Received: 20 July 2021 | Accepted: 13 November 2021 | Published online: 31 January 2022

DOI: 10.1002/agj2.20968

Agronomy Journal

SPECIAL SECTION: BIG DATA PROMISES AND OBSTACLES: AGRICULTURAL DATA OWNERSHIP AND PRIVACY

Addressing new data privacy realities affecting agricultural research and development: A tiered-risk, standards-based approach

James C. Wilgenbusch[1,2] | **Philip G. Pardey**[1,3] | **Naomi Hospodarsky**[2] | **Benjamin J. Lynch**[4]

[1] GEMS Informatics Center, Univ. of Minnesota–Twin Cities, St. Paul, MN 55108, USA

[2] Research Computing, Univ. of Minnesota–Twin Cities, Minneapolis, MN 55455, USA

[3] Dep. of Applied Economics, Univ. of Minnesota–Twin Cities, St. Paul, MN 55108, USA

[4] Minnesota Supercomputing Inst., Univ. of Minnesota–Twin Cities, Minneapolis, MN 55455, USA

Correspondence
James C. Wilgenbusch, Minnesota Supercomputing Institute, 117 Pleasant St SE, Minneapolis, MN 55455.
Email: jwilgenb@umn.edu

Assigned to Associate Editor David Clay.

Abstract

Concerns related to data ownership and privacy cut across all sectors of our economy, shape public–private research relationships, and, if left unaddressed, threaten to limit the potential gains to be had from the "big data" revolution. Rather than offer a one-size-fits-all approach to dealing with data privacy and security concerning food and agricultural research and development (R&D), we propose a three-tiered data security approach based on three tiers of risk tolerance: high, medium, and low with general guidelines explicitly mapped to standards. Data privacy and security are not costless, and so an economically informed approach that weighs the cost of a potential security breach against the benefits from accessing and using data for R&D is a more practical approach than treating all data equally from a risk management perspective. These tiers of risk must be understood in relation to standards for there to be meaningful governance of these data. We begin by characterizing the rapidly evolving nature of data privacy in an agricultural R&D context before providing an overview of the key means by which the privacy of agricultural data is presently being governed in various regions of the world. As an illustration of the approach that we propose, we apply our tiered risk and standards-based approach to the CGIAR's Responsible Data Guidelines. This approach is similar to that used by the healthcare sector to effectively implement data privacy requirements and promote an awareness among key stakeholders of the need for and importance of well-defined data privacy standards.

Abbreviations: DHHS, U.S. Department of Health and Human Services; FTC, Federal Trade Commission; GDPR, General Data Protection Regulation; HIPAA, Health Insurance Portability and Accountability Act; IEC, International Electrotechnical Commission; IP, intellectual property; ISO, International Organization for Standardization; NIST, National Institute of Standards and Technology; R&D, research and development.

1 | INTRODUCTION

Over recent years, the agricultural press and the farm organizations that represent farmer interests have paid increasing attention to the privacy, use, and ownership of farm-related data (see, e.g., American Farm Bureau, 2016; Herbold-Swalwell, 2018; McIntosh, 2018). Surveys conducted in 2014 and 2016 by the American Farm Bureau indicated that U.S. farmers "…were 'concerned' or 'extremely concerned' about which entities can access their data and whether that data could be used for regulatory purposes" (Janzen, 2019). While farmers appreciate the potential for agricultural information to improve their farming operations, a recent survey of Canadian farmers also revealed significant and increasing concerns by farmers in that country about the implications of sharing farm-originated data (Farm Credit Canada, 2019).

The data privacy and security concerns for data sourced from farms spill over and have significant consequences for agricultural research and development (R&D), whether that research is conducted by public or private entities. The very technologies that produce more farm-related data (e.g., satellite, drone, machine, and ground sensors) are also increasingly being used in experimental settings both on farms and on research stations. Likewise, rapidly expanding applications in the data sciences (e.g., artificial intelligence and machine learning techniques and their specialties such as neural networks or natural language processing) are lowering the cost of making more scientific and commercial sense of the deluge of agricultural data (Goldfarb & Tucker, 2019). Moreover, scientific data from public (university and government) agencies are increasingly being pooled with, or used to ground-truth, on- and off-farm crop- and animal-related data for data science purposes or to enable the development and deployment of new agricultural devices and applications driven by data.

As the data revolution in the food and agricultural sciences gathers pace, the concerns over data privacy and security, and their implications for innovation in the food and agricultural sectors, are bound to multiply. A likely driver of these increasing data policy, intellectual property and practice concerns is the notion that data has potential economic value and thus how best to create and share that value (Jones and Tonetti, 2020). These same economic drivers arose in the 1970s and 1980s as technological developments in the biosciences (e.g., gene sequencing, gene modification, and gene editing) unlocked new potential value in genetic resources that hitherto had been "freely and openly shared." This spurred a growth in the rules, regulations and IP related to the genetic resources used in agriculture (Binenbaum et al., 2003; Nottenburg et al., 2002; Wright and Pardey, 2006), all of which had and continue to have profound research freedom-to-operate and international trade implications for genetic innovations in agriculture. In addition, these concerns reach well beyond data concerning just the phenotypic (e.g., yield or quality) performance of crops and animals in farm or experimental field settings. The data revolution also encompasses the generation, analysis, and deployment of crop, animal, and microbial genomic information; all sorts of weather and environmental data; as well as food and agricultural management and socio-economic data (National Academies of Sciences, Engineering, and Medicine, 2019). Moreover, the source of data relevant for innovation in the food and agricultural sectors stretches well beyond the farm, involving data elements along the entire value chain linking farms to markets.

Not only are the sources and potential applications of data in agriculture proliferating, the entities performing the research are changing profoundly as well. As Alston and Pardey (2021) reveal, the private sector now performs over half (51.4% in 2015) of the world's food and agricultural R&D, well up from the one-third private share in 1980. Moreover, the private presence in food and agricultural R&D is moving well beyond the rich countries to involve research undertaken elsewhere in the world, particularly in agriculturally large, middle-income countries such as China, India, and Brazil. This is expanding the awareness and necessity to address the intellectual property (IP) and other privacy and contractual concerns related to public–private research relationships, many of which involve the sharing of sensitive farm- or firm-originated (e.g., agri-business) data. These developments are coming at a time when many public funding agencies are requiring more formal, and often more open-access, data management practices for the results of research that arise from the projects they fund (e.g., National Science Foundation, 2002; USDA–National Institute for Food and Agriculture, 2019; U.S. Agency for International Development, 2021). These IP pressures, in conjunction with the new scientific opportunities arising from innovations in the data sciences themselves, have given rise to new principles and guidelines affecting the stewardship and management of

Core Ideas

- Dealing with data privacy and security concerns is central to the future of much agricultural R&D.
- Concerns over how the privacy of agriculture data are protected limits their availability.
- A one-size-fits-all approach to data privacy will not effectively address stakeholder concerns.
- The healthcare sector offers examples of how to balance privacy with accessibility
- Technical standards framed by risk management will help develop trust among stakeholders.

scientific data. This includes the findable, accessible, interoperable, and reusable (FAIR) standards described by Wilkinson et al. (2016), or the FAIR(ER) data practices implemented by the GEMS informatics platform (GEMS Informatics Initiative, 2021) that in addition promotes the ethical use of data (that respects IP and privacy aspects of data) and also strives for replicable results from the reuse of data.

Concerns over data privacy are certainly not new. They are widespread and affect nearly every sector of our economy, and the approaches to addressing them are in some cases more mature and could serve as models for the agriculture sector. For example, within the healthcare sector, the Health Insurance Portability and Accountability Act (HIPAA) (Health Insurance Portability and Accountability Act, 1996) has governed the data privacy and security provisions for safeguarding medical information in the United States for over 20 years. Importantly, HIPAA is not prescriptive and therefore does not provide anything like a check list or a mapping of guidelines to standards, which can be used to develop specific implementations. Several years after HIPAA was signed into law, the U.S. Department of Health and Human Services (DHHS) released the HIPAA Privacy Rule (Standards for Privacy of Individually Identifiable Health Information, 2002) and HIPAA Security Rule (Centers for Medicare & Medicaid Services) (Health Insurance Reform: Security Standards, 2003) to establish technical and nontechnical standards and to operationalize the protection of an individual's electronic protected health information. Recognizing the sensitivity of electronic protected health information and the increased risk of cyber-attacks, the DHHS Office for Civil Rights created a "crosswalk" between HIPAA security rule and National Institute of Standards and Technology (NIST) cybersecurity framework (DHHS, 2014) in an attempt to address cybersecurity gaps and to assist healthcare organizations in increasing their attention to securing health data. We propose a similar approach for agricultural data, in which the general guidelines found in the various approaches to governing agricultural data privacy are mapped to well-defined standards in the context of varying levels of risk.

