

The International Society of Precision Agriculture presents the
**16th International Conference on
Precision Agriculture**
21–24 July 2024 | Manhattan, Kansas USA



Authors:

Bierman, Don - My Ag Data, USA

Craker, Ben - Ag Data Coalition, USA

**Who are the data stewards: Moving data driven
agriculture forward**

A paper from the Proceedings of the
16th International Conference on Precision Agriculture
21-24 July 2024
Manhattan, Kansas, United States

Abstract.

Nearly a decade ago agricultural equipment manufacturers, service providers, retailers, land grant universities, and grower organizations came together to begin discussing the growing needs for producers to manage their farm data. This discussion was partly fueled by the industry shifting from moving data via physical media to cloud API connections. Several initiatives including the Agricultural Data Coalition (ADC) were subsequently launched focusing on addressing data privacy and security concerns and were soon followed by regulatory action in some parts of the world such as the General Data Protection Regulation (GDPR) in the EU. Since, there have been many discussions, concerns shared, technological advances in the industry, and other changes; however, many of the same issues remain related to data privacy and access. In this paper, several of the more complex use cases about data ownership, access, possession, and permissions will be reviewed. Having a detailed understanding of these situations is imperative to creating solutions to these complex issues that can also be applied to more straightforward situations. The results of recent surveys and interviews conducted by the ADC with different stakeholders will be discussed. The survey was partially focused on understanding what technologies are being used, not just installed on equipment. For example, if a machine has GNSS guidance or automatic section control farmers use it, however several indicated their machines are equipped with telematics but are not utilizing the system. Through the interviews the need for better data management and interoperability was laid bare when discussing data entry for sustainability programs. There is currently a tremendous effort required to collect, organize, and manually enter information presenting a huge barrier to adoption of these programs. This is a prime example of the need for data permissions and access, since such a broad variety of data is needed, and from the interviews it is not the farmer gathering and entering the information. These data points are informative not only to perceptions of different actors in the industry but also provide insight into the real world. Understanding the current state of who is collecting, archiving, and distributing data is critical in informing recommendations on how to address these issues in the future. Knowing who wants data, who is paying for data, and who is deriving value from data is key to understanding how permissions will be managed. The ever-growing pressures to document all aspects of agricultural production to provide consumers with information on where their food comes from will continue to highlight these data access issues. The need to produce and document climate

The authors are solely responsible for the content of this paper, which is not a refereed publication. Citation of this work should state that it is from the Proceedings of the 16th International Conference on Precision Agriculture. EXAMPLE: Last Name, A. B. & Coauthor, C. D. (2024). Title of paper. In Proceedings of the 16th International Conference on Precision Agriculture (unpaginated, online). Monticello, IL: International Society of Precision Agriculture.

smart commodities, similarly, will also hinge on different actors agreeing to grant permission to a multitude of data sources.

Keywords.

Data, privacy, ethics, control, ownership, stewardship, access, permission, sharing

Agricultural data stewardship over the years

Over the years, there have been many initiatives and efforts focused on data privacy and security for the agricultural industry, nearly since the advent of digital farming technologies. Through these projects, different best practices have been identified, scenarios analyzed, and recommendations made, generally with the involvement of many different industry players spanning the equipment manufacturers, data consumers, and farm organizations. However, there are still many questions and concerns within the industry about data use, privacy, sharing, and storage. In this paper we set out to explain some of the fundamental complexities driving these ongoing concerns. By better understanding the root causes we hope to move the industry forward by recommending some best practices and other courses of action.

Data stewardship, what does it mean

Often a good place to start when addressing a broad area of concern and confusion is to agree on terminology. Ensuring everyone is working from the same definitions helps make the communication of thoughts and ideas clearer. Data sharing, ownership and access often have their meanings conflated. For the purposes of this paper the definitions from OECD (OECD, 2019) for these concepts will be used:

- Data Sharing – refers to the provision of data by the data holder, on a voluntary basis. It includes the re-use of data based on commercial and non-commercial conditional data-sharing agreements, as well as open data.
- Data Ownership – implies stewardship over data, including rights related to access, creation, modification, sale, storage, and licensing. It involves determining who has control over the data and its associated privileges (Pierson, 2019)
- Data Access - pertains to the ability to retrieve and utilize data. It involves mechanisms that enable authorized users to obtain relevant information for analysis, modeling, and decision support.

When concerns regarding data ownership and control originally surfaced, there was a large push to assert the farmer owns the data. Generally, people agree with the opinion, it is the farmer's data, they own it and should be able to do as they wish with it. However, owning data is different than owning many other assets. Ownership of physical things is fairly straightforward since there is usually only one actual thing, and making copies of it generally is difficult, if not impossible. However, for data it is easy to generate a nearly limitless number of copies, making control of the data very important. Unlike physical things, sharing data can result in essentially a loss of control of the asset even if the owner retains the original since copies can be made and widely distributed.

During a panel discussion at a Hermann Lecture (Huermann Lectures 2023) the question was posed from the audience on whether the industry should stop using the term sharing. For many people, sharing conjures up the image of children with a toy. They can give the toy to their friend, their friend plays with it for a while and gives the toy back, and most of the time this behavior is encouraged. With data, it's quite a different situation. When you provide the data to someone, they're able to take it and make copies of it. They could potentially give it to numerous other entities without the person who originally provided access really knowing what's happened, which is quite different than sharing a toy.

