Senator Jerry Moran
Questions for the Record
Subcommittee on Consumer Protection, Product Safety, Insurance & Data Security
"Technology in Agriculture: Data-Driven Farming"
November 14, 2017

Question 1. In your testimony, you state that as participation in the data community increases to a critical mass, farmers' bargaining power with the data service providers likely will be greatly reduced and a majority of the value will be enjoyed by the providers. You then state that for farmers to take maximum economic advantage of Big Data tools, large numbers of farmers must "buy-in" and participate in the data community. Where is most of this value enjoyed by the data service providers derived from? What can be done to mitigate the disparate levels of value received, especially for producers?

Witness response to Question 1, part A: Where is most of this value enjoyed by the data service providers derived from?

Before discussing the source of the value enjoyed by data service providers through data compilation and deployment of Big Data tools, it is important to discuss the value of data at the farm level in the form of Small Data. At the individual farm level, farmers may take advantage of today's data collection and analysis tools to run on-farm experiments with respect to seed varieties, fertilizer applications, moisture management, and so on. They can also use data to calculate crop shares for rent, look for improved efficiencies in equipment, management, conservation practicies, and so on. In these uses, farmers capture all of the data's value on-farm in the form of increased returns or reduced costs.

While these "micro" level benefits can be considerable for the individual producer, there can also be significant value derived through the use of Big Data systems and analytical techniques at the "macro" level when data from hundreds or even thousands of operations are aggregated. As mentioned in the written testimony, Big Data analytics can be used to much more rapidly develop hybrids by running trials across multiple soils and environmental conditions, developing more accurate and robust models for predicting risk factors such as weather patterns and production numbers, and development of improved agricultural equipment. All of these items can be derived from the aggregation of data from farmers and provide value back to those farmers.

When one examines the potential economic values of data for the data service providers, there are five primary sources to consider:

1) Data service providers can derive revenue from the services they provide directly to the farmer that provided the data. This could be in the form of service fees for things like data collection and validation, creating a repository of the farmer's data that can be easily shared with other parties to whom the farmer would like to provide access (such as a crop consultant, landlord, etc.), or in the case of a data service provider that also serves another role such as a crop consultant, providing reports, prescriptions, or recommendations to the farmer based upon the data.
2) The data service provider could derive revenue from using the data to market goods or services to the farmer. For example, if the data service provider is a subsidiary of a seed company, the farmer's data could help the seed company make seed recommendations that are a good fit for the farmer's operation. This form of focused marketing can sometimes benefit the farmer as well. For example, consider

purchasing a mixing bowl for your kitchen on Amazon. Based on that purchase, Amazon suggests a whisk and baking sheet that are commonly purchased with your mixing bowl that have gotten favorable reviews from people who purchased all three items simultaneously. If you also make the suggested purchases and enjoy them, it is a "win-win" transaction for both you and Amazon. Similarly, purchasing a seed variety that performs well for your farm and increases profits creates a win-win for you and the seed company. However, companies can also use a farmer's data to provide recommendations while extracting more profit from the producer. As another example, seed companies already use information about producers to know which seed varieties are better suited to the farmer's land and charge them more for that variety than they would charge another producer. With increasing access to producer data, input suppliers could continue to derive more precise information about a producers willingness to pay (or ability to pay) for their inputs and adjust their pricing accordingly.

3) Data service providers might provide data products or services based on farmers' data to other companies. For example, data service providers might sell reports or predictive models to insurance underwriters might help them price crop insurance products. These reports or models might be derived from farmers' data but their sale would not necessarily involve the transfer of the farmers' data.
4) Data service providers can function as data aggregators and then sell the farmer's data to third parties, deriving revenue from those sales. In some cases, such data service providers may pay the farmer for their data, but in other cases they may charge the farmer a fee for their data collection services (while providing some analytical services or reports back to the farmer) and also derive a fee from the sale of the data to other entities.
5) Eventually, if and when enough farms join data networks, a fifth use could come of that data – use of that information for significant transactions in commodities markets. A hypothetical example would be a data service provider who had access to a sufficiently large sample of farms to make accurate predictions of eventual U.S. crop yield who then takes positions in the commodities markets well before anyone else would be able to access that information.