Our paper draws on the significant strides that have been made in the health sciences. We begin by briefly characterizing the rapidly evolving and complex nature of data privacy in an agricultural R&D context before providing an overview of the key means by which the privacy of agricultural data is presently being governed in various regions of the world. As a practical illustration of the approach that we propose to deal with data privacy, we apply our tiered-risk and standards-based procedure to the CGIAR's responsible data guidelines (CGIAR, 2020). To our knowledge, this is the first mapping between agriculture data privacy governance guidelines and the technical and nontechnical standards that can assure the degree of privacy being sought.

2 | THE CHANGING PRIVACY REALITIES CONCERNING AGRICULTURAL R&D-RELATED DATA

At its core, R&D, not least research directed to agriculture and food, involves generating data for the purposes of advancing science or promoting technical change in the agri-food sector (Alston & Pardey, 2021). However, for much agricultural research, including increasing areas of R&D conducted by the public sector, the days of open and unfettered access to and unencumbered rights to share and use significant amounts of the data generated by science are long gone. For many years, public-sector scientists conducting contract breeding or experimental (yield or crop management) trials for commercial companies have been subject to restrictions regarding precisely what data can be shared with whom and for what purposes. That private presence in the public agri-food sciences continues to grow, now reaching into areas well beyond their historical focus on genetic, fertilizer, and related experimental testing and trialing. Now, the increasing use of sophisticated robotic, drone, and other data capture devices for (field) research performed by public scientists are associated with an increasing and often complex contractual assignment of use rights regarding the data arising from such R&D. In some instances, use of leased instrumentation is bundled with private (artificial intelligence and other) analytic services, wherein the primary data are deemed the sole property of the instrument provider, while the public researcher is only afforded access to processed data products, which is sometimes limited to the exclusive use of the researcher who leased the instrumentation. In other instances, the public researcher retains access and noncommercial use and sharing rights to the primary data but is contractually obliged to give an exclusive license for commercial use of that data to the private instrument provider.

Agricultural R&D has had long-standing traditions of conducting on-farm research but hitherto often with limited attention given to the privacy concerns associated with the use and sharing of that farm-sourced data. The era of increasing farm instrumentation and (privately provided) analytic services, coupled with the dawn of "big data" applications in agriculture, is rapidly changing farmers' perceptions of the privacy dimensions of farm-sourced data streams, with direct implications for the accessibility and research uses of these data.

For example, on-farm R&D examining the nutrient run-off implications of various cropping and land management practices often involve coupling data generated by commercial yield monitors; machine- and human-sourced crop management practices, perhaps privately sourced third-party weather data; and privately and publicly sourced soil, terrain, and other data. This constellation of data sources and analytic

services, and the various access rights and uses to which these data might be put, illustrates the context dependent nature of the privacy concerns surrounding agricultural data (Nissenbaum, 2004, 2010). If the farm-sourced data were only used to inform the farmer who provided it, their willingness to share may take one form, but if the data were being passed along to third parties (including, say, making the data openly available in support of research publications), then the farmer may well have a different view regarding the privacy attributes of their data. As Nissenbaum (2019) noted, even if the raw data were kept in the private possession of the researcher who collected it, the generation and use of higher-order data (also known as research results) arising from lower-order data (also called primary research data) adds even more nuanced and often highly consequential notions concerning data privacy. In and of itself, farmers may be willing to share (anonymized) yield, soil attributes, crop management, and related data for certain prescribed research purposes. However, while georeferenced versions of these same data elements open up many new analytic opportunities—both via direct application of the data for process modeling purposes or for use as ground-truth data linked with satellite-sourced, remote-sensed data to calibrate artificial intelligence models—revealing identifiable data involving farm boundaries can raise new privacy concerns if farmers fear regulatory consequences associated with the findings arising from such research.

Not only is the notion of (data) privacy context sensitive, as Acquisti et al. (2016) notes, it is also difficult to define and means different things to different people. Nonetheless, Altman (1975) concluded that the many notions of privacy at root pertain to the boundaries between the self and others, between unshared or shared, or, in fact, publicly accessible. Moreover, data privacy, or the different dimensions thereof, often involves critical economic elements regarding the value of information and the distribution of costs and benefits associated with data access and use. Acquisti et al. (2016, 443–444) succinctly observed that "… at its core, the economics of privacy concerns the trade-offs associated with the balancing of public and private spheres between individuals, organizations, and governments.… In some [instances] privacy protection can decrease individual and societal welfare; in others, privacy protection enhances them. Thus, it is not possible to conclude unambiguously whether privacy protection entails a net positive or negative change in purely economic terms: its impact is context specific."

These context-sensitive economic and other privacy attributes pertain as much to data collected in the service of (agricultural) science and innovation as they do to any other forms and uses of data. Thus, it follows that a one-size-fits-all-approach to managing the access and use rights to R&D data is unlikely to be effective or efficient, leading us to offer a more flexible, tiered-risk and standards-based approach to dealing with data privacy concerns. Before doing so, we briefly survey the state of play regarding efforts to protect agricultural data privacy.

3 | MECHANISMS TO PROTECT AGRICULTURAL DATA PRIVACY

The growing number of diverse approaches used to address data privacy concerns related to agricultural data can be daunting (Ferris, 2017; Sanderson et al., 2018; Stubbs, 2016; Wiseman et al., 2019). At a high level, these approaches can be divided into voluntary codes of conduct, laws and regulations, and legally binding contracts (Archer & Delgadillo, 2016). Voluntary measures can be understood as suggested best practices; whereas laws, regulations, and contracts set out mandatory measures that typically include a range of penalties as a result of noncompliance. The CGIAR Platform for Big Data in Agriculture and Responsible Data Guidelines (CGIAR, 2020) is an example of a voluntary measure or a voluntary code of conduct for data practices and is explicitly intended to be "aspirational in nature" and "an aid for responsible decision making" (CGIAR, 2020). The CGIAR guidelines are organized around a standard data life cycle, which gives researchers a familiar framework to apply a mix of high-level (e.g., "Don't ignore ethical practices/standards …") to low-level (e.g., "… use two-factor or multifactor authentication.") good practices. The good practices are presented as 'Tips' for what to do and what not to do (Figure 1). We will use these tips as the basis for our standards mapping described later in this paper.

Similar voluntary codes of conduct have been created to serve specific geographic regions (Table 1). In Europe, eight organizations (European Farmers and European Agri-cooperatives, European Agricultural Machinery Association, European Organisation of Agricultural, Rural and Forestry Contractors, European Space Agency, Fertilizers Europe, European Feed Manufacturers' Federation, European Crop Protection Association, European Forum of Farm Animal Breeders, and European Council Of Young Farmers), each of which is comprised of their own member organizations, recently published the *European Union Code of Conduct on Agricultural Data Sharing by Contractual Agreement* (Anonymous, 2018). The EU Code broadly applies to the agri-food sector and covers a diverse set of data managed and generated by this sector. While the EU code is voluntary, its signatories encourage, "…all parties involved in the agri-food chain to conform according to these jointly agreed principles" (Anonymous, 2018). Similar voluntary codes have also been created in the United States (American Farm Bureau, 2016) and in New Zealand (Anonymous, 2016). Another example of a voluntary or recommended code of conduct includes the recommendations to "Address Privacy and Security" developed by the Principles for Digital Development group