Should we now be focusing more on using terms like access permissions instead of talking about sharing data? The core concept being the different terminology will result in people thinking differently about providing access to data. This is likely a reaction to many companies

being opposed to saying the farmer owns the data within their data privacy and sharing agreements. If companies are unwilling to say a farmer owns the data, surely, we should change how we think about sharing that data. However, it is important to consider why these companies do not want to use the term ownership. With regard to agricultural data, ownership can be very difficult to determine in many circumstances. This will be demonstrated with the complex use case later in the paper. The legal perspective was to focus more on who has access or control of the data, since that was a much better metric to be able to manage and insert into a legal document. Much of the industry did not like that nuance, many people prefer to be able to clearly state the farmer owns the data without qualifiers.

Another concept that has more recently entered into conversations is FAIR data (Wilkinson et al., 2019). Fair stands for Findable, Accessible, Interoperable, and Reusable. These are excellent principles that should be applied to all data. However, it is important to note that FAIR data does not mean free, or completely open data. It is possible to have data that is easy to find, easy to access, interoperable and reuseable but still have the ability to permission access to it. In the context of farmers providing access to data with researchers this is a very important point since the farmer may not want the data to be used beyond the specific project they are providing. Special care must be taken to make sure the data is provided with a clear understanding of who will have access and what happens with the data when/if the research is published.

Use cases and scenarios

To help explain the role of data access (sharing) relative to ownership and how it impacts the demands the industry is putting on growers, it's helpful to walk through different use cases of how this data is generated, who has access, and who needs access. This helps to explain where the gaps may be and what best practices or management decisions could be made to help improve the situation.

Data Sources, where does it come from

There are several key actors who either collect or generate data themselves or provide the equipment and technology that does. It is important to understand who they are and what their interests are.

- The farmer, who is generally at the center of discussions around data access and permissions. They are the entity or individual responsible for the management of the farm, whether row crops, specialty crops, or livestock. They are interested in having any data relating to their operation to inform management decisions or meet reporting requirements.
- The owner of the land on which the farming operation takes place, increasingly, is a third party who may not be very involved in the production taking place on the property. They do, however, have a strong interest in the data generated on their land. They are managing an asset and want to be sure it is being maintained so they can continue to derive value from it.
- Equipment manufacturers provide modern connected equipment capable of capturing data during an operation. This would include machines like tractors, soil probes, application drones, planters, fertilizer applicators, irrigation systems, feed mixers, weather stations, harvesters, biosensors, and other mobile and fixed equipment. This would also include component manufacturers that make sensors or other hardware that is added to other equipment as aftermarket solutions.
- Somewhat adjacent to equipment manufacturers are machines and technology platforms such as drones that may be used for remotely collecting observations, like imagery. Additionally, satellite or crewed aerial vehicle platforms for collecting imagery would fall into this category.

- Another somewhat broad category is advisor or service provider. These are the people and organizations focused on providing agronomic insights and recommendations. They might use some of the other categories of equipment to generate information such as soil samples, imagery, or other information sources such as crop scouting.
- Somewhat analogous to the field-based services providers are the nutritionists and other advisors making recommendations for animal operation about things like diet formulations.
- Another variation in the service provider category is the custom operator. Sometimes they may be the same organization as the one making a recommendation, but in this case, they also have their own fleet of equipment to apply products or perform other operations for the farmer.
- Also, just like some of the service providers will also use other types of equipment, there are also many software tools the farmers, advisers and other organizations will use to analyze and act upon collected data.

This list is not exhaustive, but hopefully provides a base understanding of the different actors, the equipment and technologies involved in generating, analyzing, and utilizing agricultural data.

Who owns and more importantly – controls – data

While it may sound like a minor difference, the distinction between ownership and control is at the heart of concerns from the legal perspective on using the term ownership. In the following examples, which are common scenarios in modern agriculture, there are a variety of people that could own data generated during a farm operation. Therefore, it is often more practical to discuss who controls the data. Often what people are concerned with is who has access to the data and how it may be used. Discussing the ownership of data does make sense in some cases, but when it comes down to it, laying out guidelines for access and control will likely have more tangible results. One way to address this issue is to make sure there is an agreement up front between the different parties that have potential ownership over any data that will be generated. If everyone shares the same understanding, hopefully in writing, it really helps to clarify the ownership versus control dynamics.

Simple example

Let's start with a very simple use case: A farm operation owns their own land and is conducting a field operation with their own equipment that's logging the data locally, storing it on the terminal in the machine (for instance, fertilizer application), the data is logged by the machine on a terminal in the cab and generally transferred via physical media such as a USB stick to a software program running locally on the farm computer. The farmer is in control of that data the entire way through. They own the land, equipment and software used and are collecting and transferring the data themselves, so there's really no question on whose data that would be. This makes it very straightforward and easy to understand both the ownership and control aspects of the data in this scenario. This can apply to a wide variety of crops, from broad-acre corn or soybeans, dairy production, specialty crops, to fruits and vegetables.

Complex example

In the modern agricultural industry with the advent of connected machines, shifts in dynamics around land ownership and renting, as well as services provided by custom operators the simple use case is not common. It is much more likely to be a variation on this complex use case. The issues to clearly determine who owns the data are driven by the multitude of parties that can make a legitimate claim to at least part of the data generated. As illustrated in *Figure 1*, the parties include the farmer managing a field that is owned by a separate entity. The farmer in this case has contracted with a custom operator to apply a crop nutrition product to the field based on a prescription map generated by an agronomic advisor. The custom operator has equipment with an integrated OEM telematics system, which means the machine logs data as it moves across the field for real time performance monitoring and logistics purposes. An aftermarket solution is installed on the machine to control the actual application of product based on the prescription map that also logs the as-applied data. This means there are two separate streams of data coming off the same machine stored in separate cloud platforms. As a result of the different parties and systems involved the ownership of the information collected during the operation is more difficult to determine.

