stakeholders and providing a balanced source of information.

This paper explores the various dimensions of data ethics and stewardship within the agriculture domain, emphasizing the need for responsible practices that balance innovation with social, economic, and environmental well-being.

2.3 Thesis

Data ethics by design refers to the proactive integration of ethical principles and considerations into the entire lifecycle of data management and technology development. This approach ensures that ethical issues such as privacy, fairness, transparency, and accountability are addressed in system and process development from the outset, rather than as an afterthought. By embedding ethical principles into their design processes, data exchange partners can build systems that respect data originators' rights, comply with regulations, and foster trust among stakeholders.[1] This approach is particularly critical in agriculture, where trust plays a foundational role in business relationships with many agreements made based on informal understandings e.g., a handshake, rather than formal contracts. Additionally, the complexity of agricultural data flows means a data originator may rely on a trusted service provider to process and analyze their data, sometimes through multiple software tools and third-party integrations. With trust between people at the center of this complex industry, adherence to ethical principles is a fundamental requirement for broad adoption of technologies that require data sharing. There are numerous examples of these principles both from within the agriculture industry as well as other segments. A collection of frameworks of principles such as FAIR, CARE, TRUST and others can be found in Appendix C with links to more information on each resource.

[1] "Data Ethics: What It Means and What It Takes | McKinsey." Accessed July 15, 2025. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/data-ethics-what-it-means-and-what-it-takes.

3 Understanding Agricultural Data

3.1 Overview

An unprecedented number of choices for integrated systems and solutions have driven the agriculture industry towards the collection and use of data from tractors, combines, environmental sensors, irrigation equipment, grain carts, unmanned aerial systems, and more. To facilitate the demand for these highly integrated systems and to reap the benefits while reducing risks, it is important to consider how data will be managed. Agricultural data has increasing value to its originators. At the same time, data originators are growing more aware of and sensitive to what these data may tell others about their operations and how others may be monetizing their data without their knowledge or informed consent.[2] Data originators want transparency regarding the lifecycle of their data including collection, use, sharing options, ability to delete or remove, and effective security controls.

3.2 Data lifecycle

As referred to earlier ethics must be considered at every stage in the lifecycle of data taken from the "Data Lifecycle model" by the University of Wisconsin Data Governance Program:[3]

1. **Plan:** The initial phase of the data lifecycle, where strategies for data management are established before collection or acquisition. This includes defining governance roles, access controls, usage policies, and retention guidelines, while ensuring compliance with relevant laws, rules, and regulations. Accountability for data stewardship is also determined at this stage.
2. **Create:** The stage where data is generated, captured, or recorded from various sources, including manual entry, automated sensors, or system outputs. Ensuring accuracy, completeness, and proper metadata documentation at this stage is essential for downstream usability, integrity, and governance.
3. **Manage:** The stage encompassing data storage, security, integrity, and retention from creation to destruction. IT specialists, including data architects and risk managers, design the infrastructure to ensure proper management. Best practices for archiving, records retention, and digital preservation guide decisions on how long to keep data, balancing legal requirements, policy, and future utility.
4. **Use:** The stage where data is organized, transformed, analyzed, and interpreted to generate meaningful insights. Documentation, data pipelines, and reproducible workflows ensure transparency, aiding future users in

[2] Josephson, Anna, and Melinda Smale. "What Do You Mean by 'Informed Consent'? Ethics in Economic Development Research†." *Applied Economic Perspectives and Policy* 43, no. 4 (2021): 1305–29. https://doi.org/10.1002/aepp.13112.

[3] Data, Academic Planning & Institutional Research. "Introduction to the Data Lifecycle." Accessed July 15, 2025. https://data.wisc.edu/data-literacy/lifecycle/.

understanding modifications made during analysis and fostering trust in the data:

5. **Share:** The stage focused on enabling data reuse, replicability, validation, and transparency. It involves data curation through preparation, selection, and contextualization to support effective sharing. Transmission methods and access controls vary, and for long-term availability, responsibility may shift to a trusted repository for preservation and continued accessibility.
6. **Collect/Reuse:** The stage where data literacy skills are applied to locate, evaluate, and understand data while ensuring compliance with access conditions. Effective reuse depends on grasping the data's purpose, history, and lineage to maintain accuracy and relevance.
7. **Destroy:** The final stage where data is securely and permanently removed to prevent unauthorized access or misuse. This involves methods like deletion, shredding, degaussing, or cryptographic erasure, ensuring compliance with legal, regulatory, and policy requirements.
8. **Close out:** Any data remaining is kept according to predefined retention schedules. Only data deemed to be essential should be kept, and retain/archive what is required by law and what might be needed for future use.

3.2 Categories of agricultural data

Agricultural data can be broken down into six categories previously identified by AgGateway and then integrated by the American Farm Bureau Federation into the core Ag Data Transparent principles. [4] These categories are expanded upon in Appendix A.

3.3 Standardizing data formats

With the wide variety of data types, sources, and uses, there is a growing need to provide standardized interfaces and semantics to enable the accurate and efficient exchange of data among systems. As the market matures and the need to share data increases, trading partners realize that developing and maintaining one-off interfaces with each individual trading partner is unsustainable. AgGateway provides the antitrust and intellectual property framework to enable competitors, trading partners, and other interested parties to collaboratively solve these data exchange problems. When stakeholders agree on data exchange basics, they can redirect resources to developing new and innovative tools that leverage data and create value.

3.4 Applications of agricultural data

There are many existing uses for data, ranging from enabling better informed decision making by farm managers to accurately and efficiently reporting on production practices and outcomes. Some uses enable the data originator to generate greater returns in their operation by making better decisions, where other uses are driven by

[4] Ag Data Transparent. "Core Principles." Accessed July 15, 2025. https://www.agdatatransparent.com/principles.

opportunities or requirements from a government or other entity. A review of a simple and complex data sharing use case may help put the various perspectives in context but is out of scope for this document.[5]

3.5 Non-technical factors hindering use of agricultural data

Many data originators hesitate to share information when the benefits and intended use are unclear, leading to uncertainty and reluctance. Additionally, soft or unfunded mandates create indirect pressure—while sharing data isn't required, withholding it can result in lost market access, increased costs, or diminished profitability. In some cases, organizations may deliberately avoid sharing data out of concern that doing so could set a precedent for future regulatory burdens without adequate support or compensation. Addressing these concerns requires clearer governance, transparent incentives, and a data-sharing environment that balances accessibility with sustainability. [6]

[5] Bierman, Don, and Ben Craker. "Who Are the Data Stewards: Moving Data Driven Agriculture Forward." *A Paper from the Proceedings of The, International Society of Precision Agriculture*, July 21, 2024, 14. https://www.ispag.org/resources/publications/proceedings/?action=abstract&id=10206&title=Who+Are+the+Data+Stewards%3A+Moving+Data+Driven+Agriculture+Forward&search=authors.

[6] The ODI. "Our Theory of Change." Accessed July 15, 2025. https://theodi.org/about-the-odi/our-theory-of-change/.