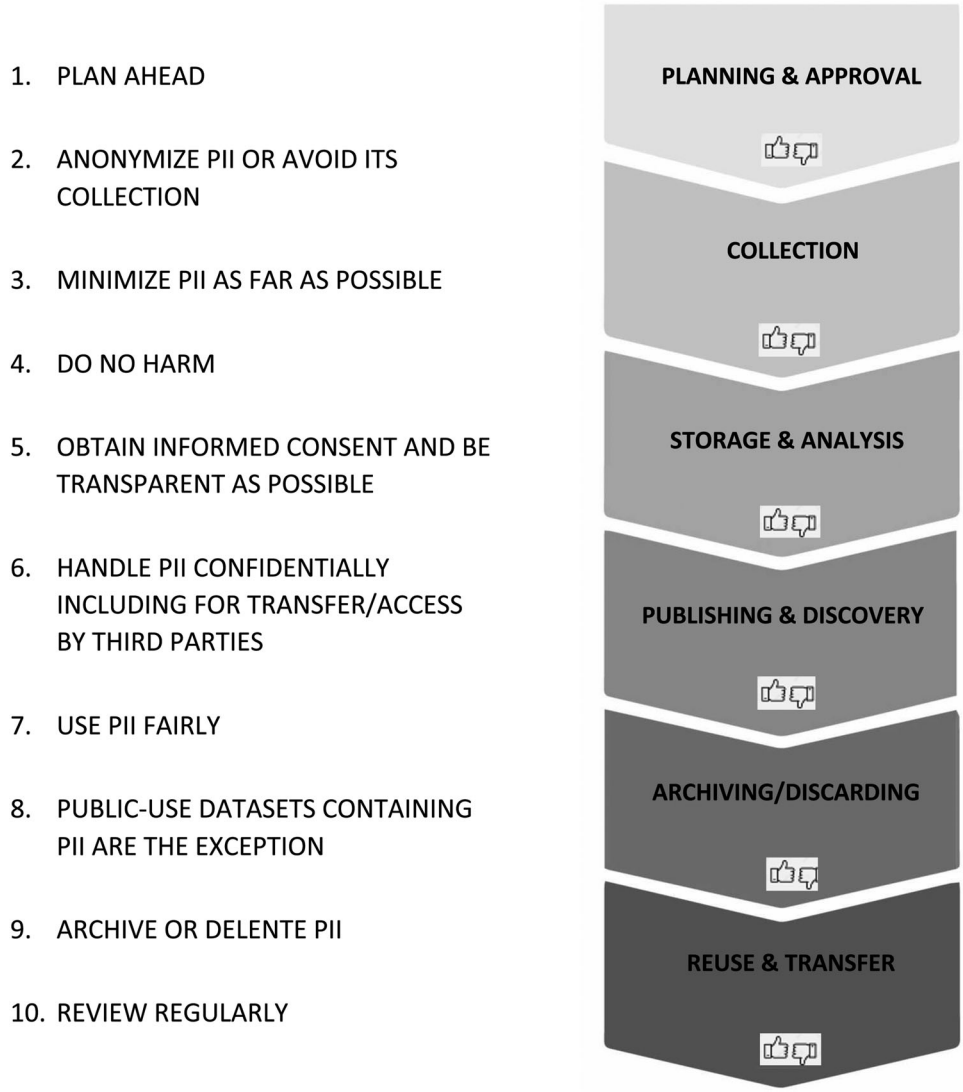

FIGURE 1 A schematic depicting the CGIAR's platform for big data in agriculture-responsible guidelines

(PFDD, 2020). In 2000, the African Union adopted an African Union Convention on Cyber Security and Personal Data Protection (African Union, 2014), and in December 2018, the World Bank Group posted a Personal Data Privacy Policy that was operationalized in May 2020 (Tafara, 2020; World Bank, 2020). None of these guidelines, policies, or conventions make direct mention of food or agriculturally related data and are principally or exclusively concerned with the protection or privacy aspects of personally identifiable data.

The many different forms of voluntary codes of conduct used to protect the privacy of agriculture data make it difficult for stakeholders at all stages of agricultural innovation and production to know how to comply with the growing set of diverse expectations, especially those who operate in a multicountry context (e.g., the CGIAR) or engage in joint research conducted, say, by land grant universities or the USDA involving international partners where data are shared across national borders. Furthermore, it is unclear whether these voluntary codes of conduct are having the desired effect. Sanderson et al. (2018, p. 15) concluded that "…the question of what ag-data codes really achieve remains to be answered." Others are less ambiguous and argue that "…the current regulatory environment is not sufficient to protect sensitive agricultural data…" (Ferris, 2017, p. 331) because state law in the United States is not uniform "…and voluntary industry standards are simply that—voluntary" (Ferris, 2017, p. 331). Beckerman (2019) and Ferris (2017) proposed solving this problem by creating federal regulation aimed specifically at protecting agricultural data in the same way that HIPAA (The Health Insurance Portability and Accountability Act, 1996) governs the healthcare industry and the Gramm–Leach–Bliley Act (1999) regulates the financial services industry in the United States. In 2018, the United States introduced new legislation called the Agriculture Data Act (2018), which would apply to data that are relevant to covered

TABLE 1 Key organizing principles for a sample of voluntary data codes of conduct

CGIAR platform for big data in agriculture & responsible data guidelines	European Union code of conduct on agricultural data sharing by contractual agreement	American Farm Bureau's 'Privacy and Security Principles for Farm data'	New Zealand's 'Farm Data Code of Practice'
Planning and approval; Collection; Storage and analysis; Publishing and discovery; Archiving and discarding; Reuse and transfer	Attribution of the underlying rights to derive data (data ownership); Data access, control, and portability; Data protection and transparency; Privacy and security; Liability and intellectual property rights	Education; Ownership; Collection, access, and control; Notice; Transparency and consistency; Choice; Portability; Terms and definitions; Disclosure use and sales limitation; Data retention and availability; Contract termination; Liability and security safeguards	Disclosures: Corporate Identity; Rights to data; Security standards; Data access; Data Sovereignty
			Practices: Rights to data; Data interchange and access; Security; Regulatory compliance

Note. Source: Developed by authors based on information taken from CGIAR (2020), General Data Protection Regulation (2016), American Farm Bureau (2016), Farm Data Accreditation, Ltd. (2014).

conservation practices. If passed, this law will likely precipitate the development of specific requirements for how the privacy of covered data is protected. Such data protection standards may be relevant to other types of agri-food data, which makes it important to keep track of the development of this bill in the years to come.

While agriculture data are not explicitly protected by law or regulation, some legal and regulatory frameworks can be used to protect the privacy of agriculture data. For example, in the United States, section five of the Federal Trade Commission (FTC) Act (United States, 2018) seeks to protect consumers against unfair or deceptive acts or practices in or affecting commerce and therefore could be used to protect agriculture data. That said, Ferris (2017) argues that there are a number of reasons why this is unlikely to happen in practice. Given the FTC's broad scope and limited resources, Ferris points out that the FTC is more likely to exercise its enforcement activities on high-profile cases where the potential consequences of a violation are very serious and the likelihood for a successful prosecution is very high. Ferris (2017) claims that cases involving agriculture data privacy do not meet these expectations, so there is little reason to believe that FTC enforcement would be an effective legal mechanism to use for the protection and enforcement of agriculture data privacy.

Following high-profile events like the Facebook–Cambridge Analytica scandal, more and more U.S. states are beginning to enact legislation to protect data that are considered private (Beckerman, 2019). While these state-based data privacy laws appear to have the best interest of an individual's privacy in mind, the lack of uniformity in the way data privacy is treated across states is leading to questions and some doubts as to whether state data privacy laws are actually helping to protect privacy in general (Beckerman, 2019; Ferris, 2017). One notable exception at the state level is Minnesota's Agricultural Data statute (Agricultural Data Act, 2018). Similar to the proposed Agricultural Data Act (2018), the Minnesota Agricultural Data statute legally defines a class of agricultural data as private. Such a measure gives the University of Minnesota and the Minnesota Department of Agriculture a way to protect grower (and other identifiable) data from open access requests. This overarching measure of privacy protection has helped to address grower concerns that the data from their farms, which, for example, is provided for research, could be accessed by a competitor or other interested party to obtain an economic advantage or by an environmental organization to seek legal action.

The United States is certainly not alone in enacting legislation around data privacy. Perhaps most notable, starting in 2016, Europe enacted the General Data Protection Regulation (GDPR) (2016). Similar to U.S. law, the GDPR does not explicitly protect agriculture data; rather, the regulation only applies to personal data, which under GDPR is considered as

> "… any information relating to an identified or identifiable natural person ("data subject"); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person." (GDPR, 2016, Art. 4.1).

Therefore, agriculture data can only be protected if the data cannot be separated from personal information (Janzen, 2018). This required link to a person before agriculture data are afforded protection is similar to laws protecting data privacy in China and Brazil, two important countries in agriculture production and data (Archer & Delgadillo, 2016).

The last approach to agricultural data privacy governance that we will briefly discuss is contractual. Archer and Delgadillo (2016) do a thorough job discussing the specific legal elements that should be contained in a contract and they also discuss some of the data-related issues that may arise when organizations engage in a contractual agreement that spans multiple countries. Our concern regarding the use of data licenses or contracts—including the privacy or IP clauses that often now form part of the data management plans embedded in standard research funding or collaboration agreements—to govern agriculture data privacy is the same for any of the other governance mechanisms that we have discussed so far and is the primary focus of this work. That is, contracts and all other means governing data privacy must clearly define the technical and nontechnical standards that will be used to reasonably ensure the privacy of the data. Without such standards, it is hard to know whether a future data privacy breach resulted from lack of adherence to these standards or simply whether the assault on privacy was particularly egregious. In other words, contracts provide a very flexible means to establish these expectations; however, without an unambiguous mapping of these expectations to well-defined standards, the implementation of data protection measures will vary widely and will make it impossible for stakeholders to know how well their data are being protected. Our mapping of the CGIAR guidelines to specific standards discussed in the following section provides a tangible example of how a contractual partnership—including research funding, material transfer, (customized) data use or master research agreements—can establish a mechanism to objectively evaluate whether the requirements of a contract compare favorably with best practices used in other industries. In the absence of laws and regulations for protecting agricultural data privacy, a contractual approach that clearly identifies risks coupled with an explicit map to technical and nontechnical standards used to safeguard data is likely the best approach to establish common expectations and mitigate general risks related to data privacy (Archer & Delgadillo, 2016).