A potential sixth value source is from the data service provider positioning itself as an acquisition target with the purchasing company getting either the data it holds and/or the subscription relationships with the farmers it serves. Economic theory and historical precedent both suggest that we will see an evolution in the agricultural data industry starting with a large number of service providers vying to engage farmers because, as Metcalfe's law suggests, the value of their data networks will increase as a square of the number of their network participants. Better-capitalized firms or firms with another competitive advantage will acquire other firms until eventually only a handful of dominant service providers – or even a singular monopolistic provider – emerges. In the course of this evolution, the more farmers and data a company can acquire, the more attractive they become as an acquisition target. While some firms are certainly pursing the strategy of becoming one of the dominant providers, it is equally certain that other firms are seeking simply to be acquired.

Witness response to Question 1, part B: What can be done to mitigate the disparate levels of value received, especially for producers?

Research continues to determine both the value of agricultural data in the aggregate and what proportion of that value is captured by the farmer relative to others in the value chain. As with the current USDA estimate that farmers capture 15.6¢ of the food dollar, it is likely farmers will not capture a large proportion of data values since they are relatively small, "atomistic" players in the market with little bargaining power and face significant barriers to the kind of collective action necessary to increase that bargaining power.

Having said that, farmers' bargaining power may be at a maximum right now. As mentioned in the response to Part A of Question 1, most data service providers recognize they are in a race to acquire access to the data of as many farmers and their acres as they can, as quickly as possible. Some are approaching this with the strategy of a telemarketer telling a prospect "sign up today, because this offer will be gone tomorrow!" However, farmers should be thinking like economists, and carefully weighing the benefits presented by any particular service provider with the value they can receive from that provider's services (or, indeed, the payment the provider is offering to secure the farmer's data). To that end, farmers should ask five questions of any prospective data service provider:

1) How many growers/farms/fields/acres are in the data service provider's data community? The higher the number, the greater the value it can potentially provide to the individual farmer.
2) What analytics conducted on the community will benefit my farm? This aims at the direct ability of the data service provider to increase the farmer's ability to make profitable decisions, regardless of any external benefits.
3) What data quality control standards are being used? If the data service provider is not taking strong measures to ensure the quality of the data in their community, it cannot provide reliable insights to the producer. To quote the age-old computer principle: "garbage in, garbage out."
4) What uses will be made of my data? This question has a number of implications discussed in the responses to other questions below, but here its purpose is to help the producer gauge what potential values the data service provider may be trying to capture that will not directly benefit the producer (or may actually increase costs for the producer).
5) What assurances can the data service provider make that the farmer's data will not be provided to third parties without the ability of the producer to share in the revenues from that transaction? Consider the analogy of a farmer giving an unrestricted easement to an oil pipeline company. The farmer may get a payment for the easement, but the oil pipeline company may be able to sell co-located easements to a natural gas company, a telephone company, a fiber optic company, and so on, without the farmer having any ability to capture the value realized from those transactions. Similarly, if a farmer does not have an agreement restricting "downstream" uses of his or her data, they only have one opportunity to capture the value of that data.

While farmer's negotiating power may be at a peak, it may also be a matter of timing to make sure farmers do not completely miss out on the ability to capture the value of their data. As the industry continues to evolve, we will likely see a progression going from a) farmers paying data service providers for their services, b) data service providers realizing they need more and more farmers in their network and thus reducing their costs, potentially to nothing, to join, c) companies

actually paying farmers for their data, but then d) companies securing a critical mass of farms or acres to have sufficiently robust data networks that they no longer need additional famers to join.

In the end, the most effective means of helping farmers secure the maximum amount of their data's value may be educational efforts to help them determine the value of their data, evaluate data service providers and their agreements, and make informed choices about their data sharing relationships.

Question 2. You note the irony in the growth of proprietary data network protocols that lead to complaints about the lack of interoperability of farm equipment systems also providing greater protection against data breaches. What measures can be taken to continue improving the interoperability of data collection without sacrificing security?

Witness response to Question 2:

There are two primary reasons computer viruses are far more problematic for computers using a Windows operating system than an Apple OS system: first, far more machines use Windows, and second, the nature of Windows architecture permits more access points to its code than Apple OS. At the same time, one could also make the argument that all computer users have benefitted from consolidation in the computer industry in that so many programs are now interoperable and data can be shared among what is now billions of other users with relatively little friction. By analogy, farmers and data service providers may both benefit from consolidation in the industry and the increased interoperability it provides.

Such consolidation may also "consolidate" security concerns as more eggs are kept in fewer baskets. This is the Windows problem again – since there are far more Windows-based computers in the world, virus writers devote more resources to viruses that target them. Thus, as consolidation continues in the agricultural data sector, increased research and development efforts will be needed to make sure that these "fewer baskets" are guarded by increasingly robust security tools. Congressional support of these research efforts would benefit not only agricultural producers but a vast number of Americans who rely on data security to keep their personal information secure.