4. Ethical Considerations

4.1 Privacy & confidentiality

Data privacy and confidentiality have become critical concerns in the age of big data, cloud computing, and pervasive technology. As a result, systems need to prioritize protecting data originators' and data consumers' rights to ensure data are used as intended . Below is a breakdown of some key areas of ethical consideration:

4.1.1 Data use in accordance with data originator expectations

One of the primary ethical concerns in data privacy is ensuring that data are used in a manner consistent with the data originator's expectations. When data originators provide their data, they generally expect data to be used for a specific purpose and remain secure from unexpected use which may harm them. Ethical principles include:

- **Informed consent**: Data originators must be clearly informed about how their data will be used, what benefits they can expect in return for the use of their data, and their consent must be obtained for such uses. Ambiguous terms of service or hidden purposes for data collection violate this principle. In the research or medical fields there are often review boards that help define what informed and consent mean, in the agricultural data context these terms are not consistently defined. For the purposes of this paper the below definitions are provided:
    - Informed - the data originator knows what data is being collected, what purposes it will be used for, what parties will have access, and if or how they will be compensated for providing access. This should include an understanding of the lifecycle of the data for the specific use so they know to what extent and when their data can be returned or removed once provided.
    - Consent - the data originator freely, and actively (opts in) agrees to providing access to the data after being informed of the use.[7]
- **Transparency**: Data originators should have clear, accessible information about what data are being collected, why data are being collected, and how data will be processed. Ethical practices require that data exchange partners be open about their intentions and limitations.
- **Respect for autonomy**: Data originators should be provided opportunities to opt-in to having their data utilized, as opposed to assuming that they consent `by default' to have their data collected and analyzed. It is unethical to coerce or manipulate people into sharing data without giving them meaningful alternatives.

[7] *Topic: Informed Consent | American Sociological Association*. n.d. Accessed July 15, 2025. https://www.asanet.org/topic-informed-consent/.

Failing to adhere to these principles leads to breaches of trust and exposes data exchange partners to legal risks, such as violations of privacy laws exemplified by the General Data Protection Regulation (GDPR) and California Consumer Privacy Act (CCPA).

4.1.2 Actively managing data movement

Data movement between data exchange partners or data originators poses potential ethical challenges, especially when it involves sensitive information. Actively managing data movement should prioritize the protection of privacy, security, and consent. Ethical considerations include:

- **Data minimization**: Only the data necessary for a specific purpose should be collected and exchanged. Holding excessive or unrelated information increases the risk of harm in the event of a breach.
- **Secure transmission**: Data exchanged between data originators or data exchange partners must be encrypted and protected against interception or unauthorized access. Data exchange partners should ensure robust security measures are in place during data movement.
- **Accountability and auditing**: Data exchange partners should monitor data movement to detect unauthorized access or misuse. Auditing mechanisms should track who accessed the data and when, ensuring transparency and accountability throughout the process.
- **Third-party risk**: When data are shared with third-party service providers, ethical responsibility extends to ensuring that those providers adhere to similar privacy and security standards. This means vetting partners, establishing clear data-sharing agreements, and regularly auditing compliance.

4.1.3 Anonymization

Anonymization is a critical practice aimed at protecting data originators' identities when data are used for research, marketing, or other purposes. However, ethical issues arise when the process of anonymization is either insufficient or improperly managed. Ethical considerations around anonymization include:

- **Effectiveness of anonymization**: It is not enough to remove obvious identifiers like names and addresses. Data should be anonymized to a level where data originators cannot be re-identified, even when the dataset is combined with other sources. Techniques like differential privacy and homomorphic encryption are being increasingly used to provide stronger anonymization guarantees. [8]
- **Re-identification risks**: Even anonymized data can sometimes be re-identified by analyzing patterns, especially if combined with external datasets. Ethically, data exchange partners must evaluate the potential risk of re-identification and take extra precautions if there is a risk of harm.

[8] Brandao, Luis, and Rene Peralta. "Privacy-Enhancing Cryptography to Complement Differential Privacy." *NIST*, no. post 11 (November 2021). https://www.nist.gov/publications/privacy-enhancing-cryptography-complement-differential-privacy.

- **Data utility vs. privacy**: While anonymization helps protect privacy, it can reduce the utility of data. Ethical decision-making requires balancing privacy with the practical benefits of data use. In contexts like sustainability research, where de-identified data can advance practices and policies, the ethical question becomes how to maximize data utility without unreasonably compromising privacy and ensuring the data originator is informed of the risks.
- **Ongoing monitoring**: Anonymization is not a one-time solution. As technology evolves, new methods of re-identifying anonymized data may emerge. Ethical data exchange partners need to regularly reassess the effectiveness of their anonymization techniques and make adjustments as necessary.

4.2 Ownership & control

In the digital age, the concepts of data ownership and control have taken on increasing importance. As agricultural producers' data are increasingly recognized as valuable assets, questions surrounding who owns the data and who has the right to control data usage are central to ethical debates. The ethical considerations around data ownership and control focus on fairness, consent, and responsibility, balancing the interests of data originators, data exchange partners, and society.

4.2.1 Ownership rights and data originator control

A fundamental ethical issue in data ownership is the extent to which data originators own their data and maintain control over data use. Ethical principles around ownership and control include:

- **Data as personal property**: Ethically, data originators should have the right to claim ownership of the data they generate, such as personal information, preferences, behaviors, and interactions. This perspective emphasizes that data are not just assets to be used by data exchange partners but are a form of property belonging to the originator.
- **Right to access and portability**: Data originators should have the right to access their data held by data exchange partners. In line with ethical norms and privacy regulations like the GDPR, data originators should also have the ability to transfer their data to another data exchange partner if desired, maintaining control over how and where their data are used.
- **Right to erasure (right to be forgotten)**: A key ethical right for data originators is the ability to request deletion of their data. This becomes particularly important in cases where data are no longer needed or where the data's continued existence could harm the data originator (e.g. outdated or misleading information). Ethical data exchange partners should provide clear and accessible paths for data originators to request the deletion of their data and provide assurances or evidence to the requester that the request was completed. This capability is mandated by law in some localities, such as the European Union and California.
- **Informed consent for data use**: Data exchange partners must obtain informed consent from data originators before collecting or using their data. The consent

should be specific, informed, and voluntary, allowing data originators to understand what they are agreeing to and enabling them to revoke consent if their preferences change. Proof of informed consent should be maintained for auditors.

- **Data control versus ownership:** Occasionally ownership and control of data are used interchangeably, but they are different concepts with different implications. The owner of data may have very little control of their data. Likewise there are entities with a great deal of control over how data are used or shared but do not have any actual ownership of the data. When these terms are used it is important to be clear about the differences and the associated responsibilities and limitations of the owner and the controller, these expectations and responsibilities should be clearly defined in the legal agreements between the parties.

4.2.2 Control over data by data exchange partners

When data exchange partners collect and manage data, they assume significant control over how the data are used, shared, and stored. The ethical concerns in this area revolve around ensuring data exchange partners act as responsible stewards of the data rather than exploiters. Key considerations include:

- **Data stewardship**: Data exchange partners must recognize their role as custodians of data, with the responsibility to protect it from misuse, unauthorized access, and breaches. Ethically, this entails implementing robust security and traceability measures and respecting the privacy rights of data originators.
- **Power imbalance**: Often, there is a significant power imbalance between data originators and data exchange partners, where originators have limited visibility and control over how their data are used once collected. Ethically, data exchange partners should not exploit this imbalance by using data for purposes beyond what was originally agreed upon or expected, even if technically legal.
- **Transparency and accountability**: Ethical data exchange partners must be transparent about how they collect, use, and share data. This includes making privacy policies clear, explaining what data are collected, who data are shared with, and for what purposes. Additionally, data exchange partners should be accountable for any misuse or mishandling of data, taking responsibility for data breaches or ethical violations. Institutions storing or utilizing data should establish a publicly-accessible point of contact.
- **Monetization and exploitation**: Many data exchange partners benefit from originator data, raising ethical concerns about whether data originators should be compensated or receive some benefit for the use of their data. Ethical considerations suggest that data exchange partners should not monetize originators' data without transparently providing them with some form of reciprocity, whether that be compensation, services, or other benefits.