4 | TECHNICAL AND NONTECHNICAL STANDARDS AND PROCEDURES

The development of good or best practices implies that a set of standards already exists by which comparisons to general practices can be made. Standards not only make it possible to objectively order one approach over another, but they also help to unambiguously describe what methods will be used when it comes to protecting data privacy. For example, even a relatively specific sounding action like 'anonymizing data' could mean different things to different people if left without the reference to existing standards and definitions (Nayak et al., 2016). For example, the U.S. NIST outlines the following five ways that data can be anonymized (McCallister et al., 2010, Sect. 4.2.4).

1. Generalizing the data—making information less precise, such as grouping continuous values
2. Suppressing the data—deleting an entire record or certain parts of records
3. Introducing noise into the data—adding small amounts of variation into selected data
4. Swapping the data—exchanging certain data fields of one record with the same data fields of another similar record (e.g., swapping the ZIP codes of two records)
5. Replacing data with the average value—replacing a selected value of data with the average value for the entire group of data.

If left undefined, both the data provider and the data recipient could be very surprised by the results of the deidentification process (see also Massell et al., 2014).

To implement the guidelines-to-technical-standards mapping we conducted and described in this paper, we drew from standards developed by a joint technical committee of the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) and by NIST. The ISO/IEC joint technical committee was formed in 1987, while NIST became the new name of the U.S. National Bureau of Standards in 1988. The beginning of the National Bureau of Standards dates back to 1901. Both of these organizations develop standards that generally transcend political boundaries and thus their work is frequently cited as a means for defining, or at least benchmarking, requirements in other countries. The CGIAR Platform for Big Data in Agriculture and Responsible Data Guidelines (CGIAR, 2020) were mapped to five ISO/IEC standards (specifically, ISO/IEC 27001:2013, ISO/IEC 27002:2013, ISO/IEC 27017:2014, ISO/IEC 27018:2015, ISO/IEC 27701:2019) (ISO/IEC, 2013a, 2013b, 2014, 2015, 2019) and two NIST standards (McCallister et al., 2010; Ross et al., 2017) to unambiguously define what is meant by the more general standard. Before we explicitly mapped the CGIAR guidelines to these standards they could be interpreted in many different ways. A brief description of the standards used in this case study is given in Table 2.

TABLE 2 Sources of technical information for mapping CGIAR guidelines to technical standards

Technical standard	Description
ISO/IEC 27001:2013	Specifies the requirements for establishing, implementing, maintaining, and continually improving an information security management system within the context of the organization.
ISO/IEC 27002:2013	Gives guidelines for organizational information security standards and information security management practices including the selection, implementation and management of controls taking into consideration the organization's information security risk environment(s).
ISO/IEC 27017:2015	Gives guidelines for information security controls applicable to the provision and use of cloud services by providing additional implementation guidance for relevant controls specified in ISO/IEC 27002; additional controls with implementation guidance that specifically relate to cloud services.
ISO/IEC 27018:2014	Establishes commonly accepted control objectives, controls, and guidelines for implementing measures to protect Personally Identifiable Information (PII) in accordance with the privacy principles in ISO/IEC 29100 for the public cloud computing environment.
ISO/IEC 27701:2019	Establishes commonly accepted control objectives, controls, and guidelines for implementing measures to protect Personally Identifiable Information (PII) in line with the privacy principles in ISO/IEC 29100 for the public cloud computing environment.
NIST 800-171	Protecting controlled unclassified information in nonfederal systems and organizations
NIST 800-122	Guide to Protecting the Confidentiality of Personally Identifiable Information (PII)

Note. Source: Developed by authors drawing on information from the ISO/IEC and NIST (McCallister et al., 2010; Ross et al., 2017) standards listed in the table.

5 | STANDARDS-COUPLED SECURITY TIERS TO SAFEGUARD AGRICULTURAL DATA PRIVACY

Safeguarding agricultural data would be relatively easy if there was not an interest in making these data more broadly available to advance R&D objectives that align with the public good. For example, one could simply save the data to a hard drive and place the drive in a locked safe. Where the consequences of failing to make data private are dire, such means may be warranted. However, making data more broadly available for current or future use is more than a mere interest or even guiding principle in much of the agriculture sector. Many public research, funding, and policy agencies either strive for or have mandates to make data open. For example, the Global Open Data for Agriculture and Nutrition initiative, launched in 2012, seeks to lower barriers to agricultural and nutritional data so that anyone in the world may access these data without restriction. The Global Open Data for Agriculture and Nutrition presently has over 700 members, including the USDA and the CGIAR (Adams, 2021; Global Open Data for Agriculture and Nutrition, 2021). There is also an open-access approach inscribed in the policies and guidelines that undergird the development of the CGIAR Platform for Big Data in Agriculture (see, e.g., CGIAR, 2016). As alluded to earlier, a one-size-fits-all approach is neither necessary nor desirable when it comes to applying approaches to data privacy in agriculture. As a demonstration of this approach, we mapped well-defined standards to the CGIAR Responsible Data Guidelines (CGIAR, 2020) that address the guidelines data privacy aspirations while balancing the need for access to the data, as required by the CGIAR Open Access and Data Management policy (CGIAR, 2016).

Fortunately, balancing access to data with the protection of its privacy is not a new enterprise. That is to say, there is a rich set of mature experiences and examples, especially from the healthcare sector (Horvitz & Mulligan, 2015; Lane et al., 2014; O'Keefe & Rubin, 2015; Rodwin & Abramson, 2012), and from U.S.–based research universities (Redd et al., 2019). To face the new demands around data in the agriculture sector, it is critical that we borrow, or at minimum learn, from the experiences coming from other sectors. For example, U.S.–based public research universities are actively working on ways to contend with an expectation of openness to data tempered by an overarching ethical framework that respects, values, and, in some cases, requires data privacy (e.g., University of Minnesota, 2020).

The practice of balancing access rights over data with concerns over data privacy fits firmly into a broader framework of balancing benefits with risks (Stine et al., 2008). A first step in the application of this framework is to broadly classify data according to the pecuniary or nonpecuniary harm that could be caused to individuals (also known as research subjects) and the organization hosting the research if the privacy of these data were to be compromised. Broadly classifying data in this way, while being cognizant of the contextual nature of the types and degrees of harm involved in a data breach, requires that everyone involved at all stages of the research data life cycle be aware of the risks associated with their data and the policies and procedures used by their organization to handle these data. Stakeholders [e.g., the 'key players' in Stubbs (2016) and the 'actors' in Nissenbaum (2019)] must, at a minimum, be able to identify what types

TABLE 3 Three tiers of risk for agricultural data

Low-risk data	Medium-risk data	High-risk data
Data are considered public; Data have been fully de-identified, or subject has consented to make data public; The loss or unintentional alteration of these data would not result in material harm to the subject or institution	Data are considered private; Data have been fully de-identified; The loss or alteration of these data would result in significant material harm to the subject or institution	Data are considered private; Data contain personal identifiable information; The loss or alteration of these data would result in catastrophic material harm to the subject or institution

of data require special treatment and know who can provide help when questions emerge about the data that they are generating or charged with managing (D'Arcy & Greene, 2014; Geller et al., 2010; Hu et al., 2012). This is to say that agriculture research, like nearly all other research domains, requires a team-based approach where experts in areas related to technical data security standards are a part of the team and are consulted throughout the process. The practice of classifying and managing data will also reflect an organization's appetite for risk, so explicitly considering data privacy and security in the context of a standard risk management framework is a necessary first step to good data management practices.

Our focus going forward is to assume that these data have been classified and fall under one of three tiers of risk: high, medium, and low for the protection of agricultural data types (Table 3). These tiers of risk and their associated security protocols map to the putative effect or risk to a research subject and to the organization hosting the research if the privacy of the data protected under each tier were to be compromised either willfully or by failing to meet the relevant standards. This approach is described in great detail in NIST Federal Information Processing Standards Publication 199 (NIST, 2004) (see also Table 4). It is regularly used by practitioners operating in a variety of economic sectors to address data privacy concerns and is a good example of how well-defined frameworks already exist and can be leveraged by the agriculture sector to address data concerns of many types and at many scales.