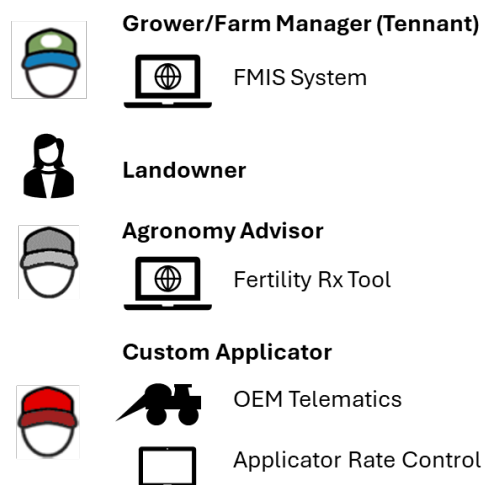


Figure 1 – Entities and systems involved in the complex example.

Each of these different actors involved complicate the situation by having a potential ownership claim on the data:

- During the operations the machine generates performance and logistics data that is vital to the custom operator. They can use this information to monitor their machines, see how well operators are performing and plan logistics of the operation. Additionally, they may use some of the information to bill the farmer for their services. With the custom operator owning the machines that are generating the data they have the most control over the data and a claim to its ownership as well. The service provider is also then responsible for transferring any logged data from their machine to whoever needs or wants it.
- Each OEM manufacturer would like to have access to the data the machines and hardware generate to understand how they are used, the performance, and other information to enable additional services such as predictive maintenance scheduling for the equipment. Additionally, they can use the data to inform product development decisions by better understanding how the machines are used. Many OEMs require the owner of the equipment to provide access to the data for their telematic services to even function properly. Not agreeing to the terms results in dramatically reduced functionality, while this does not constitute an ownership claim on the data it does have a similar end result.
- The agronomic advisor could also have an ownership stake in the data, their expertise was used to generate the variable rate map. How the rate is determined for the different areas of the field is really their “secret sauce”, so they have a strong ownership claim on the prescription which ties closely with the as-applied record.
- Obviously, the farmer would like to have access to that information since it would include the actual as-applied record of what product was applied where in the field, this is crucial for operations managing on a sub-field level. Understanding what areas yielded better and what areas yielded less with different amounts of inputs can inform management decisions for the future. For a farm operation, that data is key to being able to make informed, data-driven decisions. Since they orchestrated the recommendation,

application, and ultimately harvest of the crop, they have a stake in owning the data generated.

- Similarly, with the field being rented, the owner of the land has an interest in understanding how the field is being managed. Having the record of crop nutrition application information can help them understand if the nutrients available in the soil are being maintained or depleted. They have a vested interest in understanding what's happening in the field and if the renter is responsibly managing it. Since it is their field where this operation is taking place, they have a reasonable claim of ownership on the data.

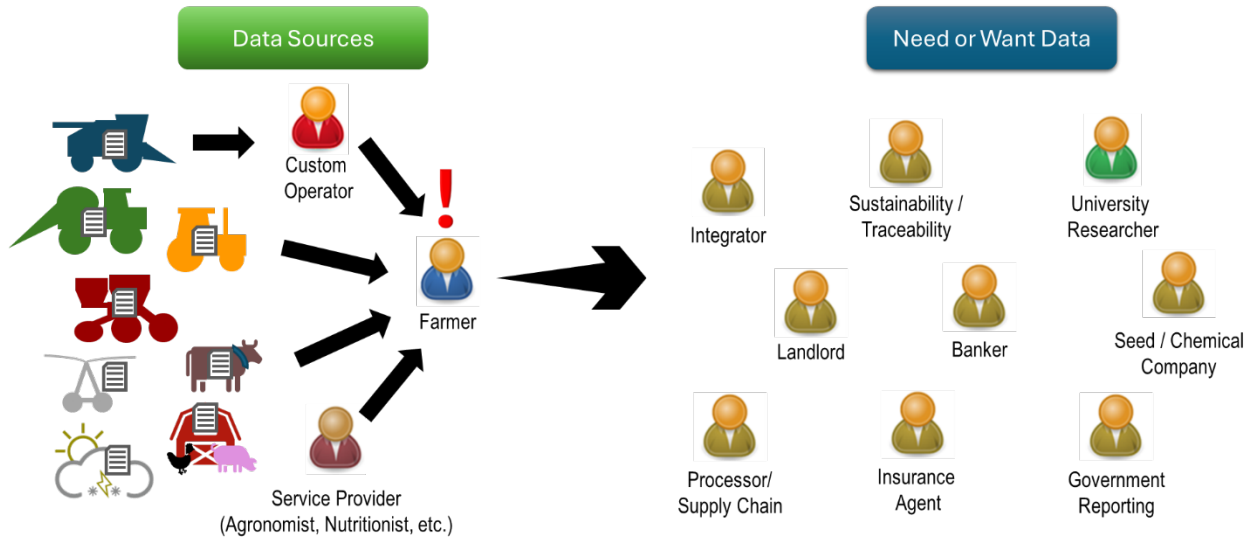


Figure 2 – Data sources and entities requesting data from the farmer.

Data uses

Understanding where the data comes from and who may have an ownership claim is helpful. But what are the uses of the data by the various actors and how does that influence decisions about access? Figure 2 above shows where the data may come from on the left, but the right shows a sample of parties that are interested in receiving data from the farm operation. Below is a list of different uses for the data that further builds out the motivation for each party to try and ensure they have access to the data and how they generally manage it.