4.2.3 Ethical responsibility for shared data

In cases where data are shared among multiple data exchange partners or across platforms, ethical concerns arise around control and responsibility for protecting the data. Key considerations include:

- **Clear data ownership terms**: When data are shared between entities, such as between a data exchange partner and a third-party vendor, there must be clear agreements regarding who owns the data and who has the right to control it. This reduces disputes and ensures all parties understand their responsibilities.
- **Shared responsibility for security**: Data exchange partners that share data with third parties must ensure that those third parties maintain the same level of security and ethical responsibility as the original collector. Ethical principles suggest that data exchange partners should not absolve themselves of responsibility once data are handed off but should instead engage in careful oversight and monitoring.
- **Preventing unethical data sharing**: Data cannot be ethically shared with third parties that intend to use it for undisclosed purposes, discriminatory practices, or unauthorized marketing. Data exchange partners must establish strict guidelines for how data can be shared and with whom, ensuring that any shared data remains aligned with the original consent provided by the data originator.

Ethical considerations of data ownership and control focus on safeguarding the rights of data originators, ensuring data exchange partners act as responsible stewards, and addressing the challenges of jurisdictional boundaries. Upholding these principles ensures that data are handled with respect, transparency, and fairness, fostering trust between data originators and data exchange partners while protecting privacy and fostering continued innovation.

4.3 Equity & fairness

Data equity and fairness are critical ethical considerations in the collection, use, and analysis of data. As data increasingly influences decision-making in agriculture, ensuring fairness and equity in these processes is essential to avoid reinforcing existing biases or creating new forms of discrimination. The ethical challenges revolve around the fair treatment of data originators of all types and sizes, mitigating bias, and ensuring equal access to the benefits of data-driven innovations.

4.3.1 Fair representation in data

One of the most significant ethical challenges in data equity is ensuring that the data collected represents all groups fairly. Without adequate representation, data can perpetuate systemic biases and lead to unfair outcomes for underrepresented groups. Ethical considerations include:

- **Inclusivity in data collection**: Ethical data collection practices should ensure that data are gathered from a diverse range of data originators. When certain

groups or types of operations are underrepresented in datasets, algorithms and analyses based on those datasets may be skewed or biased, leading to recommendations that are inaccurate or of limited applicability.

- **Avoiding sampling bias**: Sampling bias occurs when certain groups are systematically excluded from the data collection process, leading to skewed results. Ethical considerations demand that researchers and practitioners actively work to identify and correct for sampling biases to ensure that conclusions drawn from the data are fair and applicable across diverse populations.
- **Balancing different interests**: Ethical data practices must strike a balance between majority and minority interests. Decisions based on data should not disproportionately benefit or harm one group over another. Special attention should be given to ensuring that marginalized or vulnerable data originators are fairly represented.

4.3.2 Mitigating algorithmic bias

Algorithmic bias occurs when the algorithms used to analyze data produce biased or unfair results, often reflecting the biases present in the original data or algorithm design. Ethical considerations around algorithmic bias are vital in ensuring fairness in data-driven decision-making. Key issues include:

- **Bias in training data**: Machine learning algorithms are only as fair as the data they are trained or tested on. If training data contain historical biases or reflect societal inequalities, the algorithm may perpetuate or even amplify those biases. Ethical considerations require researchers and developers, regardless of affiliation or objective, to critically assess their datasets and apply techniques to mitigate bias if needed, such as re-sampling, re-weighting, or fairness-aware machine learning.
- **Transparency in algorithm design**: Ethical data practices demand transparency in how algorithms are designed and used. This means being open about the inputs, methodologies, and potential biases of algorithms, and allowing for external scrutiny. Transparency also enables affected data originators to understand how decisions are made and advocate for fairer outcomes.
- **Impact on decision-making**: Algorithms are often used to make decisions where biased outcomes can have serious consequences. Ethical considerations require that these algorithms be regularly audited to ensure they produce equitable results. Independent validation and verification (IVV) by an external and unbiased party is a best practice to ensure systems are properly designed. Additionally, decision-makers should not rely solely on algorithmic outputs but should consider human oversight to prevent unfair outcomes.

4.3.3 Determining access to data and resources

In addition to the responsibilities of data exchange partners in stewarding originator data, these organizations have a broader societal role in promoting policies and

practices which improve access for all data originators to be able to realize the benefits of data and data-driven technologies. Key considerations include:

- **Digital divide**: The digital divide refers to the gap between those who have access to data, digital technologies, and the internet, and those who do not. Closing this divide helps ensure that all data originators have equitable access to data resources and the opportunities that result, such as education and economic advancement.
- **Data democratization**: Data democratization is the process of making data accessible and understandable to a wide range of users, not just experts or data exchange partners. Ethically, data should not be concentrated in the hands of a few powerful entities. Instead, access to data should be shared broadly to ensure that all data stakeholders can benefit from data insights. This could include ensuring public institutions like universities have a means to access data for research and educational purposes with adequate permissions and security.
- **Data literacy**: Data literacy is the ability to understand and use data effectively. Without adequate data literacy, data originators may be excluded from the decision-making processes that shape their lives. In addition to the limitations on connectivity and access to data related to the digital divide, there is also an educational component. People require training and education to be able to better understand and use data. This is a key aspect of informed consent, in ensuring data originators have sufficient understanding of data systems and technologies and potential uses to be able to make decisions about providing access to their data. Data literacy also helps data originators be better positioned to derive value from their data.
- **Balancing commercial interests with social good**: Many data exchange partners collect and analyze data for profit, which can create ethical tensions when commercial interests conflict with broader social good. Ethical considerations call for balancing profit motives with a commitment to advancing society in general. Data exchange partners should consider how their data practices impact the public and whether they are contributing to the well-being of all data originators.
- **Engaging data originators in data use**: Ethical data practices involve actively engaging the data originators whose data are being collected and used. This means involving data originators in decision-making processes, allowing them to voice their concerns, and ensuring they have a say in how their data are utilized. By prioritizing data originator engagement, data exchange partners can better ensure that their data practices align with the needs and values of the data originator themselves. Having engagement and input from data originators during their creation makes policies more practical and acceptable to all parties.

In summary, ethical considerations of data equity and fairness focus on ensuring that data collection, analysis, and usage are inclusive, unbiased, and benefit all data stakeholders. Addressing issues of fair representation, mitigating algorithmic bias, ensuring equal access to data, and using data for social good are central to creating a

data-driven society that is just and equitable. Ethical data practices not only prevent harm but also promote trust and long-term sustainability in data-driven initiatives. Trust is critical since it can be easily broken and very difficult to rebuild; parties need to prioritize proper and transparent communication.

5 Stewardship Challenges

A useful definition of data stewardship is the set of practices that enable an organization's data to be accessible, usable, safe, and trusted. It includes overseeing every aspect of the data lifecycle: from creating, preparing for using (i.e., preprocessing), analyzing, storing or archiving, and deleting data. These practices are followed in accordance with an organization's data governance principles and policies.

As organizations increasingly rely on data to drive decision-making, they face a variety of challenges spanning technical, legal, and ethical domains. These challenges include the integration of disparate data sources, ensuring data quality and accuracy, managing conflicting legal obligations, safeguarding intellectual property, and ensuring fair and ethical use of data. Addressing these challenges is vital for organizations aiming to leverage data while maintaining trust and regulatory compliance.

5.1 Technical challenges

Not all the issues related to the ethical use of data are driven by perceptions of fairness or even compliance with policies or regulations. There are many technical aspects that have a large impact on how data are shared from system to system, or if data consumers are able to derive value from a dataset while meeting the ethical expectations.

5.1.1 Data integration and interoperability

Being able to effectively use data is heavily dependent on a receiving system being able to ingest data so it can be used. Often this means getting data into a format the system is able to use and aligning on common terms and concepts so data from multiple sources can be accurately combined or compared.