This approach is especially valuable because it establishes the security categories for both the information (e.g., data) and the systems that host this information. Leveraging the effort related to classifying both the information and the information systems is a common approach because in practice it helps to inform procedures, which in turn are used to implement solutions. For example, it should come as no surprise that information classified as high risk should only be stored on information systems that meet the security standards suitable to host high-risk information. It follows that as the risk to a subject (including the costs arising from a data breach) increase, the standards and procedures used to protect that person's (or organization's) privacy will also become more stringent or strict. While this approach may seem obvious, the goal of explicitly mapping general data privacy guidelines to well-recognized data standards would not be feasible if it were not possible to transfer the tiers of effect from the exercise of classifying information to the practice of managing information systems (e.g., Levenstein et al., 2018; Sweeney et al., 2015). More specifically, the criteria under each data-security tier are used to inform what standards are mapped to a specific guideline as, for example, found under the CGIAR's Responsible Data Guidelines (CGIAR, 2020). As might be expected, systems designed to support low-risk data may have fewer required standards linked to the CGIAR Responsible Data Guidelines than systems designed to support high-risk data. Similar tiered-data security classification schema exist for many U.S. research universities (Supplemental Table S1).

Another less commonly considered dimension of the risk management framework is an organization's security objective (NIST, 2004). For example, if the security objective for the information an organization is charged with managing is data availability, then the technical and nontechnical standards used to safeguard these data will be very different from those whose security objective is to protect confidentiality. This more nuanced approach provides more flexibility simply because practitioners are not required to fit all of their data under a single security objective category. NIST Federal Information Processing Standards Publication 199 (2004) demonstrates what the potential effects to a subject or organization might look like if the security of information were to be compromised by juxtaposing the effects with three security objectives: confidentiality, integrity, and availability (Table 5).

In short, we advocate for a three-tiered system because a finer-grain system with more than three tiers becomes impractical to implement, while a more coarse-grained system with fewer tiers either does not afford sufficient protections to some data or makes protecting data prohibitively burdensome (e.g., expensive and complicated). In our judgement, the number of effective categories and the general risk-management framework described above offer sufficient flexibility to address the majority, if not all, current risks associated with agricultural data. Importantly, by not being overly complicated, this approach also encourages better compliance by more closely reflecting what can be practically implemented

TABLE 4 Description of putative effects from breaching three levels of data security

Potential effect	Definitions
Low	The potential effect is low if the loss of confidentiality, integrity, or availability could be expected to have a limited adverse effect on organizational operations, organizational assets, or individuals.[a]
	A limited adverse effect means that, for example, the loss of confidentiality, integrity, or availability might (a) cause a degradation in mission capability to an extent and duration that the organization is able to perform its primary functions, but the effectiveness of the functions is noticeably reduced; (b) result in minor damage to organizational assets; (c) result in minor financial loss; or (d) result in minor harm to individuals.
Moderate	The potential effect is moderate if the loss of confidentiality, integrity, or availability could be expected to have a serious adverse effect on organizational operations, organizational assets, or individuals.
	A serious adverse effect means that, for example, the loss of confidentiality, integrity, or availability might (a) cause a significant degradation in mission capability to an extent and duration that the organization is able to perform its primary functions, but the effectiveness of the functions is significantly reduced; (b) result in significant damage to organizational assets; (iii) result in significant financial loss; or (c) result in significant harm to individuals that does not involve loss of life or serious life threatening injuries.
High	The potential effect is high if the loss of confidentiality, integrity, or availability could be expected to have a severe or catastrophic adverse effect on organizational operations, organizational assets, or individuals.
	A severe or catastrophic adverse effect means that, for example, the loss of confidentiality, integrity, or availability might (a) cause a severe degradation in or loss of mission capability to an extent and duration that the organization is not able to perform one or more of its primary functions; (b) result in major damage to organizational assets; (c) result in major financial loss; or (d) result in severe or catastrophic harm to individuals involving loss of life or serious life threatening injuries.

Note. Source: Developed by authors based on information from Stine at al. (2008).

[a]Adverse effects on individuals may include, but are not limited to, loss of the privacy to which individuals are entitled under law.

by various stakeholders and enforced by various oversight groups.

6 | A CASE STUDY: MAPPING THE CGIAR PLATFORM'S RESPONSIBLE DATA GUIDELINES TO TECHNICAL STANDARDS

The lack of clear a map between data privacy guidelines and a set of standards is among the causes of confusion for practitioners working to provide data services to the agricultural sector and contributes to the lack of trust among various data producers and owners (Wiseman et al., 2019). Mapping the general guidelines contained in various frameworks used to govern data security to specific technical and nontechnical standards helps to resolve the ambiguity and codifies what will actually be done to protect the privacy of data. To illustrate an application of this approach, we map the 10 high-level functional guidelines described under the CGIAR's Responsible Data Guidelines (CGIAR, 2020) to specific standards and regulations (Table 2) based on the potential effect that could result if the confidentiality of the data (Table 3) that the CGIAR is entrusted with managing were to be compromised (Figure 2). The product of this approach is a mapping of each guideline (or guideline subcategory if one existed) to relevant technical standards for each security tier (Figure 3). Table 6 is a sample of the CGIAR's Responsible Data Guidelines (CGIAR, 2020) and the standards to which they are mapped. Because of the large number of guidelines, the complete mapping is presented in Supplemental Table S2.

More concretely, Table 6 lays out the guidelines to standards mapping. The first three columns identify the data management function, high-level guideline, and, where relevant, a more specific subguideline (also referred to as 'tips') as described in the CGIAR's Responsible Data Guidelines (CGIAR, 2020). Columns four and five identify the specific standard (Table 2) according to the three risk categories (Table 3). In all cases, there was an agreed set of standards for each of the CGIAR guidelines, and in most cases, we were also able to differentiate those standards into the three risk tiers. Standards identified in a lower tier extend to the tiers above them. Therefore, if no standard is listed in the high category, it is because it inherits the low or medium standards.

The format and even the overarching goal of this approach draws heavily on examples from other fields especially the healthcare profession. Our mapping closely reflects the crosswalk approach taken by the DHHS (Table 7) to manage the privacy of healthcare data, in which higher-level data management functions contain categories and subcategories of increasingly specific recommended practices, which are mapped to specific standards. We recognize that there are differences between the privacy concerns for healthcare data and the privacy concerns for agriculture data. That said, there are also many commonalities along with important lessons that can be learned by examining and selectively using approaches

TABLE 5 Data security objectives and the tiered consequences of a data breach

Security objective	Potential effect		
	Low	**Moderate**	**High**
Confidentiality			
Preserving authorized restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information.	[44 U.S.C., SEC. 3542] The unauthorized disclosure of information could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized disclosure of information could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals. The unauthorized disclosure of information could be expected to have a **severe or catastrophic**	adverse effect on organizational operations, organizational assets, or individuals.
Integrity			
Guarding against improper information modification or destruction, and includes ensuring information nonrepudiation and authenticity.	[44 U.S.C., SEC. 3542] The unauthorized modification or destruction of information could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized modification or destruction of information could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized modification or destruction of information could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.
Availability			
Ensuring timely and reliable access to and use of information.	[44 U.S.C., SEC. 3542] The disruption of access to or use of information or an information system could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals	The disruption of access to or use of information or an information system could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The disruption of access to or use of information or an information system could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.

Note. Source: Taken from NIST (2004, p. 6).