Farm management

Growers and farm managers use software tools commonly referred to as Farm Management Information Systems (FMIS) to track, make, and implement decisions on the farm. Farmers can use them to track inventory, log spatial information and generate profitability maps for their fields. Most also have the ability to generate variable rate prescription maps and often serve as the primary archive of farm records, especially for equipment generated data.

There are a variety of other tools that perform a similar function for managing dairy herds and other livestock operation. Some FMIS systems even have a livestock component that is often an add-in feature set. These tools provide very similar functionality, just focused on livestock data instead of data from fields. One difference today that is beginning to shift, especially in the dairy sector, is the tools are more limited in scope of functionality. Where an FMIS is generally capable of performing accounting functions, payroll management, as well as reading and generating controller files for precision ag use, livestock tools are more targeted. Often there is one tool for general heard management and milk production tracking. A sperate tool is used to generate rations to send to the mixing equipment.

Regardless of the tools used, it is clear farmers need data to make decisions about how to run

their operations. There are countless different variables farmers try to account for when planning and executing operations. Having reliable and readily available information about what has and has not worked in the past is crucial to improving decision making going forward.

Precision ag services

Often utilizing data to inform and drive management decisions on the farm is done with the input of an advisor. They collect a variety of data such as soil test information, historical yield data, and other information to develop variable rate maps or seeding recommendations. It is generally beneficial to enter the as-applied and final yield information back into these systems to verify the success of the recommendation and further refine models for future seasons. In many instances ag service providers will use more than one of these tools as well. They generally have one tool they prefer for making seeding recommendations and the subsequent prescriptions maps but have a separate tool for fertility recommendations and maps. Sometimes there are also incentives to load data into different supplier tools to meet rebate requirements, though they may not use those tools to perform any analysis or generate prescription maps. Most precision ag advisors will have another tool they use primarily for converting various file formats so they can be loaded into the other tools that may not support all file types used in the industry.

Similar to the agronomic recommendations, livestock producers have relationships with service providers as well. They function in much the same way, collecting various information from different sources to make dietary recommendations and generate ration mixes. They also might analyze herd health and performance to make breeding or other recommendations about herd health. As mentioned above, these service providers will also likely have multiple tools they use for generating different recommendations and analysis.

Regulatory reporting

Generally, there are two broad categories of regulatory reporting. One, often simpler reporting channel is overall production information used for statistical purposes. This is what rolls up to county, state, regional, or country wide reports of area planted, yield, and other general production information. Most of the time this information consists of summaries, or totals for fields or areas defined by the reporting agency that do not necessarily align with how the farmer manages their operation.

The other broad category of reporting requires a little more specific information but is still often at a summary or totals per field or herd level of granularity. If there is any involvement of a subsidized insurance program where the government offsets the insurance premium, participation will require information. The government will want data on that field to be able to prove that it met the requirements for the program, this can be as simple as area and planting date. However, in more complex conservation programs there may be a need to report additional detail. In many circumstances the grower will provide their as-applied or yield data to their insurance provider who will take that precision agricultural data and translate it into the summary values and other reporting information needed. Depending on the regulatory body involved they may not accept controller files from a machine but only support manually supplied information.

Research

For decades farmers have relied on land grant universities and extension programs to help them learn and understand the value and effectiveness of different practices and technologies. With the advent of precision ag the ability to conduct larger scale trials and research has increased. Many institutions and extension programs work with collaborating growers to run trials in real world scenarios across entire fields. As a result, the data generated by farming operations now can be used as part of research and trials to understand how practices and products perform in a normal production agricultural environment, outside the lab or in smaller trail plots. Often farmers and other advisors in the industry prefer this type of information since it is more reflective of what happens in the “real world”. There is a perception that smaller trials in test plots can be influenced by various factors and results may not be the same at scale.

GHG reporting, value add commodities

When agricultural commodities are contracted and sold to a Consumer Product Goods (CPG) company that is either marketing their end-product as being sustainably grown or if they are offsetting their greenhouse gas (GHG) emissions by driving changes in their supply chain they have a strong interest in any data generated in farming operations. They need to have accurate documentation of not only what yield came from the field, but what management practices were used, what was applied to the field, at what rate and when. This is being driven largely by the purchaser of the end-product wanting to know more about what happened in the supply chain starting in the farmers field with regard to GHG emissions.

Land valuation

Many landowners' information about production practices is being treated like mineral rights were a generation ago. A farmer could rent the ground and farm it, but the mineral rights were a separate "asset" that was controlled by the landlord. Having records of the production practices and yields achieved is valuable in much the same way, this data can be used when land is sold to demonstrate production history and potentially drive a higher price. Additionally, the information about practices is especially relevant if participating in carbon programs, many of the programs have long term implications for what can be done on the land they own that needs to stay with the land much like a right of way to ensure the permanence of the carbon credits generated.

Monetizing data

For several years now there have been discussions and hopes within the agricultural community for farmers to be able to directly monetize the data generated in their operations. The assumption has been that if a third party can collect and aggregate data sets, they would be able to sell the data to different organizations. This has proven to be a difficult business model for a variety of reasons. Some issues are the ownership and control issues discussed in this paper, others relate to the time and effort involved with collecting quality data from farming operations. Still other problems come down to getting enough growers in a given geography or other categorization to participate in order to generate a large enough data set.