- **Heterogeneous data formats**: Data often come from various sources in different formats, making it difficult to integrate seamlessly. This challenge is heightened when data originate from incompatible systems, including databases, applications, or platforms that lack standardized communication protocols.
- **Semantic discrepancies**: Even when data formats align, the meaning and context of data fields may differ. For example, the same term may be used differently across datasets (e.g., "customer" could mean a retail buyer in one system and a business client in another).
- **Lack of unified standards**: With diverse schemas, metadata, and taxonomies in use, data interoperability is difficult, especially when integrating datasets across industries or regions with varying regulations and practices.

5.1.2 Data quality and accuracy

The old adage of garbage in, garbage out is especially true with data. If the data used are inaccurate or out of date, any insights or analysis generated by a system will not be reliable. For data to be valuable, time and effort needs to be made in the collection of the data, often something incumbent upon the data originator. If they do not put the effort in up front to create quality data, it will cause issues and potentially make the data unusable downstream.

- **Data inconsistency and duplication**: Inaccurate data entries, duplicates, or incomplete records degrade data quality. Handling vast amounts of data from multiple sources amplifies these issues, requiring validation and cleansing processes to ensure high quality.
- **Timeliness and update lag**: Ensuring that data are up-to-date is critical, especially in real-time systems. Delays in data synchronization across systems can cause discrepancies, reducing the accuracy of analysis and quality of decision-making.
- **Source reliability**: When aggregating data from third parties or external platforms, assessing the accuracy and reliability of these sources is a major challenge. Incorrect or biased data can skew outcomes.

5.1.3 Tension between data preservation and rights to data deletion

People like to hold on to things as long as possible if they perceive that these things may one day have value. With modern technology it can be incredibly easy and cost effective to archive a huge amount of data. However, storing data for long periods of time does have ramifications for the ethical use and management of that data which should be considered.

- **Data retention vs. compliance**: Many organizations aim to retain data for long periods to support research, historical analysis, and to generate business insights. However, regulations such as the GDPR and CCPA grant individuals the right to request the deletion of their data. Balancing these conflicting needs is difficult, especially for legacy systems not built with privacy regulations in mind or situations where data traceability is challenging.
- **Permanent deletion processes**: Implementing processes that ensure complete and irreversible deletion of data is technically challenging. It requires changes to backups, archival systems, and log files, which may retain traces of data for longer than intended.

5.1.4 Separation of personally identifiable information

Most regulations regarding data protection and rights focus on personally identifiable information (PII). In the agricultural industry there is a lot of information that can be used independently from PII to enable analysis, service insights and inform decisions. For that to happen, systems need to be able to reliably separate PII from other data.

- **Data entanglement**: PII is often deeply intertwined with non-sensitive data, making separation difficult. Complex data structures, especially those in large relational databases, can embed PII in multiple locations or across interconnected records.
- **Granular access control**: Ensuring that only authorized individuals or systems can access PII while allowing broad use of anonymized or de-identified data requires sophisticated access control mechanisms. Maintaining such controls consistently across all data pipelines and platforms can be error-prone.
- **Re-identified PII**: In some cases, non-PII data can be combined or analyzed in ways that inadvertently reveal PII, necessitating continuous monitoring and updating of how data are handled.

5.1.5 Data sustainability

As with many technologies data can have a shelf life if not properly managed. System providers need to be aware of changes in technology and have plans and funding in place to ensure the data they steward are able to be used not only today but also in the future. Or, if the data will at some point no longer be used, plans should also be in place for when and how these data will be archived or removed from a system.

- **Long term system viability:** Systems need to plan for how they will be managed for the long term. Decisions need to be made and resources allocated to ensure the system is able to effectively support ongoing stewardship principles as technologies advance, creating enduring value to the data. This requires thoughtful financial planning to ensure that appropriate long-term support resources are available.
- **Data lifecycle planning:** Data systems need to be designed with the full lifecycle of data in mind. Systems should not only be able to ingest and efficiently organize data, but also provide a means for long term storage, to update and translate data as technologies advance, and to mark data as obsolete or remove it from the system when it is no longer useful.
- **Data interoperability over time:** Data standards and interoperability were discussed previously, but the issue is not a one time task. As systems, technologies, and the broader landscape of our connected world continue to grow and evolve, data and systems must adapt. Regular reassessment of data formats and system interoperability are key to data retaining its value and utility.

5.2 Legal & regulatory

Agricultural data are often in a grey area when it comes to regulations. Often data falls somewhere between PII, intellectual property, or general business information. Depending on the classification there are very different laws and requirements that are applicable. Additionally these regulations are not consistent across the world, nor fixed. There are continually new laws being passed with some countries or sub-national jurisdictions taking fundamentally different approaches to regulation. This leads to a very complicated environment for any organization to navigate, especially if they have operations in multiple countries.

5.2.1 Compliance with rapidly changing and conflicting data protection laws

Primarily focused on PII, there are rapidly changing regulations around the world.

- **Diverse legal frameworks**: Different regions have varying data protection laws, such as the General Data Protection Regulation (GDPR) in the European Union, the California Consumer Privacy Act (CCPA) in the U.S., and other national or sub-national (state) privacy laws. These regulations may impose conflicting requirements, making compliance across jurisdictions complex or even impossible.
- **Rapid evolution of laws**: Data privacy regulations are continually evolving in response to new technologies, threats, and societal concerns. Organizations must stay up-to-date with these changes and adjust their policies and procedures to remain compliant. This can be resource-intensive and difficult.
- **Arbitration agreements**: Many data use agreements include clauses requiring parties to resolve disputes through arbitration. Arbitrators have a bias to rule in favor of entities who pay them. Dispute resolution should be performed in a manner that is equitable for all interested parties.
- **Data sovereignty**: Data sovereignty refers, in a legal context, to the idea that data are subject to the laws and governance structures of the country where the data are collected or processed. Ethically, data exchange partners should respect local regulations and not attempt to bypass stricter data protection laws by storing or processing data in more lenient jurisdictions. They should be transparent with data originators about where data are being stored or shared.
- **Cross-border data transfers**: Laws like GDPR place restrictions on data transfers to countries that do not offer an "adequate" level of protection. Navigating international data flows and ensuring lawful data transfers, such as via standard contractual clauses (SCCs) or Binding Corporate Rules (BCRs), is a significant challenge for global organizations. Ethical data exchange partners need to ensure that data remain protected to a standard that satisfies the requirements of each relevant jurisdiction.

5.2.2 Intellectual property rights

While intellectual property rights have a long track record, their application to data is relatively new. While some forms of data easily integrate into existing regulatory frameworks, other aspects of agricultural data are not as obvious, requiring organizations to use care when developing new systems and utilizing data.

- **Data ownership ambiguity**: Determining who owns data can be unclear, particularly when data are collected, processed, and transformed by multiple parties. For example, in research collaborations, questions may arise about whether the raw or derived data are owned by the data subject, the organization collecting the data, or the party analyzing it. Similar ambiguity holds when data are collected by a third party when the land is managed by one party and owned by yet another party

- **Protection of proprietary data & trade secrets:** Intellectual property laws may not always offer adequate protection for data, especially in cases where the data have been anonymized or aggregated. Organizations need to ensure that data sharing agreements include robust provisions to protect proprietary data from misuse or unauthorized disclosure.
- **Data reuse and commercialization**: When data are shared for specific purposes (e.g., research or business analytics), it may later be repurposed or monetized in unanticipated ways. Organizations must define clear guidelines in contracts about permissible use and prevent unauthorized data commercialization.

5.2.3 Agreement type definitions and purpose

Agriculture has a strong history of agreements being made by a handshake between two parties that know and trust each other. With the adoption of more and more technology, legal agreements are becoming more common and in many cases more necessary. Having an agreement where expectations from both parties are clearly stated is a good first step, but with many organizations there are multiple parties and several layers of agreements that are necessary.