The increasing automation of agricultural data collection and transmission tasks might actually serve as a means of increasing data security. The second reason Windows computers suffer from so many security issues is the Windows operating system was not built with security as one of its primary concerns, since its foundations were built before networked computing was a primary use of PCs. As a result, there are far more access points to its code that virus writers can use to insert malicious instructions. Conversely, there are fewer access points in Apple OS – a system built from the ground up for networked, multi-user applications that thus requires explicit user permission for code to be activated. The analogy in agricultural data is that the tighter integration of data networks across agricultural equipment creates fewer intervention opportunities for third parties. The market will likely continue to drive this integration. With that said, the issue of user's data security needs to be continuously brought to the attention of hardware and software developers so they can keep security as a foundational principle in their designs.

Another tool that might aid the security of farm data networks may not seem to be directly connected to the security issue, but certainly is: the expanded development of broadband cellular networks in rural areas. Cellular transmissions are encrypted by the nature of the processes that make cellular communication work; this has the most beneficial side effect of adding to the security of the transmission. However, in many rural areas, sufficient cellular signal or bandwidth is not available to make use of many current agricultural data technologies, requiring farmers to manually

download and transmit their data using much less secure methods. Expansion of rural broadband cellular access would have the dual benefit of making agricultural data tools more accessible and profitable for farmers while also reducing the security risks associated with collecting and transmitting agricultural data.

Question 3. In your testimony, you note that the current legal framework for ownership of agricultural data is inadequate for the transfer and aggregation of agricultural data. Is the agricultural data space unique enough to require specific legislation regarding ownership and property rights, or is a novel combination of existing property ownership laws more appropriate and adequate?

Witness response to Question 3:

One of the first questions farmers always ask about their data is "do I own it?" It is a natural question to ask, since farmers depend on access to land whether it is owned or leased, and thus are closely attuned to property rights. However, traditional notions of ownership break down to some extent with agricultural data. Thus, the better question for the farmer to ask may be "what rights do I have with respect to my data, and what rights to others have with respect to it. The most critical element of this may be what rights (and abilities) does the farmer have to exclude others from access to the data.

If one thinks about it, the notion of "privacy" is really a function of one's ability to exclude others from access to information. For example, HIPAA's provisions regarding health information naturally couples with the notion of privacy in one's health matters, as they pertain to matters of one's own body. Thus, HIPAA provides a legal right to exclude others from access to health information without explicit consent for the disclosure of that data. The Fair Credit Reporting Act (FCRA) covers matters of financial information, which also couples with traditional notions of privacy and the right to exclude others from access to one's financial records.

However, the barriers start to blur a bit with the FCRA in that financial transactions mean reaching out to and communicating with another party; we recognize that others may have some limited right to access our financial information if that information is relevant to their financial risk in some respect, such as loaning us money. Therefore, we allow credit bureaus to collect financial information and to disclose that information to others so they may make credit decisions about us.

The credit reporting analogy is important to understanding whether we need a specific legal framework for agricultural data. Whenever someone requests credit, they are required to ask if the potential borrower will give consent to the lender to access a credit report. Thus, for a third party to make use of the financial data, they must have the consent of the person about which the data is collected. Conversely, though, in many farm data agreements, the farmer may not have the right to approve or deny access to their data to "downstream" users.

One could argue that agricultural data is significantly different from HIPAA-protected health information (about the workings and condition of one's own body) or financial data, but one should also consider the fact that farmers lack the ability to protect their information from disclosure. Put another way, it would be impossible for farmers to try and keep confidential much of their production information and practices because they can literally be seen from aerial or satellite imagery. With relatively little effort one can tell how many acres a particular farmer owns and what proportion of it is in what crop for a given year (compare this to a bicycle maker or a coffee shop – it would be extraordinarily difficult to determine the volume or nature of their business, or to even tell one business from the other without continuous monitoring of the business at high image resolutions, simply because their businesses have roofs and farms cannot). It could be argued then that perhaps farmers should be given more protections than other businesses because one could derive a significant amount of financial information about them from publically available resources, to say

nothing of the improved ability to do so if one coupled data sources from a data service provider that transferred that information without the farmer's consent.