However, with the advent of carbon programs and other incentives to adopt "sustainable" farming practices there is now a new avenue for monetizing data. It is not in selling a large, aggregated data set, but in documenting production practices. Proving adoption of a particular farming practice is now something that can generate additional revenue for the farmer. The concept of climate smart commodities, and more specialized versions like sustainable aviation fuel are examples where a farmer is paid more for a crop that is indistinguishable from the conventionally produced commodity. The data about how the commodity was produced is the only aspect of a truckload of grain that earns a higher price than a traditional truckload. If the information is lost or unavailable the price of that load goes down.

Technology adoption, relationship to permissioned data access

Current state of industry perceptions, survey findings

In a recent survey conducted by the Agricultural Data Coalition through the National Agricultural Producers Data Cooperative, the group attempted to understand some of the perceptions around data access permissions, as well as what technologies were actually being employed on farms. Survey responses were limited, but in combination with several interviews with different actors within the industry, as well as some poll questions in a webinar, a good picture was generated. Often in the media and other sources, we hear that farmers do not want to share their data, and it's exceedingly difficult to get access to information about farm operations. What the survey found was a little bit different. In general, it seems that most farmers do not object to providing access to their data if it is to someone they know, someone they trust, or most importantly, they understand what they're getting for it. For example, when providing data to a service provider for a recommendation on fertility or a particular seed variety or seeding rate, farmers have relationships with those providers, understand what they're getting for it, and

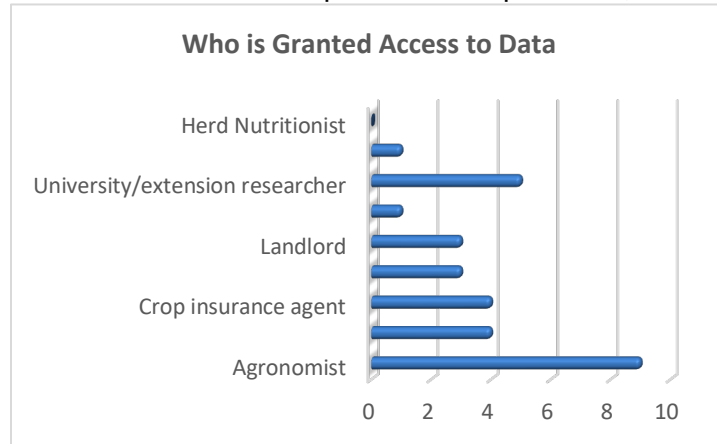


Figure 3 – Survey results of who is provided access to the farmers data.

seem to be willing to share their data with those organizations. Obviously, this is not a blanket statement that applies to every single farmer. There are going to be farmers that no matter what, do not want to share any information with anyone, but they are the outliers. For new startup companies that don't have the relationship or a clear value they are providing to the farmer result in more issues with data access. The risk-to-reward ratio is not what the farmer is comfortable with if they are not clear who they are giving their data to, what they are doing with it, and what they are getting in return.

Even with the evidence farmers are comfortable providing access to their data generally to their service providers they work with, it would be interesting to further dig into this with farmers. As mentioned in the complex use case, and explanation of the precision ag uses of data it is clear many of the service providers are often using multiple software tools, some provided by the input manufacturers themselves, some via third parties. Whether or not a farmer truly understands how many different systems their data gets put into for a service like a seeding recommendation was not clear within the context of this survey, so what the feelings farmers have about the actual process and tools involved may be different.

Technology use and relation to data control

The survey and interviews also captured information regarding what technologies are actually being used on farm shown in *Figure 4*. In recent years, new tools and products are being included from the factory with equipment. Systems like the connectivity hardware and subscriptions are being installed by default, similar to how air conditioning is now just included with a new car and not an option that has to be selected. The equipment industry has gone through several shifts with regard to technology pricing and machine configurations. When most new technologies enter the market, they start as an aftermarket add on, then move to integrated hardware sometimes requiring a one-time software unlock, finally to today's annual subscription-based pricing. This shift follows many other product categories where manufacturers need a way to generate a recurring revenue stream to cover the costs of ongoing software updates instead of selling "dumb iron" once and only needing to provide physical parts and no subsequent bug fixes or updates.

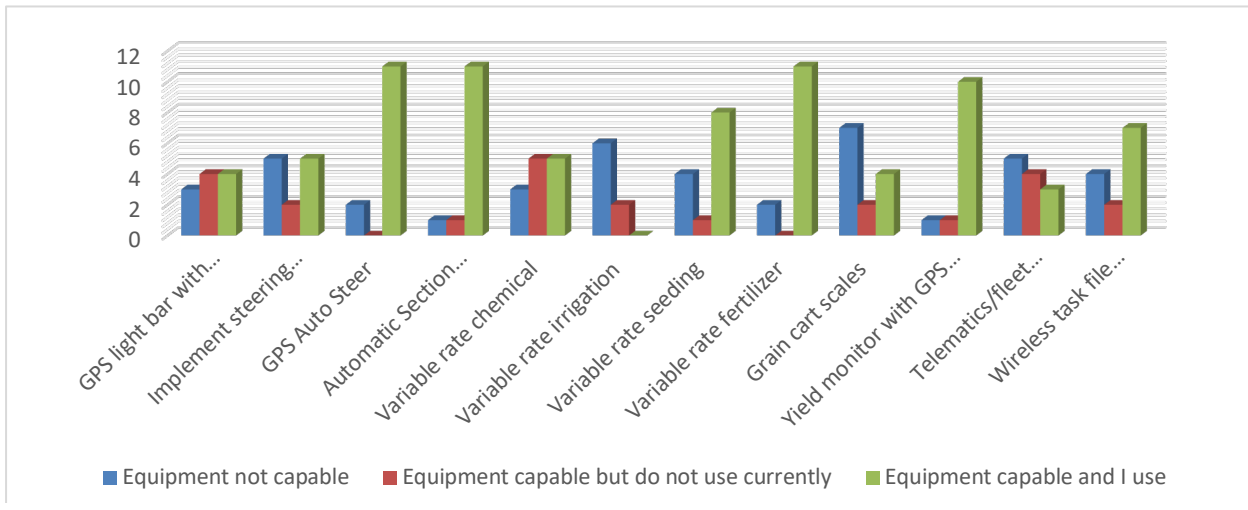


Figure 4 – Survey results for whether various technologies are installed on a machine and if they are used

This migration from aftermarket add-on to standard equipment makes understanding which of these tools are truly being utilized important since it directly impacts what data is being generated and accessed. Depending on if the data take is transferred through the wireless connectivity solution, or if it is moved via a USB stick has a huge impact on who has control and access to data.