- **Varying contract types and obligations**: Organizations often enter into a variety of data sharing agreements, including data licensing agreements, non-disclosure agreements (NDAs), and data processing agreements (DPAs). Each agreement type comes with its own legal implications and obligations, such as confidentiality, scope of use, and liability, which must be clearly defined and adhered to.
- **Clarity on data usage purposes**: It is essential that agreements specify the exact purpose for which data will be used. Regulatory frameworks like GDPR require that data be collected and processed for specified, legitimate purposes. Ensuring that contracts explicitly define permissible uses of data helps avoid legal disputes and ensures compliance with purpose limitation principles.
- **Data processing obligations**: In situations where third parties are involved in data processing (e.g., cloud providers or analytics firms), organizations must ensure that these processors adhere to relevant legal requirements. This includes ensuring that processors follow data protection laws, maintain data security, and respect the terms outlined in the processing agreements.

5.2.4 Mitigating risks of data misuse

- **Privacy violations and surveillance**: One of the biggest social and ethical risks in data stewardship is the potential misuse of data by exploiting this data to gain unintended information about a data originator. Organizations must ensure that the data they collect are not used to infringe on individuals' privacy or to track behavior without consent. Ethical data stewardship requires strong safeguards against using data beyond the original intent.

- **Algorithmic bias and discrimination**: This topic was discussed above in the Equity and Fairness section, but bears mentioning again as a significant technical challenge.
- **Data commodification and exploitation**: As data become increasingly valuable, there is a growing risk of exploitation, where originators' data are treated as a commodity to be exploited without their awareness or fair compensation. Ethical challenges arise when organizations prioritize profit over user rights, potentially leading to unethical data-sharing practices, exploitation, or manipulation.

6 Recommendations

6.1 Data agreements

Data privacy and use policies create an agreement about how user data are collected, used, disclosed, and managed between data originators, data exchange partners, and data exchange partners' relationships to third parties. Notices help to clarify expectations, understandings, and policies regarding agricultural data sharing. Data agreements and notices are binding contracts. Things to consider when drafting data agreements include making the documents as transparent, or easily understood and unambiguous, as possible. These agreements should also be publicly accessible so a potential user can review any relevant policies and agreements prior to purchasing or using a product or service. Users should also be notified if any changes are made to agreements or policies they have accepted. This includes requiring users to re-opt-in if there have been substantive changes to how data will be collected, used, disclosed, or managed.

Having policies and agreements is a key first step, but organizations also need to be sure the commitments made in an organization's legal documents are understood internally. Leadership within the organization must also be committed to implementing data ethics by design. Nearly every job function from developers, customer support, sales, and marketing may be impacted by the content of the policies, so it is imperative the organization provides resources and training to ensure everyone understands what is included and are notified if and when any changes take place.

Most organizations will have multiple legal agreements, policies, and statements that generally reference each other and cover different aspects of data sharing, use and related topics. The construction of these documents is mostly driven by preferences of the team or lawyers involved in their creation. However, there are some examples to consider including the Ag Data Transparent template data use agreement[9] or the OpenTEAM Boilerplate Data Hosting and Storage Agreement[10].

Regardless of the name, number, or structure of agreements, Appendix D includes a good starting checklist of topics that should be considered in the creation or modification of any legal agreements between a system provider, the customer/user, or third party data processors.

[9] Ag Data Transparent. “Ag Data Use Model Agreement.” Accessed July 15, 2025. https://www.agdatatransparent.com/model-agreement.
[10] “Boilerplate Data Hosting and Storage Agreement - Ag Data Use Agreements.” Accessed July 15, 2025. https://openteam-agreements.community/hostingandstorage/.

7 Conclusion

In summary, the complex landscape of data stewardship demands careful navigation of technical, legal, and ethical challenges. Organizations must find ways to integrate diverse datasets, preserve data quality, and adhere to conflicting legal frameworks, all while protecting individual privacy and upholding ethical standards. By addressing these issues, data stewards can foster inclusivity, prevent data misuse, and ensure that data management aligns with both regulatory requirements and societal values. A strategic, responsible approach to data stewardship is essential for building trust and ensuring that data is used in ways that benefit all stakeholders.

For further inquiries, visit http://www.aggateway.org or contact member.services@aggateway.org

8 References

Ag Data Transparent. n.d. “Ag Data Use Model Agreement.” Accessed July 15, 2025. https://www.agdatatransparent.com/model-agreement.

Ag Data Transparent. n.d. “Core Principles.” Accessed July 15, 2025. https://www.agdatatransparent.com/principles.

American Sociological Association. n.d. “Topic: Informed Consent.” Accessed July 15, *2025*. *https://www.asanet.org/topic-informed-consent/*.

Becker, Sophia, Don Bierman, Alexander Bucksch, et al. “The NAPDC: Stakeholder Input and Strategic Directions.” *Preprint, OSF*, December 19, 2023. https://doi.org/10.31219/osf.io/tkg96.

Bierman, Don, and Ben Craker. “Who Are the Data Stewards: Moving Data Driven Agriculture Forward.” Paper presented at the *International Conference on Precision Agriculture*, July 21, 2024. https://www.ispag.org/resources/publications/proceedings/?action=abstract&id=10206&title=Who+Are+the+Data+Stewards%3A+Moving+Data+Driven+Agriculture+Forward&search=authors.

Brandao, Luis, and Rene Peralta. “Privacy-Enhancing Cryptography to Complement Differential Privacy.” *NIST*, no. post 11 (November 2021). https://www.nist.gov/publications/privacy-enhancing-cryptography-complement-differential-privacy.

Carroll, Stephanie Russo, Ibrahim Garba, Oscar L. Figueroa-Rodríguez, et al. “The CARE Principles for Indigenous Data Governance.” *Data Science Journal* 19, no. 1 (2020). https://doi.org/10.5334/dsj-2020-043.

University of Wisconsin–Madison. n.d. “Data, Academic Planning & Institutional Research. Introduction to the Data Lifecycle” Accessed July 15, 2025. https://data.wisc.edu/data-literacy/lifecycle/.

Food and Agriculture Organization of the United Nations. n.d. “EU Code of Conduct on Agricultural Data Sharing by Contractual Agreement.” Accessed July 15, 2025. https://www.fao.org/family-farming/detail/en/c/1370911/.

Fruchterman, Jim, Katy McKinney-Bock, and Steve Francis. 2024 “Better Deal for Data White Paper.” *Better Deal for Data*, April 2, 2024. https://bd4d.org/better-deal-for-data-white-paper/.

Global Open Data for Agriculture and Nutrition. “The Codes of Conduct.” Accessed July 15, 2025. https://godan-world.netlify.app/www.godan.info/codes/list.html.

Josephson, Anna, and Melinda Smale. “What Do You Mean by ‘Informed Consent’? Ethics in Economic Development Research” *Applied Economic Perspectives and Policy* 43, no. 4 (2021): 1305–29. https://doi.org/10.1002/aepp.13112.

Lin, Dawei, Jonathan Crabtree, Ingrid Dillo, et al. “The TRUST Principles for Digital Repositories.” *Scientific Data* 7, no. 1 (2020): 144. https://doi.org/10.1038/s41597-020-0486-7.

McKinsey & Company. n.d. “Data Ethics: What It Means and What It Takes.” Accessed July 15, 2025. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/data-ethics-what-it-means-and-what-it-takes.

Open Data Institute. n.d. “Assuring Data Practices.” Accessed July 15, 2025. https://theodi.org/about-the-odi/our-theory-of-change/.