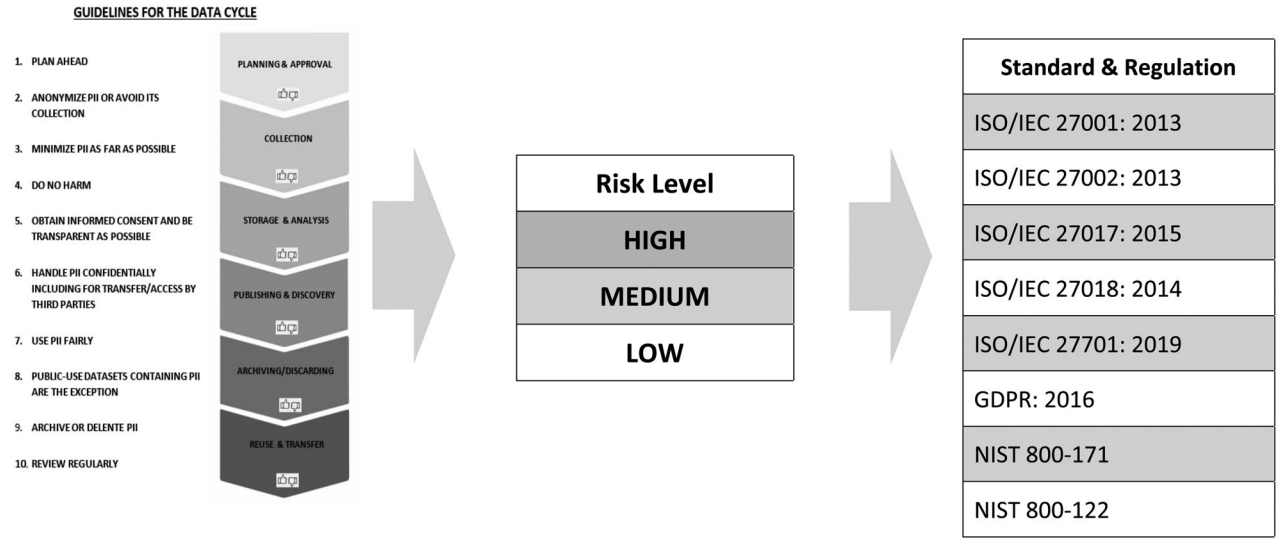

FIGURE 2 A risk-based schema for mapping high-level CGIAR responsible data guidelines to specific data standards and regulations

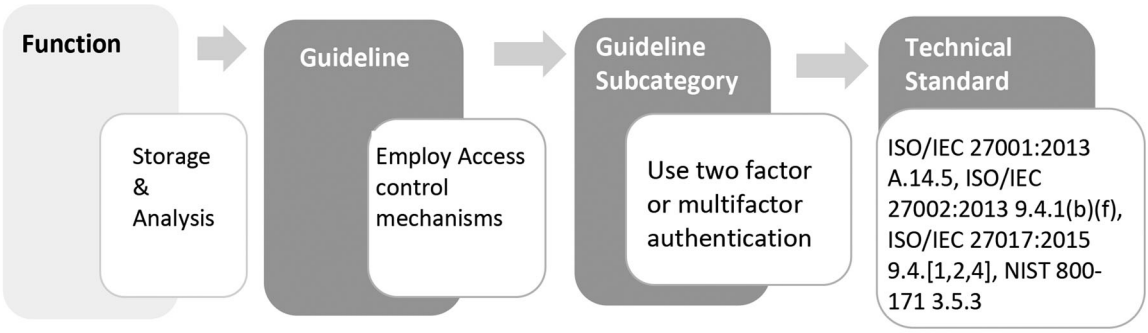

FIGURE 3 An example of mapping CGIAR data functions and guidelines to specific data standards and regulations

TABLE 6 A sample of the CGIAR's responsible data guidelines mapped to relevant technical and nontechnical standards for three security tiers. Note: The complete set of guidelines mapped to standards can be found in Supplemental Table S2 in the online supporting material

Function	Guideline	Subcategory	Risk level	Applicable standards
Planning and approval	Create a data management plan	Compliance requirements (including necessary forms for obtaining consent, and ethics clearance, if applicable)	H	NIST 800-122:4.1.1; ISO/IEC 27701:2019 A.7.2.2, A.7.3.1
			M	GDPR Art 6
			L	ISO/IEC 27001: A.18.1.1
Collect	If you cannot anonymize, minimize the PII and pseudonymize to reduce the disclosure risk	–	H	NIST 800-122:4.2.3; ISO/IEC 27701:2019 A.7.4.5
			M	GDPR 25
			L	–
Storage & analysis	Ensure appropriate IT & security controls to protect confidentiality of PII at rest and in transit	Store data in secure locations, devices, or servers	H	GDPR Art 5.5(f), 6, 24.2, 27.1, 27.2(a-b), 27.[3-5], 32.1(a-c), 32.2, (shortened); ISO/IEC 27018:2014 9-12 (Shortened); NIST 800-122 3.2.6, 4.2.1, 4.3
			M	ISO/IEC 27002:2013 6.1.[1-2], 6.2.[1-2], 8.1.[1-4], 8.2.[1-3], (shortened); ISO/IEC 27017:2015 8.1.1, 8.1.2, (shortened); NIST 800-171 3.1.[1-22],3.4.[1-9],3.5.[1-11],3.7.[1-6] (Shortened)
			L	ISO/IEC 27001:2013 A.8.1.[1-4], A.8.2.[1-3], A.8.3.[1-3] (shortened); ISO/IEC 27017:2015 8.1.1, 8.1.2, 8.2.2, 10.1.1, 10.1.2, (shortened) 18.1.5; NIST 800-171 3.1.[1-2],(shortened)

Note. Source: Developed by authors.

to data privacy and security from other sectors of the economy, and this is clearly one that translates well. Another important lesson learned from the healthcare sector is that practical guidance describing more specifically how to protect certain types of healthcare data lagged well behind the laws designed to protect these data. This is an important consideration especially in light of the many voluntary codes of conduct and the increasing number of contractual arrangements with customized data use agreements that currently govern the use of a great deal of agricultural related data, be that data obtained from U.S. or international sources.

Beginning to define and standardize approaches for safeguarding agricultural data is a good first step. Not only does a standards-based approach promote trust between data producers or owners and agriculture researchers, it also helps application developers and data repositories managers build

TABLE 7 An example of mapping HIPAA security standards and implementation specifications to NIST cybersecurity categories

Function	Category	Subcategory	Relevant control mappings[a]
PROTECT	–	PR.DS-7: The development and testing environment(s) are separate from the production environment	COBIT 5 BAI07.04; ISO/IEC 27001:2013 A.12.1.4; NIST SP 800-53 Rev. 4 CM-2; HIPAA Security Rule 45 C.F.R. § 164.308(a)(4)[4]
	Information protection processes and procedures (PR.IP): Security policies (that address purpose, scope, roles, responsibilities, management commitment, and coordination among organizational entities), processes, and procedures are maintained and used to manage protection of information systems and assets.	PR.IP-1: A baseline configuration of information technology/industrial control systems is created and maintained	CCS CSC 3, 10; COBIT 5 BAI10.01, BAI10.02, BAI10.03, BAI10.05; ISA 62443-2-1:2009 4.3.4.3.2, 4.3.4.3.3; ISA 62443-3-3:2013 SR 7.6; ISO/IEC 27001:2013 A.12.1.2, A.12.5.1, A.12.6.2, A.14.2.2, A.14.2.3, A.14.2.4; NIST SP 800-53 Rev. 4 CM-2, CM-3, CM-4, CM-5, CM-6, CM-7, CM-9, SA-10; HIPAA Security Rule 45 C.F.R. §§ 164.308(a)(8), 164.308(a)(7)(i), 164.308(a)(7)(ii)

Note. Source: Department of Health and Human Services (2014).

[a]Mappings to other standards come from the NIST Cybersecurity Framework, Appendix A and are provided for reference.

systems that maximize the benefits of the current agricultural data deluge while also respecting and protecting its privacy. For example, Sweeney et al. (2015) described a tag-based approach whereby existing technologies could be used to scale this approach for a number of different disciplines. Likewise, the measures described in this paper and applied to the CGIAR Platform's Responsible Data Guidelines can be used more generally in a wide range of emerging agricultural data governance approaches that increasingly form part of research funding or collaboration agreements.

7 | CONCLUSION

Agricultural innovation is a global enterprise, and data centric approaches are increasingly being used to unlock the gains that we should expect from these innovations. That said, justifiable concerns regarding the way that the privacy of agriculture data is or is not being protected limits the availability of these data for the public good. Such concerns highlight the need for the development of consensus concerning how data privacy is governed in the agricultural sector. More to the point, this consensus will likely come by mapping data management guidelines to well-known technical and nontechnical standards so that data owners better understand how their privacy is being protected and can more objectively compare one approach with another. This mapping of guidelines to standards should be considered in light of the context sensitive pecuniary or nonpecuniary risks associated with a data privacy breach. Mappings of agriculture data privacy guidelines, such as the CGIAR's Responsible Data Guidelines (CGIAR, 2020), to technical standards are at best rare and to our knowledge do not exist, which may make this mapping the first instance of its kind.

The agriculture sector involves a large and diverse set of stakeholders and so a one-size-fits-all approach to addressing privacy will not effectively address stakeholder concerns. Moreover, it will likely inhibit rather than advance the movement of data into the public domain by applying strong technical and administrative controls where they may not be needed. The implementation of a standards-based approach to data access and management is best considered by aligning the type of data being managed to the degree of risks posed to stakeholders if those data were no longer private. While this tiered approach is more nuanced, frameworks for assessing such risks and applying appropriate standards to manage the privacy of data already exist and, as we illustrate in this paper, can be used as a transparent and practical basis for addressing data privacy concerns in the agricultural sector. The specific implementation details will certainly differ when applied to cases related to agriculture, but the overarching framework does not need to be reinvented. Waiting to do something should not be considered an option. Data management guidelines linked to common technical standards that are considered in light of tiered risks posed by the consequences of breaches in data security will help develop trust among stakeholders and advance the innovation promised by the Big Data revolution in agriculture.