However, if one did desire to provide enhanced protections for agricultural data and allow farmers to exclude others from their data without explicit consent, two significant barriers loom. First, one would have to define the type of agricultural data subject to the protection. This would be challenging, given the broad diversity of data that can be collected and transmitted on farms today. Second, it would be difficult to define when such consent would be needed. Would the farmer have to give consent to any data transfer to another party? There would be significant transactional costs in such an approach. Further, there are doubtlessly data uses that will be available in the near future that might not even be conceivable today; it would be quite challenging to give informed consent in an up-front data use agreement when one doesn't know what data uses might be possible in the future.

The current legal framework might be serviceable as an interim tool to help provide farmers some grounds for the excludability of agricultural data, and enhancements to that framework may be possible. In the near term, though, perhaps one way to help farmers maintain control of their data is additional research into encryption algorithms that give farmers a key that would be required to access the information – this would put more control over downstream uses back in the hands of the farmers, and also give them an increased ability to participate in the value received for data transactions.

Question 4. As more and more firms enter the agriculture-technology space and interact with data used by and/or generated by farmers, the need for clarity and consistency on privacy principles is growing. For these new entrants, can you suggest any best practices these firms should engage upon to ensure their data privacy procedures properly convey the data's expected use?

Witness response to Question 4:

Much use is made in the agricultural data industry of the word "transparency" but there can often be much ambiguity in what that term means. The greatest value of that term, in the witness' opinion, is to err on the side of disclosure to the farmer when discussing the internal and external uses that the data service provider will be making of the farmer's data. Those uses should be disclosed clearly in language that is understandable by farmers with a wide range of experiences and educational backgrounds. One such example may be a Truth in Lending Act (TILA) disclosure. Though an agricultural data use agreement might not bear a clear analogy to a lending transaction, TILA makes clear the potential impacts of the lending transaction and the borrower's rights and remedies. Data service providers could benefit from making sure their data use agreements have similar levels of clarity.

Another principle beyond the clarity of the disclosure is its frequency. Using another financial analogy, individuals can use credit monitoring services to receive notifications when someone makes a credit inquiry about them. Data service providers could also provide notices when an external entity has made a request to access the farmer's data, or when a new internal use is made of the data. Robust notification procedures can also help farmers take protective actions in the event of a data breach.

As mentioned in the response to Question 3, there arises the issue of informed consent when a new data use arises that was not contemplated by the original data use agreement. Though it might increase transactional costs, the simple answer to this problem is to require disclosure of a potential new use and secure the farmer's consent to the use before it is implemented. The counterpoint to this approach, however, is that its increased transactional costs might make companies implementing it less competitive than those who do not.

Finally, as new companies enter the agricultural data sector, they would do well to avail themselves of the efforts of farm groups, existing data service providers, and equipment manufacturers to develop consensus on the principles that should govern agricultural data management. The Privacy and Security Principles Farm Data developed by the American Farm Bureau Federation and the Ag Data Transparency Evaluator are two good starts for companies to use in developing their operating policies and procedures. Both of these tools continue to develop, and the dialogue can provide greater benefits to the agricultural industry with increased participation from more farmers and data firms.

Question 5. As more and more firms enter the agriculture-technology space and interact with data used by and/or generated by farmers, the need for clarity and consistency on privacy principles is growing. For these new entrants, can you suggest any best practices these firms should engage upon to ensure their data privacy procedures properly convey the data's expected use?

Witness response to Question 5: [It appears this question may be an inadvertent copy of Question 4]

Question 6. While much of the data we discussed in the hearing is generated on farm and captured by farmers or their equipment, significant quantities of data is publicly available and critically important to inform risk modeling, yield prediction, etc. in both the public and private sector. How can we encourage the continued use of this type of data, and even grow our sources, while ensuring that farmers understand their role in this process?

Witness response to Question 6:

Perhaps the best steps that can be made toward this goal are to continue funding of research and extension efforts through our Land Grant universities to help producers understand the value of the data resources to their decision making processes. For example, the witness is currently the principal investigator on a Southern Risk Management Education Center grant funded through USDA-NIFA to develop a handbook and decision tools that can help producers understand the value of agricultural data tools and help them make informed choices about their uses (SRMEC Agreement 21667-19).

Additional research on how agricultural data systems can be made more robust, reliable, and accurate can also add to the volume and quality of publically-available data. For example, the most commonly-logged seed variety on planter data systems is the variety that comes first alphabetically on the system's drop down list. This means that producers sometimes inadvertently (although potentially carelessly or intentionally) select data inputs that are inaccurate, which in turn affects all downstream uses of their data. Research of tools to help improve data accuracy will not only increase profitability for producers, but will also improve the data and decision tools available to the industry.