The results of the survey and interviews showed that automatic steering or guidance has a very high use rate. Basically, if it's on the machine, farmers are using it because they see there's an immediate and direct benefit. Similarly, automatic section control is widely adopted. If the machine has it, farmers are using it. Again, this falls into that same category as guidance where it is an easy, immediate impact that is clear to see. The survey results traced very closely with the adoption of the technologies shown in *Figure 5* from the Purdue/CropLife Precision Dealer Survey (Erickson, B., & Lowenberg-DeBoer, J. (2023)).

Less adopted technologies and relationship to data access

When it comes to other technologies, the adoption rate was not quite as high. Variable rate technology had a lower adoption, and through the interviews the reason seems to be that it doesn't just work out-of-the-box like guidance and section control where there is little planning necessary. The operator can show up in the field, turn it on, and it will provide a benefit with not much additional effort. With variable rate technology, planning up front is required. The system requires a prescription map. That implies some analysis needed to be done on other data to determine rates as well as subdividing the field.

This is an extra logistics challenge to make sure the prescription map is generated properly and makes it to the cab of the machine. Then when the machine is in the correct field with the correct product loaded, the operator needs to load the map and engage the system. It requires more effort to use compared to the other technologies mentioned.

Another technology that is being included in the base machine by the equipment manufacturers is the connectivity hardware and subscription. These systems are installed on machines by default now but are not utilized as much by the growers in the survey. It's unclear as to why this is, it does seem farm operations would benefit from being able to remotely check where machines are, how they are performing, whether fuel is needed, or if routine maintenance is due. It may be the value of these systems is greater for the manufacturer and dealers since it enables them to provide services helping ensure machines are operating at peak efficiency. The systems also help with managing logistics, dealers can monitor machine hours to help ensure routine maintenance happens at the right time and can be scheduled in advance, all while making machines easier to find for remote service.

Regardless of why the telematics systems are not being utilized, that connectivity piece is important from the data control perspective. As mentioned earlier it helps not only provide the position and other information in the OEM cloud for fleet logistics, but generally, it is the same connection that transfers prescriptions and work records to and from the machine. Equipment manufacturers require data exchange with machines be done through their platforms, largely citing security concerns over remote access to machines. That means a grower leveraging the cloud connection completely changes who has access to and control over the data going to and from machines in the field.

Drivers for future adoption

One aspect uncovered during the direct interviews was regarding how data is being used and who has access to it, with respect to greenhouse gas reporting programs. Several of the service providers interviewed helped facilitate participation in programs for some of their farmer customers. The organizations were providing a service to select customers. In one instance, the service provider was helping with all the data collection, organization, and actually entry into the greenhouse gas reporting program. This consisted of entering large amounts of data into a spreadsheet to calculate the GHG reduction and payments for the carbon credits. They estimated it took roughly half an hour per field the farmer had enrolled to manually enter all the data into the tool. This did not include the time spent collecting and organizing all the information beforehand. The program required entry of a wide variety of information covering historical practices, production, fertility, crop rotation, fuel use, as well as other details. This manual process and time required will be a barrier to broader adoption of these programs.

A different scenario was described by another service provider that was interviewed. For the [Proceedings of the 16th International Conference on Precision Agriculture](#) 21-24 July, 2024, Manhattan, Kansas, United States

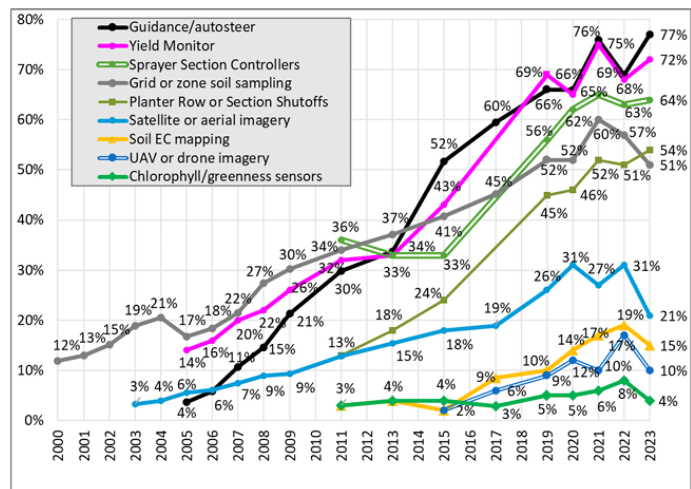


Figure 5 – Producer use of precision technology from 2003 to 2023
CropLife Purdue University Precision Dealer Survey.

program they were assisting their grower customers with, all of the data entry was being managed by the sustainability program personnel, so they did not have nearly as much time invested in manual data entry. However, there was still a significant effort going through all the different production history, crop rotation, and other history for the farm operation to provide the data the program staff needed to enter.