ACKNOWLEDGMENTS

The content of this paper benefited greatly from the comments of two anonymous reviewers and discussions with participants at the Big Data Workshop for Agriculture titled, "BD AI, Blockchain, Workforce Development and Data Privacy", held at the Microsoft Campus, Fargo, North Dakota on September 10, 2019; the 2019 CGIAR Big Data in Agriculture Platform Convention titled "Trust: Humans, Machines and Ecosystems" held at ICRISAT, Hyderabad, India on

October 18, 2019; and the Data Ownership Dialog, co-located with the ASA-CSSA-SSSA International Annual Meeting held in San Antonio, Texas held on November 13, 2019. We are also thankful for the support from Brian King and Connie Chan-Kang in the preparation of this work. Additional support was provided by the GEMS Informatics Initiative at the University of Minnesota and CGIAR Platform for Big Data in Agriculture.

AUTHOR CONTRIBUTIONS

James C. Wilgenbusch: Conceptualization; Methodology; Supervision; Writing – original draft; Philip G. Pardey: Conceptualization; Project administration; Supervision; Writing – review & editing; Naomi Hospodarsky: Data curation; Formal analysis; Methodology; Resources; Writing – review & editing; Benjamin J. Lynch: Data curation, Resources, Software, Writing – review & editing.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

ORCID

James C. Wilgenbusch https://orcid.org/0000-0001-9464-1578
Philip G. Pardey https://orcid.org/0000-0002-8012-1341

REFERENCES

Acquisti, A., Taylor, C., & Wagman, L. (2016). The economics of privacy. *Journal of Economic Literature*, *54*, 442–492. https://doi.org/10.1257/jel.54.2.442

Adams, J. (2021). *Open data: Enabling fact-based, data-driven decisions*. USDA Blog Archives. https://www.usda.gov/media/blog/2018/07/13/open-data-enabling-fact-based-data-driven-decisions

African Union. (2014). *African Union convention on cyber security and personal data protection*. African Union. https://au.int/en/treaties/african-union-convention-cyber-security-and-personal-data-protection

Agricultural Data Act, Minnesota Stat. § 13.643 (2018). https://www.revisor.mn.gov/statutes/cite/13.643

Agriculture Data Act of 2018 to Food Security Act of 1985, S. B. 2487, 115th Cong. (2018). https://www.congress.gov/bill/115th-congress/senate-bill/2487/text

Alston, J. M., & Pardey, P. G. (2021). The economics of agricultural innovation. In C. B. Barrett & D. R. Just (Eds). *Handbook of agricultural economics* (vol. 5). Elsevier.

Altman, I. (1975). *The environment and social behavior: Privacy, personal space, territory, crowding*. Brooks/Cole Publishing.

American Farm Bureau. (2016). *Privacy and security principles for farm data*. American Farm Bureau. https://www.fb.org/issues/technology/data-privacy/privacy-and-security-principles-for-farm-data

Anonymous. (2016). *New Zealand farm data code of practice, version 1.1*. http://www.farmdatacode.org.nz/wp-content/uploads/2016/03/Farm-Data-Code-of-Practice-Version-1.1_lowres_singles.pdf

Anonymous. (2018). *EU Code of conduct on agricultural data sharing by contractual agreement*. http://cema-agri.org/sites/default/files/publications/EU_Code_2018_web_version.pdf

Archer, J. K., & Delgadillo, C. A. (2016). Key data ownership, privacy and protection issues and strategies for the international precision agriculture industry. *13th International Conference on Precision Agriculture*. 31 Jul.–4 Aug. 2016. St. Louis, Missouri. International Society of Precision Agriculture. https://hbfiles.blob.core.windows.net/files/2f53c518-a374-460f-a40e-a82ace4b8605.pdf

Beckerman, M. (2019). Americans will pay a price for state privacy laws. *The New York Times*. https://www.nytimes.com/2019/10/14/opinion/state-privacy-laws.html

Binenbaum, E., Nottenburg, C., Pardey, P. G., Wright, B. D., & Zambrano, P. (2003). South-North trade, intellectual property jurisdictions, and freedom to operate in agricultural research on staple crops. *Economic Development and Cultural Change*, *51*, 309–335. https://doi.org/10.1086/346177

CGIAR. (2016). *CGIAR open access and data management policy*. CGIAR System Management Office. https://cgspace.cgiar.org/handle/10947/4488

CGIAR. (2020). *Responsible data guidelines: Managing privacy and personally identifiable information in the research project data lifecycle*. CGIAR Platform for Big Data in Agriculture. https://bigdata.cgiar.org/responsible-data-guidelines/

D'Arcy, J., & Greene, G. (2014). Security culture and the employment relationship as drivers of employees' security compliance. *Information Management & Computer Security*, *22*, 474–489. https://www.emerald.com/insight/content/doi/10.1108/IMCS-08-2013-0057/full/html

Department of Health and Human Services. (2014). *HIPAA security rule crosswalk to NIST cybersecurity framework*. https://www.hhs.gov/sites/default/files/nist-csf-to-hipaa-security-rule-crosswalk-02-22-2016-final.pdf

Farm Credit Canada. (2019). *Producers embrace technology, but want control over their data, FCC Survey*. https://www.fcc-fac.ca/en/about-fcc/media-newsroom/news-releases/2019/producers-embrace-technology-but-want-control-over-their-data.html

Farm Data Accreditation, Ltd. (2014). *New Zealand farm data code of practice*. ver 1.1, Cl 4. http://www.farmdatacode.org.nz/wp-content/uploads/2016/03/Farm-Data-Code-of-Practice-Version1.1_lowres_singles.pdf

Ferris, J. L. (2017). Data privacy and protection in the agriculture industry: Is federal regulation necessary? *Minnesota Journal of Law Science & Technology*, *18*. https://scholarship.law.umn.edu/cgi/viewcontent.cgi?article=1422&context=mjlst

Geller, G., Boyce, A., Ford, D. E., & Sugarman, J. (2010). Beyond 'Compliance': The role of institutional culture in promoting research integrity. *Academic Medicine*, *85*, 1296–1302. https://doi.org/10.1097/ACM.0b013e3181e5f0e5

GEMS Informatics Initiative. (2021). *FAIR(ER) Data*. University of Minnesota. https://agroinformatics.org/features/fair2-data/

General Data Protection Regulation. (2016). *General data protection regulation, official legal text*. European Commission. https://gdpr-info.eu/

Global Open Data for Agriculture & Nutrition. (2021). *GODAN Partners*. GODAN. https://www.godan.info/partners-logos

Goldfarb, A., & Tucker, C. (2019). Digital Economics. *Journal of Economic Literature*, *57*, 3–43. https://doi.org/10.1257/jel.20171452

Gramm–Leach–Bliley Act, Pub. L. No. 106-102 (Nov. 12, 1999), 113 Stat. 1338 (1999). https://www.govinfo.gov/content/pkg/PLAW-106publ102/pdf/PLAW-106publ102.pdf

Health Insurance Portability and Accountability Act of 1996 (HIPAA), Pub. L. No. 104-191, 100 Stat. 2548 (1996). https://www.govinfo.gov/app/details/STATUTE-110/STATUTE-110-Pg1936/summary

Health Insurance Reform: Security Standards 2003, 68 Fed. Reg. 8333 (February 20, 2003) (to be codified at 45 CFR Parts 160, 162, & 164). https://www.govinfo.gov/app/details/FR-2003-02-20/03-3877/summary

Herbold-Swalwell, E. (2018). *Ownership of your data a big deal*. Farm Progress. https://www.farmprogress.com/regulatory/ownership-your-data-big-deal

Horvitz, E., & Mulligan, D. (2015). Data, privacy, and the greater good. *Science*, *349*, 253–255. https://doi.org/10.1126/science.aac4520

Hu, Q., Dinev, T., Hart, P., & Cooke, D. (2012). Managing employee compliance with information security policies: The critical role of top management and organizational culture. *Decision Sciences*, *43*, 615–60. https://doi.org/10.1111/j.1540-5915.2012.00361.x

International Organization for Standardization– International Electrotechnical Commission. (2013a). *Information technology—Security techniques—Information security management systems—Requirements* (ISO Standard No. 27001:2013). https://www.iso.org/standard/54534.html