OpenTEAM. n.d. “Boilerplate Data Hosting and Storage Agreement – Ag Data Use Agreements.” Accessed July 15, 2025. https://openteam-agreements.community/hostingandstorage/.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, et al. “The FAIR Guiding Principles for Scientific Data Management and Stewardship.” *Scientific Data* 3, no. 1 (2016): 160018. https://doi.org/10.1038/sdata.2016.18.

Appendix A: Categories of Agricultural Data

Farm management data

- Business operations
    - Financial & tax
    - Operating & land loans
    - Office files
    - Capacity / timing data
    - Farm labor and contracts
    - Human resources
- Supply chain data
    - Point of sale data
    - Partnerships
    - Customer data
    - Supplier data
- Transport and storage data
- Commodity prices (input and output pricing)
- Reporting and compliance data

Agronomic data

- Crop seed data
    - Genetics data
    - Production attribute data
- Planting data
    - Recommendation data
    - Prescription data
    - Work order data
    - As planted data
- Yield data
    - Attribute data
    - Quality data
- Disease and pest management data
    - Crop protection data (herbicide, insecticide, fungicide)
    - Crop protection use and application rates data
    - Biological (herbicide, insecticide, fungicide)
    - Prescription data
    - Work order data
    - As treated / as applied data
- Crop nutrition data
    - Sampling data
    - Application and use of biological fertilizer and soil amendment data
    - Application and use of synthetic fertilizer data

    - Prescription data
    - Work order data
    - As Treated / as applied fertility treatment data
- Pollinators
- Harvested commodity preservation and handling data

Land data

- Conservation data
- Tillage practice data
- Water management data
- Soil and fertility data
    - Soil test data
    - Nutrient management data
    - Waste management data
- Environmental and ecological data
    - Watershed data
    - Topological data
    - Elevation data and derivatives
    - Drainage data
    - Irrigation data
- Geospatial information system (GIS), global navigation satellite system (GNSS), & field boundary data
    - Ground-based machine data
    - Uncrewed aerial system (UAS) data
    - Sensor collection system (EC/EM) data
    - Remote sensing including radar, spectral, & lidar data

Machine data

- Rolling and fixed assets data
- Machine health and operation technique data
    - Energy & fuel use
    - Machine load
    - Equipment reference data
    - Equipment function
    - Milk parlor equipment information (system, not animal data)
- Other livestock equipment information (robots, sensory equipment, etc)

Climate and weather data

- Weather stations
- Soil probes
- Sensor data
- Temperature and humidity data (on farm)

Livestock data

- Identification and pedigree information
- Performance information (production, reproduction, longevity, feed efficiency, rib eye area, etc.)
- Quality information (somatic cell, milk components, marbling, etc.)
- Health information (health events, treatments, etc.)
- Genetic and genomic information
- Management information on individual animals from on-farm management machines or systems (milking speed, robot efficiency, feed intake, rumination, animal weights, etc.)

In addition to these general classifications, one must also consider the form in which this data exists. For example:

- **Raw data** - data as-captured, without interpretation, modification, or calculation
- **Processed data** - data which may have been modified, contextualized, converted, interpreted, combined with other data, checked for quality, etc.
- **Anonymized data** - data which has been processed to remove data elements which might allow the data to be connected back to a particular entity as the source or subject
- **Aggregate data** - raw or processed data which is combined from multiple sources and only the combination remains; data aggregation is one possible path to anonymization
- **Derivative data** - processed data which has been modified and does not include the source data, perhaps via calculation, interpretation, or combination of multiple sources; aggregate data is a type of derivative data

Appendix B Data Use Categories

Category	Description	Typical Stakeholders	Key Ethical Considerations
Operational Use (First-Party)	Direct use by the data producer	Farmers, data collecting platforms (OEM), data analysis tools (FMIS)	Full control by data producer; presumed consent.
Advisory & Agronomic Services	Use by third-party advisors in support of farmer operations.	Farmers, agronomists, consultants	Informed consent, data-sharing transparency
Regulatory Compliance	Data submitted to meet contractual or governmental reporting requirements	Government agencies, auditors	Purpose limitation, data minimization, transparency
Sustainability & Environmental Accounting	Use to evaluate and report environmental impacts and ecosystem services.	GHG reporting programs, downstream consumers/processors, government agencies	Transparency, informed consent, informed use, anonymization
Legal, Audit & Dispute Resolution	Use in auditing or resolving disputes, typically under legal or contractual frameworks.	Regulators, trading partners, lawyers	Data use agreements, informed use, access controls.
Supply Chain Traceability	Data used to track origin, inputs, and practices across the value chain	Processors, retailers, consumers	Transparency, data integrity, informed consent, access control
Financial & Risk Management	Data used for insurance, credit scoring, or investment decisions	Banks, insurers, investors	Fairness, non-discrimination, data accuracy, informed use
Bench-marking & Performance	Aggregated data used to compare performance across farms or regions	Producer groups, co-ops, bench-marking services	Anonymization, aggregation ethics, benefit-sharing, contextual integrity
Research & Innovation	Data used in academic or private R&D to improve practices or develop new technologies	Universities, startups, public research bodies	Informed consent, open access vs. IP rights, equity in benefit-sharing

Market Intelligence	Data used to assess trends, pricing, or demand in agricultural markets	Traders, analysts, agribusinesses	Transparency, data asymmetry, market manipulation risks, access equity
Product Development & Sales	Data used for R&D to tailor or market products and services to farmers	Ag input companies, tech providers	Consent, bias, data monetization transparency
Infrastructure & Resource Planning	Data used to plan rural infrastructure, water use, or land management	Governments, NGOs, utilities	Public interest vs. private use, data equity, community consent
Public Policy & Advocacy	Data used to inform policy decisions or support advocacy efforts	NGOs, policymakers, advocacy groups	Representation, bias, transparency, accountability
Education & Extension	Data used in training, outreach, or extension services	Extension agents, educators, farmer networks	Contextual relevance, accessibility, data literacy

Appendix C: Stewardship Principles

C.1 FAIR principles: Findability, Accessibility, Interoperability, Reuse[11]

The FAIR principles provide a framework for managing data to maximize its societal value.

- Findability: Data must be easily located through standard search mechanisms, with clear metadata and identifiers that enable users to discover it effectively.
- Accessibility: Data should be accessible under clear conditions, ensuring that authorized users can retrieve it with minimal barriers, often through open standards or authenticated processes.
- Interoperability: Data must be compatible with other datasets and systems, allowing seamless integration and use across various platforms and technologies through common formats and vocabularies.
- Reuse: Data should be prepared for future use, with well-documented and openly shared information that allows others to understand, replicate, and build upon it.

Ultimately, the FAIR principles help make data more valuable and impactful by fostering openness and collaboration.

C.2 CARE principles: Collective benefit, Authority to control, Responsibility, Ethics[12]

The CARE principles focus on the ethical management of data, particularly for Indigenous and marginalized communities, to ensure fairness and respect in data stewardship.

- Collective Benefit: Data should be used to support the welfare of communities, ensuring that the benefits of data use and sharing are equitably distributed to improve societal well-being.
- Authority to Control: Communities, particularly Indigenous groups, must have the right to govern how their data is collected, used, and shared, asserting their sovereignty over data concerning their people and culture.
- Responsibility: Data stewards have an obligation to ensure that data is handled responsibly, safeguarding against misuse and fostering trust through transparency, accountability, and care.

---

[11] Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, et al. "The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data* 3, no. 1 (2016): 160018. https://doi.org/10.1038/sdata.2016.18.

[12] Carroll, Stephanie Russo, Ibrahim Garba, Oscar L. Figueroa-Rodríguez, et al. "The CARE Principles for Indigenous Data Governance." *Data Science Journal* 19, no. 1 (2020). https://doi.org/10.5334/dsj-2020-043.

