Received: 16 September 2021 | Accepted: 30 November 2021

DOI: 10.1002/agj2.20974

Agronomy Journal

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

Cultivating trust in technology-mediated sustainable agricultural research

Ankita Raturi[1] | **Jennifer J. Thompson**[2] | **Victoria Ackroyd**[3,4] | **Carlene A. Chase**[5] | **Brian W. Davis**[4] | **Robert Myers**[6] | **Aurelie Poncet**[7] | **Paula Ramos-Giraldo**[8] | **Chris Reberg-Horton**[8] | **Roderick Rejesus**[8] | **Alison Robertson**[9] | **Matthew D. Ruark**[10] | **Sarah Seehaver-Eagen**[8] | **Steven Mirsky**[4]

[1] Agricultural and Biological Engineering, Purdue Univ., 225 S. University St., West Lafayette, IN 47907-2050, USA

[2] Crop and Soil Sciences, Univ. of Georgia, 3111 Miller Plant Sciences, Athens, GA 30602, USA

[3] Dept. of Plant Sciences & Langscape Architecture, Univ. of Maryland, 4291 Fieldhouse Dr., College Park, MD 20742, USA

[4] Beltsville Agricultural Research Center, USDA ARS, 10300 Baltimore Ave., Beltsville, MD 20705, USA

[5] Horticultural Sciences Dep., Univ. of Florida, PO Box 110690, Gainesville, FL 32611-0690, USA

[6] Center for Regenerative Agriculture, Univ. of Missouri, 238 Ag Engineering Building, Columbia, MO 65211, USA

[7] Dept. of Crop, Soil, and Environmental Sciences, Univ. of Arkansas System, 1352 West Altheimer Dr., Little Rock, AR 72207, USA

[8] Crop and Soil Sciences, North Carolina State Univ., Williams Hall 1114A, Raleigh, NC 27695, USA

[9] Plant Pathology, Iowa State Univ., 351 Bessey Hall, Ames, IA 50010, USA

[10] Soil Science, Univ. of Wisconsin, 1525 Observatory Drive, Madison, WI 53706, USA

Correspondence
Ankita Raturi, Agricultural and Biological Engineering, Purdue Univ., 225 S. University Dr., West Lafayette, IN 47907-2050, USA.
Email: ankita@purdue.edu

Assigned to Associate Editor David Clay.

Abstract

We formed the Precision Sustainable Agriculture (PSA) team to conduct interdisciplinary research and technology development to improve adoption and practice of knowledge-intensive sustainable agricultural practices such as cover cropping. In this paper, we share our approach to cultivating trust among diverse stakeholders (researchers, farmers, extensionists, agricultural and information specialists, private and public entities) vested in agricultural data collection, management, and use. Our trust framework describes how we aim to be trusted with data (through preserving privacy and increasing stakeholder agency) and trusted in the process (through practicing transparency and accountability). It is operationalized through a series of social and technical infrastructures. Our project governance, stakeholder engagement tools and activities, and technology development methods aim to promote transparency and accountability in our process. We use a maturity model to govern data acquisition to ensure that only robust, privacy-preserving technologies are deployed on our partner farms and describe evolving mechanisms for handling data with varying sensitivity. Finally, we share preliminary work aimed at anticipating data use, and

Abbreviations: IRB, Institutional Review Board.; PSA, Precision Sustainable Agriculture.

identify challenges on the horizon for cultivating trust in agricultural technologies and data-driven agriculture.

1 | INTRODUCTION

Crop production faces significant challenges that destabilize food security—including pests' resistance to existing management strategies, climate change, and declining soil health. Cover crops are multi-functional sustainable agriculture tools that hold great promise for mitigating the effects of these destabilizing factors. They provide an array of agroecosystem services, such as improved soil and water quality, weed management, and enhanced nutrient cycling, which can promote cropping system resiliency and stability (Delgado et al., 2007). However, cover crops add complexity to agricultural systems (Clark, 2008), driving the need to better understand how climate, soil, and management practices interact to affect cover crop performance and subsequently affect cash crops and the environment (Kaye & Quemada, 2017). These complex processes ultimately must be shared with farmers in ways that empower them to improve their practice of sustainable agriculture.