International Organization for Standardization– International Electrotechnical Commission. (2013b). *Information technology—Security techniques—Code of practice for information security controls* (ISO Standard No. 27002:2013). https://www.iso.org/standard/54533.html

International Organization for Standardization– International Electrotechnical Commission. (2014). *Information technology—Security techniques—Code of practice for protection of personally identifiable information (PII) in public clouds acting as PII processors* (ISO Standard No. 27008:2014). https://www.iso.org/standard/76559.html

International Organization for Standardization– International Electrotechnical Commission. (2015). *Information technology—Security techniques—Code of practice for information security controls based on ISO/IEC 27002 for cloud services* (ISO Standard No. 27017:2015). https://www.iso.org/standard/43757.html

International Organization for Standardization– International Electrotechnical Commission. (2019). *Security techniques—Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management—Requirements and guidelines* (ISO Standard No. 27701:2019) https://www.iso.org/standard/71670.html

Janzen, T. (2018). *Do GDPR protections extend to ag data?*. Janzen Schroeder Agricultural Law LLC. http://www.aglaw.us/janzenaglaw/2018/5/1/gdprs-impacts-on-ag-data-platforms

Janzen, T. (2019). *Do farmers still care about ag data privacy?*. Janzen Schroeder Agricultural Law LLC. https://www.aglaw.us/janzenaglaw/2019/1/3/farmers-care-about-data

Jones, C. I., & Tonetti, C. (2020). Nonrivalry and the economics of data. *American Economic Review*, *110*, 2819–2858. https://doi.org/10.1257/aer.20191330

Lane, J., Stodden, V., Bender, S., & Nissenbaum, H. (Eds). (2014). *Privacy, big data, and the public good: Frameworks for engagement*. Cambridge University Press.

Levenstein, M. C., Tyler, A. R. B., & Bleckman, J. D. (2018). The researcher passport: Improving data access and confidentiality protection*:* ICPSR's strategy for a community-normed system of digital identities of access. ICPSR White Paper Series No.1. University of Michigan Inter-university Consortium for Political and Social Research. https://deepblue.lib.umich.edu/handle/2027.42/143808

Massell, P. B., Freiman, M. H., & McKenna, L. V. (2014). *Data masking for disclosure limitation*. U.S. Census Bureau Working Paper. https://www.census.gov/library/working-papers/2014/adrm/data-masking-disclosure-limitation.html

McCallister, E., Grance, T., & Kent, K. (2010). *Guide to protecting the confidentiality of personally identifiable information (PII)*. NIST Special Publication 800-122. National Institute of Standards and Technology. https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-122.pdf

McIntosh, M. (2018). *The Legal mess of farm data ownership*. Farmtario. https://farmtario.com/machinery/the-legal-mess-of-farm-data-ownership/

National Academies of Sciences Engineering and Medicine. (2019). *Science breakthroughs to advance food and agricultural research by 2030*. The National Academies Press. https://www.nap.edu/catalog/25059/science-breakthroughs-to-advance-food-and-agricultural-research-by-2030

National Institute of Standards and Technology. (2004). *Standards for security categorization of federal information and information systems*. FIPS Publication 199. National Institute of Standards and Technology. https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.199.pdf

National Science Foundation. (2002). *Dissemination and sharing of research results*. National Science Foundation. https://www.nsf.gov/bfa/dias/policy/dmp.jsp

Nayak, T. K., Zhang, C., & You, J. (2016). *Measuring identification risk in microdata release and its control by post-randomization*. Center for Disclosure Avoidance Research. U.S. Census Bureau. https://www.census.gov/content/dam/Census/library/working-papers/2016/adrm/cdar2016-02-measuring-identification-risk-in.pdf

Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review*, *79*, 119–157. https://digitalcommons.law.uw.edu/wlr/vol79/iss1/10

Nissenbaum, H. (2010). *Privacy in context technology, policy, and the integrity of social life*. Stanford University Press.

Nissenbaum, H. (2019). Contextual integrity up and down the data food chain. *Theoretical Inquiries in Law*, *20*, 221–256. https://doi.org/10.1515/til-2019-0008

Nottenburg, C., Pardey, P. G., & Wright, B. D. (2002). Accessing other people's technology for non-profit research. *The Australian Journal of Agricultural and Resource Economics*, *46*, 289–416. https://doi.org/10.1111/1467-8489.00185

O'keefe, C. M., & Rubin, D. B. (2015). Individual privacy versus public good: Protecting confidentiality in health research. *Statistics in Medicine*, *34*, 3081–3103. https://doi.org/10.1002/sim.6543

Principles for Digital Development. (2020). *Principles: Address Privacy and Security*. https://digitalprinciples.org/principle/address-privacy-security/

Redd, K. J., Steen, K., Nusser, S., Smith, T., Walters, T., Chasen, J., Luther, J., & Reecy, J. (2019). *Accelerating public access to research data workshop*. Association of Public and Land-grant Universities and Association of American Universities. https://www.aplu.org/projects-and-initiatives/research-science-and-technology/public-access/workshop-on-public-access-report-aplu-aau-2019.pdf

Rodwin, M. A., & Abramson, J. D. (2012). Clinical trial data as a public good. *JAMA*, *308*, 871–872. https://doi.org/10.1001/jama.2012.9661

Ross, R., Viscuso, P., Guissanie, G., Dempsey, K., & Riddle, M. (2017). *Protecting controlled unclassified information in nonfederal systems and organizations*. NIST Special Publication 800-171. National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-171

Sanderson, J., Wiseman, L., & Poncini, S. (2018). What's behind the ag-data logo? An examination of voluntary agricultural-data codes of practice. *International Journal of Rural Law and Policy*, *2018*, 6043. https://doi.org/10.5130/ijrlp.1.2018.6043

Standards for Privacy of Individually Identifiable Health Information 2002, 67 Fed. Reg. 53181 (October 15, 2002) (to be codified at 45 CFR Parts 160 & 164). https://www.govinfo.gov/app/details/FR-2002-08-14/02-20554

Stine, K., Kissel, R., Barker, W. C., Fahlsing, J., & Gulick, J. (2008). *Volume I: Guide for mapping types of information and information systems to security categories*. NIST Special Publication 800-60. National Institute of Standards and Technology. https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-60v1r1.pdf

Stubbs, M. (2016). *Big data in U.S. agriculture*. CRS Report. Congressional Research Service. https://crsreports.congress.gov/product/details?prodcode=R44331

Sweeney, L., Crosas, M., & Bar-Sinai, M. (2015). Sharing sensitive data with confidence: The datatags system. *Technology Science*, 2015101601. http://techscience.org/a/2015101601

Tafara, E. (2020). *The importance of protecting 'Privacy' in the age of digital data*. Digital Development. https://blogs.worldbank.org/digital-development/importance-protecting-privacy-age-digital-data

University of Minnesota. (2020). *Know your data and how to protect university data*. University of Minnesota. https://it.umn.edu/good-practice/know-your-data-how-protect-university

United States. (2018). Federal Trade Commission Act, Section 5 Unfair or Deceptive Acts or Practices. https://www.fdic.gov/regulations/compliance/manual/7/vii-1.1.pdf

U.S. Agency for International Development. (2021). *Chapter 579: USAID development data*. USAID. https://www.usaid.gov/ads/policy/500/579

USDA, National Institute for Food and Agriculture. (2019). *Data management plan for NIFA-funded research, education, and extension project*. USDA-NIFA. https://nifa.usda.gov/sites/default/files/resource/data-management-plan-for-research-education-extension-projects-20190926.pdf

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., …, & Mons, B. (2016). The fair guiding principles for scientific data management and stewardship. *Scientific Data*, *3*, 160018. https://doi.org/10.1038/sdata.2016.18

Wiseman, L., Sanderson, J., Zhang, A., & Jakku, E. (2019). Farmers and their data: An examination of farmers' reluctance to share their data through the lens of the laws impacting smart farming. *NJAS - Wageningen Journal of Life Sciences*, *90–91*, 100301. https://doi.org/10.1016/j.njas.2019.04.007

World Bank. (2020). *Personal data privacy*. World Bank. https://ppfdocuments.azureedge.net/ca36fdc4-5191-4d89-a49d-6189a98bad86.pdf

Wright, B. D., & Pardey, P. G. (2006). The evolving rights to intellectual property protection in the agricultural biosciences. *International Journal of Technology and Globalisation*, *2*, 12–29. https://doi.org/10.1504/IJTG.2006.009124

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Wilgenbusch, J. C., Pardey, P. G., Hospodarsky, N., & Lynch, B. J. (2022). Addressing new data privacy realities affecting agricultural research and development: A tiered-risk, standards-based approach. *Agronomy Journal*, *114,* 2653–2668. https://doi.org/10.1002/agj2.20968