- Ethics: Ethical considerations must guide data practices, ensuring that data is used in ways that respect the dignity, rights, and cultural values of the individuals and communities it represents.

The CARE principles emphasize the importance of balancing data utility with respect for community values, promoting social justice and equity in data governance.

C.3 TRUST principles: Transparency, Responsibility, User focus, Sustainability, Technology[13]

The TRUST principles are designed to ensure the trustworthy management of data repositories, fostering confidence in data stewardship.

- Transparency: Data repositories should operate with clear policies, procedures, and workflows, ensuring openness about how data is curated, managed, and shared.
- Responsibility: Organizations and individuals managing data must be accountable for ensuring its integrity, security, and ethical use, taking ownership of their role in data governance.
- User Focus: Data management should prioritize user needs, ensuring that data is easy to access, use, and understand, with services designed to meet the needs of diverse user groups.
- Sustainability: Data repositories must be managed with long-term goals in mind, ensuring the data remains available, accessible, and relevant over time through careful planning and resource allocation.
- Technology: The appropriate use of technology is essential for ensuring data security, accessibility, and interoperability, supporting the efficient management and sharing of data.

The TRUST principles emphasize the importance of reliable and ethical data management, promoting confidence and long-term utility in data stewardship practices.

C.4 GODAN (Global Open Data for Agriculture and Nutrition codes of conduct)[14]

The Global Open Data for Agriculture and Nutrition (GODAN) Codes of Conduct provide a set of guidelines aimed at promoting responsible and ethical use of open data in the agricultural and nutrition sectors. These codes are designed to ensure that open data is

[13] Lin, Dawei, Jonathan Crabtree, Ingrid Dillo, et al. "The TRUST Principles for Digital Repositories." *Scientific Data* 7, no. 1 (2020): 144. https://doi.org/10.1038/s41597-020-0486-7.
[14] "The Codes of Conduct." Accessed July 15, 2025. https://godan-world.netlify.app/www.godan.info/codes/list.html.

shared and utilized in ways that benefit farmers, researchers, policymakers, and society at large.

- Inclusivity and Participation: The codes emphasize the importance of including diverse stakeholders, particularly smallholder farmers, in data-sharing initiatives, ensuring that data benefits are distributed equitably.
- Transparency and Accountability: Data should be openly available, with clear guidelines on its collection, use, and sharing. This fosters trust among data users and contributors by ensuring transparency.
- Privacy and Security: While promoting open data, the codes stress the need to protect sensitive information, particularly personal and confidential data, ensuring it is used ethically and securely.
- Empowerment and Capacity Building: GODAN encourages building the capacity of stakeholders to access, use, and understand open data, helping communities leverage data for better decision-making and innovation.

Overall, the GODAN Codes of Conduct guide ethical, fair, and secure use of open data in agriculture and nutrition, fostering global collaboration and innovation while protecting rights and interests.

C.5 Ag Data Transparent

The Ag Data Transparent project, led by the American Farm Bureau Federation (AFBF), aims to ensure transparency and security in the collection, use, and sharing of agricultural data by establishing clear guidelines for ag tech providers.

Ag Data Transparent Principles:

- Ownership: Farmers retain ownership of their data.
- Collection, Access and Control: Farmers have control over the collection and access of their data.
- Transparency and Consistency: Clear and consistent policies on data usage and sharing.
- Portability: Farmers can easily transfer their data.
- Privacy and Security: Protection of farmers' data privacy and security.
- Contract Terms: Fair and transparent contract terms.
- Disclosure: Full disclosure of data use and sharing policies.
- Review: Regular review of data usage policies.
- Choice: Farmers have choices regarding data sharing.
- Value: Ensuring farmers benefit from their data.

C.6 Open Data Institute guide to data practices[15]

[15] The ODI. "Assuring Data Practices." Accessed July 15, 2025. https://theodi.org/what-we-do/consultancy-and-products/assuring-data-practices/.

The Open Data Institute (ODI) guide to data practices provides a framework for ethical data use, helping organizations identify and manage ethical issues in data projects. They established nine key organizational data practices:

- Accountability: Open and transparent oversight and accountability structures with clear roles and responsibilities for data.
- Privacy: Open and transparent processes for handling and sharing personal information legally.
- Security: Open and transparent processes for handling and sharing information securely.
- Standardization: Open and transparent processes outlining why and how data is collected, used and shared.
- Resourcing: Open and transparent plans and funding for the ongoing management and maintenance of data.
- Capability: Open and transparent ability to implement data processes, including both technological and human.
- Engagement: Open and transparent approaches to engagement and participation with data providers and users.
- Ethics: Open and transparent processes that outline how data is handled in accordance with a defined ethical framework.
- Permissions: Open and transparent processes for managing the permissions under which data is consumed and shared

C.7 The NAPDC: stakeholder input and strategic directions[16]

The National Agricultural Producers Data Cooperative (NAPDC) is a USDA-NIFA funded project focused on developing a neutral and secure data repository for agricultural producers, universities, and nonprofit entities to foster agricultural innovation, support technological progress, and enhance production efficiencies and environmental stewardship. In 2023 they prepared a report summarizing the strategic priorities identified at the 2023 conference for the future of the framework, many of which include principles and best practices related to data stewardship..

Major Takeaways

- Trust and transparency are essential and achievable. One way these can be cultivated and sustained is by using current and future best practices for security, privacy, and provenance while pursuing a robust communication and engagement strategy.
- Data literacy and technology education play a major role in adoption on two fronts: data literacy for agricultural producers helps them understand the impact of their choices regarding data management and access, and technology

[16] Becker, Sophia, Don Bierman, Alexander Bucksch, et al. "The NAPDC: Stakeholder Input and Strategic Directions." Preprint, OSF, December 19, 2023. https://doi.org/10.31219/osf.io/tkg96.

education will be vital for precision agriculture adoption and for new entrants into the agricultural producer space.
- Enabling analytics or predictive analyses of any kind will involve the interoperability of data, which means defining metadata standards of various kinds depending on sector, and developing tools that help identify and apply metadata uniformly.
- User-driven development and interoperability of hardware, software, and data will encourage additional agricultural producers and service providers to engage in framework development and adoption.
- Participation can be incentivized by developing a robust neutral platform that provides a clear value proposition for agricultural producers. This involves data tools and cyber-infrastructure that are easy to use, supported, and dovetail with agricultural producer needs.
- Future incentives for participation will require additional funding opportunities related to accessibility, productivity, carbon capture, precision management, and sustainability.
- Coalescence of funding opportunities and research activities around common data sharing platforms and tools is essential to avoid unnecessary duplication of effort.

Link for more information: https://agdatacoop.org

C.8 OpenTEAM: agricultural data use documents[17]

The Open Technology Ecosystem for Agricultural Management (OpenTEAM) is part of Wolfe's Neck Center for Agriculture & the Environment (Freeport, ME). OpenTEAM facilitated the creation of a set of sample documents in pursuit of their goal to enable agricultural producers to determine where their personal and agricultural data is stored, how it is processed, who has access, and how it is used.

The Agricultural data use documents include:

- Ag data glossary
  Establishes a shared language and understanding of agricultural data use concepts.
- Data fiduciary oath of care for agricultural professionals
  Serves as a trust-building document between agricultural producers and their advisors.
- Agriculturalists' bill of data rights
  Rights guaranteed to agricultural producers by participants in the OpenTEAM technology ecosystem.
- Data hosting and storage agreement
  Defines standard terms for hosting and securing agricultural producer data.

[17] "Boilerplate Data Hosting and Storage Agreement - Ag Data Use Agreements." Accessed July 15, 2025. https://openteam-agreements.community/hostingandstorage/.