While there is a potential financial benefit to growers to enroll in these programs to open new revenue opportunities for their operations, it can be a time-consuming process. Based on the current estimated payout for participation in programs and the time and effort involved, many growers see the reward as not being worth the effort. For there to be wider adoption, the process needs to become more automated and less time consuming. Much of the data needed is generated via machines in the field but is still manually entered into the GHG program tools. If they were able to interface directly with OEM clouds or ingest files from machine terminals there could be a big benefit. This path would require farmers to do a little more up-front work to ensure terminals are populated with correct product information, machine name and configuration data, accurate date and time, and proper calibration. A little time invested up front could dramatically reduce the burden on organizing all the data at the end of the season.

Either way, most growers will have an initial heavy lift to collect and organize their historical data if they have not been following best practices already. But once this one-time task is preformed their data would be organized and ready for use in GHG programs, or for other purposes such as leveraging variable rate technologies and making more data driven decisions in their operation.

Overview of current data collection, storage

One aspect that has not been discussed so far is the actual collection and storage of data. This is another key topic that can influence who has access to information, again going back to the discussion on ownership versus access. The expression “possession is 9/10 of the law” not only applies to tangible assets but also data. Whoever actually collects data and stores it in their system obviously has a much stronger ownership claim than someone who does not have the data. That was one of the interesting things that came partly out of the study and to some extent within the interviews. Many growers are absentee data owners, meaning they are not actively engaged or prioritizing the collection, management or storage of data generated in their farm operations. They rely on third parties or systems to do it for them. For the most part, early adopters or the tech savvy segment of the market is more engaged in this type of activity, but for the majority of growers are relying on one of their service providers to often collect the data. Often, even if their machine is equipped with automatic logging going through the OEM's cloud, they are going into that cloud and giving permission to their service provider to access data in their OEM account.

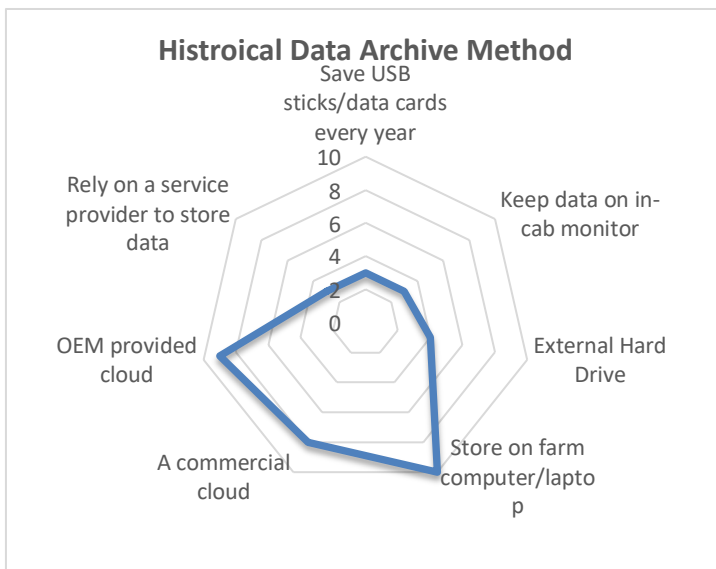


Figure 6 – Survey results for where farmers are storing data.

Growers who are concerned with permission and ownership need to take a more active role in understanding where their data is, prioritize its collection and storage, and understand how beholden they are to systems and providers managing their data. Changing this may require a shift in the industry to really move forward on issues regarding data privacy, permissions, and access.

If growers do not have their own repository of data generated on their farm, it may be a difficult discussion if they want to try out a product or service from a new startup or even a competitor to their current service provider. Growers without copies of their data need to go to their current service provider and ask for the data to be able to provide it for the new service. This situation likely drives the narrative we often hear related to issues with data access. It could be that it's not so much that the farmer is unwilling to provide access, but that the farmer is not as actively engaged in the collection and storage of data to be in a position to provide access to it quickly and easily.

Organizations like the Ag Data Coalition (ADC) have focused on solving this exact issue. Providing growers, researchers or even service providers with an independent data repository to store copies of original records and then permission access is fundamental to supporting adoption of future technologies and meeting data access needs. The concept of a central repository for all data, regardless of type or format, also can help growers as they participate in new GHG programs. Without a central archive of all original files, they would likely need to track down and organize data from multiple systems.

There are anecdotal stories that illustrate this point. For example, a grower may have Service Provider A perform soil nutrient analysis on a field. The provider goes out, collects the samples, gets the data, and likely uses that to make a crop nutrition recommendation. Then the grower may decide to try a different provider for the same service the following year. If the grower did not get a copy of the soil sample analysis from Service Provider A, they may be unwilling to give the soil test results to the grower or other provider. In that situation, the data ownership and control question are readily apparent. Who owns that soil analysis information? The grower owns the field; however, the fertility recommendation from the service provider may be all they received. If they did not agree upfront who would own, or at least receive copies of the test results they could be stuck, needing to collect new samples to work with a new provider.

This can be a barrier for the grower to try different companies or services if they first must have the field resampled because they're unable to get the data from the previous service provider. That type of interaction will not create a positive customer experience, so the expectation is over the long term not many organizations will restrict access to data in that way. Growers have multiple uses for information especially as the need to report on environmental practices increases, which could lead to growers needing the data for services their current provider does not offer. Organizations like the ADC have sought to make data storage and permissioned access a precompetitive space. The core belief is that organizations should compete on the merits of the product or services, not use their control of data to prevent customers from engaging with new providers.