In short, optimizing biological tools like cover crops is knowledge intensive: it requires farmer-friendly decision support tools that provide timely, scientifically grounded, and site-specific management recommendations. By bringing a highly coordinated transdisciplinary team to tackle these challenges—that is, a team science approach (Bennet et al., 2018)—our Precision Sustainable Agriculture (PSA) team aims to develop a precision management framework that connects farms, data, tools, and people to optimize sustainable agricultural decision-making. Meeting this objective requires the development of a robust social and technical infrastructure to cultivate trust in our approach to collecting, managing, sharing, and disseminating a massive amount of highly heterogeneous data.

Design of technology and practices to address these challenges requires a systems approach that considers intrinsic factors involved in managing cover crops while responding to the complex social, technological, and economic contexts of farming. Our work integrates heterogeneous agricultural data to inform the design of information tools for farmers and their advisors and, thus, enable the transition of cover crop research from theory to practice. The complex network of actors and stakeholders (researchers, farmers, extensionists, agricultural and information specialists, private and public entities) participating in data creation, sharing, and use necessitates trust across agricultural and technological communities. Trust within a community is an inherently social process (Lewis & Weigart, 1985), but with the increasing use of digital technologies to facilitate communication, data sharing, and other agricultural activities, there is a need for technologies that also "cultivate trusting relationships" such as those that preserve privacy, recognize data ownership and agency, support transparent and agreed-upon data sharing and access, and ensure accountability to our stakeholders.

This paper reports on our ongoing efforts to develop robust data management technologies and coordination mechanisms for trust-building that reflect the complex, multistakeholder collaborations and values at the core of our project. It provides a snapshot of our efforts to enable privacy and agency in data collection, management, and sharing, and how we practice transparency and accountability to our stakeholders. We are developing collaborative guidelines to codify these values and relationships, and sociotechnical infrastructure to manage how data can and will be shared, when, and with whom. We believe this approach to cultivating trust provides the necessary next steps needed to ensure viability of On-Farm data acquisition from public research institutions and, in fact, will allow us to expand it.

The PSA researchers' body of work spans across a broader set of sustainable agricultural practices (e.g., integrated pest management, crop rotations, soil management, nutrient management, etc.), as well multi-faceted considerations for how to improve sustainability in agriculture (e.g., policy, economics, social context, technology, etc.), though the specific examples in this article focus on the practice that brings us together—cover crops. As such, the lessons that we share translate across "sustainable agricultural research".

The PSA team can also be viewed as a collective of many smaller, interconnected projects. As such, many of the lessons that we share are scale-agnostic. A small team can adopt the open-source approach to developing software to increase transparency (Section 4.3) or conduct stakeholder engagement to improve inclusivity and accountability in the research process (Section 4.2). On the other hand, a small team may not have the need for large-scale technical infrastructure, such as a permissioned database to handle various levels of data sensitivity (Section 4.5) unless there is both sufficient variability in the sensitivity of the data and sufficient desire for data sharing to necessitate the development of such infrastructure. Section 4.4 presents a maturity model that the PSA team uses to govern the development of data acquisition protocols and tools. This maturity model may be viewed as a lens with which to evaluate when to adopt or develop technical infrastructure for smaller-scale or alternatively structured sustainable agricultural research.

2 | THE PSA NETWORK AND OUR DATA STAKEHOLDERS

We formed the PSA team to increase crop productivity, profitability, and resilience; conserve soil and water resources; and reduce the agroecological effect of farming on the environment through increased adoption and effective use of cover crops. We aim to accomplish this by providing the scientific knowledge, technical infrastructure, and outreach necessary to optimize cover crop management and accelerate adoption nationwide, thus aiding the transition to more sustainable agricultural systems. Our network currently consists of 37 core organizations across 29 states, and includes farmers and farmer advocacy groups, government agencies, universities, non-profits, and agricultural and technology companies, as well as a dynamic set of partner organizations focused on targeted research and development initiatives. Our network is consistently evolving, with the inclusion of new farmers across the United States to improve the inference domain of our science and shifting partnerships with public and private enterprise; thus, our approach to expanding the sustainability of agriculture depends on our ability to manage collaboration and trust across an ever-changing network of stakeholders. These stakeholders use the data, models, decision support tools, and cyber physical systems that come from efforts on PSA research stations and partnering farms, but we further aim to make our scientific agricultural data publicly available, allowing other researchers and developers to build on our models and decision support tools and deepen our knowledge of the role of cover crops in crop production.