In the course of their work, AgGateway Data Ethics Working Group reviewed the OpenTEAM documents, and provided feedback to OpenTEAM. OpenTEAM welcomes agricultural producers, their advisors, and agricultural technology creators in a pre-competitive space.

C.9 The Better Deal for Data[18]

Tech Matters, with funding from the Patrick J. McGovern Foundation, the Foundation for Food & Agriculture Research, Schmidt Futures, OpenTEAM, and others, is working to standardize a set of easily understood commitments which can guide adopters to policies which are respectful of data subjects, building trust and reducing friction in agricultural use cases and beyond.

The draft commitments of the Better Deal for Data (BD4D) are:

1. We are using Your Data to benefit You, Your community, humanity, and the planet; not for private gain or profit.
2. We don't claim ownership of Your Data: it remains subject to Your control.
3. We will delete Your Data, correct it, or transfer it somewhere else if You ask.
4. We will not monetize Your Data by providing it to third parties for compensation.
5. You can decide if You want to make Your Data open, or want to monetize it for Your benefit.
6. We will protect and steward Your Data and comply with applicable privacy laws, but You may have privacy obligations as well.
7. If You allow research with Your Data, we will follow best practices around the anonymization of personal data, and published research results will be made available to You for free.
8. We will be bound by legal agreements implementing these commitments, and anyone we share your data with will be similarly bound.

Tech Matters is actively soliciting community participation in the creation of the Better Deal for Data.

C.10 EU code of conduct on agricultural data sharing[19]

The voluntary "EU Code of conduct for agricultural data sharing by contractual agreement" (2018) was developed by a broad coalition of farmer and agricultural industry interests. Though very centered in EU laws and policies like GDPR, it provides both an excellent introduction to the nuances of agricultural data sharing, and specific recommendations for how rights and responsibilities should be allocated.

---

[18] Tiberio, Courtney. "Better Deal for Data White Paper." *Better Deal for Data*, April 2, 2024. https://bd4d.org/better-deal-for-data-white-paper/.

[19] "EU Code of Conduct on Agricultural Data Sharing by Contractual Agreement | FAO." Accessed July 15, 2025. https://www.fao.org/family-farming/detail/en/c/1370911/.

Key sections include:

- Data ownership: Attribution of underlying rights to determine the use of data
- Data access, control, and portability
- Data protection and transparency (opt-in / opt-out)
- Privacy and security: Interaction with the broader provisions of privacy law
- Liability and intellectual property rights

This EU Code of Conduct advocates for clear legal agreements between the various parties in the agricultural supply chain. The document includes clear definitions of important terms and case studies to illustrate its most important concepts. It concludes with a check-list for a balanced legal contract.

C.11 Other Resources and links

The above resources and summaries are not meant to be all-inclusive, they were driven by what the authors were aware, and familiar with to be able to provide a brief summary. Additional resources are listed below that may be relevant to specific areas of data management, or jurisdictions. This list is not meant to be exhaustive, but the team wanted to share as may resources as possible. Some are freely available others may require purchase or membership in an organization for full access.

Papers, guidelines, and other publications

- AgGateway data privacy and use whitepaper (2017) Data Privacy and Use White Paper v1.2
- Data governance in the dairy industry https://doi.org/10.3390/ani11102981
- New Zealand farm data standardshttps://www.datalinker.org/
- National Farmers Federation: Australian farm data code https://nff.org.au/programs/australian-farm-data-code/
- Exploring legal mechanisms for data stewardship https://www.adalovelaceinstitute.org/report/legal-mechanisms-data-stewardship/
- McKinsey: Data ethics: What it means and what it takes https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/data-ethics-what-it-means-and-what-it-takes#/
- Open Data Institute: Assessing risk when sharing data: a guide https://theodi.org/insights/guides/assessing-risk-when-sharing-data-a-guide/

Standards

- IEEE 2986-2023: Recommended practice for privacy and security for federated machine Learning https://doi.org/10.1109/IEEESTD.2024.10507779

- https://www.iso.org/standard/27001 Information security, cybersecurity and privacy protection — Information security management systems — Requirements
- https://www.iso.org/standard/71670.html Security techniques — Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management — Requirements and guidelines

Laws, legal documents, policy recommendations

- IAPP US State privacy legislation tracker https://iapp.org/resources/article/us-state-privacy-legislation-tracker/
- US FCC Task Force for Reviewing the Connectivity and Technology Needs of Precision Agriculture in the United States Report – Final Report *December 5, 2024*
- California Consumer Privacy Act (CCPA) https://www.oag.ca.gov/privacy/ccpa
- General Data Protection Regulation (GDPR) https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=LEGISSUM%3A310401_2

Appendix D: Legal Topics

The below list includes topics that should be addressed in the various legal documents an organization has for a system that stores, processes, or provides access to data. The list is not exhaustive, but a good starting point to ensure relevant topics are addressed in the various legal agreements between the provider and user of a system.

- Data
    - Data collected
        - Data type or categories included with an emphasis on any Personally Identifiable Information (PII) - see above list
        - Data (PII) collection and storage from minors
        - Use of cookies and other tracking technologies
    - Data use
        - Data use by system provider to operate system/provide service
        - Data use by 3rd parties to operate system/provide service
        - Data sharing with 3rd parties, liability for their actions
        - Data transfer for system operation
        - Data sale or distribution
    - Data ownership and control
        - Data license between users and system provider
        - Who has authority over the different types of data
    - Data management, correction and portability
        - PII access and correction
        - Data portability
        - Data deletion
        - Data retention period
        - Data security measures and data breach notification
    - PII disclosure circumstances, comply with laws/subpoenas
- License
    - License to use system
    - User Responsibilities: account security, compliance with laws
    - Intellectual property, reverse engineering, etc.
    - System or technological limitations/expectations
    - Warranty
        - As is/As available disclaimer
- Agreement
    - Questions/contact info (notification methods and contacts)
    - Modification and notification of agreement/policy
    - Assignment of agreement
    - Term and termination
        - Time period for agreement
        - Termination of services/access/subscriptions
        - Ongoing obligations, if any
    - Limitation of liability
        - Liability for third party sites linked from system/site (or disclaimer)

    - Dispute resolution, Jurisdiction
- Address sale of company: reorganization, sale, or merger
- Represent benefits or profits from sharing data

Appendix E: Terms and Definitions

California Consumer Privacy Act (CCPA) - A state statute intended to enhance privacy rights and consumer protection for residents of California, USA. It grants California consumers the right to know what personal data are being collected about them, the right to delete personal data, and the right to opt-out of the sale of their personal data

Data consumer - entities that obtain or interact with data, in this document used interchangeably with Data Exchange Partner.

Data ethics by design - the practice of embedding ethical considerations into data management processes from the outset, rather than treating them as an afterthought. It emphasizes proactive governance, transparency, and accountability in how data is collected, stored, analyzed, and used.

Data exchange partner - entities that obtain or interact with data, in this document used interchangeably with Data Consumer.

Data originator - any entity whose actions or property is being measured and the resulting data subsequently sent to another entity.

General Data Protection Regulation (GDPR) - A comprehensive data protection law enacted by the European Union that governs the collection, use, and protection of personal data. It applies to all organizations operating within the EU and those outside the EU that handle the data of EU residents. The GDPR emphasizes transparency, accountability, and individuals' rights to control their personal data

Informed consent - the process of ensuring individuals fully understand and voluntarily agree to a decision before proceeding

Integrated systems - These are agricultural systems that combine multiple farming enterprises, such as crop production, livestock, and agroforestry, to optimize resource use, enhance productivity, and improve sustainability. They generally collect data from multiple sources to aggregate and analyze to create recommendations and inform decisions on farm operations.

Third party service provider - An entity that provides products or services to a data consumer or data exchange party. They do not have a direct relationship with the Data originator but may process their data.