Recommendations and paths forward

To help resolve some of the complexity within the industry and move forward with improved understanding regarding data access, there are several recommendations. One of the key points discussed is ensuring data access is agreed to prior to utilizing a product or service. Making it clear in the agreement the grower will receive a copy of the data from the machine in a custom operation, or that the landowner gets a copy of all field operation data from the tenant can help alleviate any misunderstandings. This is a shift for much of the ag industry, which has traditionally worked off a handshake or verbal agreements. Mutual understanding of the service to be provided and who will receive a copy of the data generated is important. Especially with the shift enabling the data to be used to generate revenue on its own or to transform a commodity into a value-added product. Any agreement, whether it be with a large tech platform, or a local service provider should clearly document who will receive a copy of the data generated and what the data is able to be used for. The topics covered under the principles of the Ag Data Transparent project are good topics to make sure are addressed before entering into any arrangement with any service provider or software system. Working with a local service provider, growers may have an easier time negotiating what they are able to receive or what limitations are placed on the different parties versus a large tech platform where users generally have to accept the company's terms or not use the system.

The second recommendation, as discussed earlier, is for farmers to prioritize data collection and

management. Due to the shifting landscape and the needs identified, where multiple players are going to want access to a variety of data, it is incumbent upon farmers to take more responsibility for that information to ensure they have what data they need. Not just to manage their farm operations in the short term, but potentially to provide documentation and proof on what they did historically throughout the course of their production activities. Relying on a third party to manage their data or maintain copies of original files generated during operations leaves a key part of the farm's information up to a third party. Most other businesses would prefer to have that information stored in a system they own or control, as opposed to relying on an equipment provider or other input provider to manage their data for them. As has been seen in recent years, these companies are not immune to changing their strategy, shifting their product focus, selling off divisions, shutting down divisions, or changing product features and capabilities. It is vitally important for the farmer to maintain a copy of the original records; they might not use the data today but in the future, they may need it for some of the different uses that have been discussed.

A third recommendation is for farmers and other operators in the industry to consider the data needs of academia when it comes granting access to the data generated in their operations. Land-grant institutions, extension agencies, and other independent voices within the agricultural industry need access to data to continue providing insights and information farmers have relied upon for years. For farmers to continue to have that independent, neutral voice, where they can ask questions about the efficacy of different products or services and get answers based on real-world information they will need data. These same institutions conduct and publish research to inform policymakers and other decision-makers. It's of vital importance to ensure these institutions have access to sufficient data to make informed recommendations to government and other regulatory bodies.

Summary/conclusion

In conclusion, data stewardship is very messy with many entities having an ownership claim to parts of data generated by normal farming operations. A one-size-fits-all approach will likely be difficult to employ. As more and more technology is deployed on farms, additional data and insights will be generated that will be valuable not only to the farmer, but also downstream. End consumers of the commodities that are processed into products on the grocery store shelves are concerned about where the food came from, how it was produced, and the environmental impact of the entire production supply chain. The manufacturers of the products recognize this need and have made pledges to reduce the GHG emissions in their supply chains and direct operations. Current market dynamics offer a premium to growers able to document their production practices but it is not difficult to envision a world where having that documentation is a requirement to be able to participate in the market. A healthy and vibrant agricultural ecosystem will require the data to be accessible by all parties within the industry with appropriate permissions. Employing some of the best practices seen through the GDPR regulations in Europe, such as what data is accessible, how long it's stored, and the right to be forgotten would make sense to implement within the agricultural industry. Clear understanding of those questions will help ensure farmers and other organizations are willing to provide access to their information.

There is immense value to be gained by providing access to data within the agricultural industry, whether that be to the farmers themselves, the equipment and input manufacturers, academia, service providers or input suppliers. To help meet the needs being placed upon the agricultural industry to help with climate change mitigation as well as meeting the demands of consumers to better understand the food value chain, data flow is going to be a crucial part of this industry. Having a solid framework for providing access to this information will be imperative going forward.

Acknowledgments

Thank you to the National Agricultural Producers Data Cooperative (NAPDC) who through a sub-award of the USDA-NIFA 2021-77039-35992 grant funded the survey, webinar, and interviews that were referenced in this paper. To learn more visit AgDataCoop.org

Thank you to AgGateway, discussions in a variety of their working groups and meetups helped to

inform and educate people involved in this paper regarding data use, stewardship and other topics.

References

Agricultural Data Coalition (2023, August 1). Reality of Ag Data Today and Tomorrow [Webinar]. CropLife Industry Innovation Series. <https://www.youtube.com/watch?v=7ww4MhIUyRg>

Erickson, B., & Lowenberg-DeBoer, J. (2023)2023 Precision Agriculture Dealership Survey.

Heuermann Lectures. (2023, May 10). The Next Generation of Farm Reporting [Conference session recording]. Nebraska Innovation Campus Conference Center, Auditorium. 2023 Transformation Drive, Lincoln, NE. <https://heuermannlectures.unl.edu/next-generation-farm-reporting>

Wilkinson, M; Dumontier, M; Aalbersberg, I; et al. (15 March 2016). "The FAIR Guiding Principles for scientific data management and stewardship". Scientific Data. 3 (1): 160018. doi:10.1038/SDATA.2016.18. ISSN 2052-4463. PMC 4792175. PMID 26978244. Wikidata Q27942822.

OECD. (2019). Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies. DOI: 10.1787/276aaca8-en

Pierson, C.A. (2019). Data Ownership and Data Sharing Practices. In: Poff, D., Michalos, A. (eds) Encyclopedia of Business and Professional Ethics. Springer, Cham. https://doi.org/10.1007/978-3-319-23514-1_315-1