Drawing on de Beer's (2016) five archetypal categories of open data actors and their interactions with open data we exemplify how our team and the technologies we develop affect the types of trust relationships we intend to cultivate:

Our primary data suppliers are PSA researchers and farmers. While data collected at research stations are uncontroversial in terms of privacy, the data we collect on partner farmer sites include personal information and field data that require varying levels of privacy protections. The PSA researchers are aggregators who collate external open and proprietary data for use in our modeling and analysis. This includes the development of public utility technology, such as our Weather Data Service that aggregates weather data from public government data sources and proprietary weather data (from a private company) for use within our network. Developers fulfill a critical component of the PSA mission—from designing tools and technologies for data acquisition and management (e.g., the Weather Data Service described above) to modeling and building decision support tools (e.g., a nitrogen availability calculator). Each development team uses and transforms different PSA and external data and subsequently provides either cyber physical systems (i.e., integrated hardware and software systems, like a water-sensing system) or software (e.g., data analysis and decision support tools) for public use. Importantly, each of these technologies must be built to protect our data suppliers and aggregators, while still enabling public and private stakeholders to engage with our tools. Enrichers foster collaborations with public and private organizations alike, allowing us to expand our networks and accelerate filling knowledge gaps on climate, soil, and management interactions. Our public–private partnerships support the deployment of our data acquisition systems onto more farmers' fields, and data sharing enables the calibration of our respective models and decision tools. Public–private ventures require care to ensure that public research goals and farmers' privacy and security needs are met while still allowing commercial ventures to build upon open agricultural data. It is both a tenet of U.S. government-funded agricultural research and a shared value of PSA members that data generated from this project be released into the public domain. To facilitate open access to data, enablers such as the USDA National Agricultural Library and the Agricultural Research Service provide systems—like Ag Data Commons (USDA NAL, 2021) and the Agricultural Collaborative Research Outcomes System (USDA ARS, 2017) respectively—to host research products such as data and software. We intend to archive privacy-preserved data with such public databases, as well as make our software source code available via web-hosted code repositories via, for example, GitHub (GitHub, 2021).

In addition, our network also relies heavily on intermediaries that connect researchers and farmers, including Land Grant Institution Extension programs and outreach organizations such as the regional Cover Crop Councils, which transform research into outreach materials (e.g., fact sheets, bulletins, and tools) and events targeting farmers, regulatory decision-makers, and industry. PSA Social Science,

Core Ideas

- Sustainable agriculture is information intensive, requiring data sharing among many actors.
- Trust can be mediated through technology, but only with the input of all data stakeholders.
- Practicing privacy preserving strategies and increasing stakeholder agency increase trust.
- Transparency and accountability can aid in fostering trust in the process of data management.
- Cultivating trust is critical for success in data-driven and sustainable agriculture.

Education, and Extension Teams are partnering with such outreach groups to develop data-driven, tailored outreach materials to improve the efficacy of our efforts and facilitate the transfer of new knowledge and technologies to students and farmers.

3 | THE PSA TRUST FRAMEWORK

Across science and technology, the rise of "big data" has prompted public concerns related to the collection, use, and security of personal data (Ekbia et al., 2015). This crisis of trust has been exacerbated by data breaches across industries. Such concerns are shared by farmers due to growing public and private efforts to harness farm data to transform agriculture (Slattery et al., 2021). Farmers describe issues with transparency in data licenses, a lack of clarity around data ownership and subsequent data sharing and use, concerns about privacy protections, power imbalances between farmers and data aggregators, and a general concern around who is profiting from farmers' data (Wiseman et al., 2019). A recent report provides critical perspectives on agricultural data shared by farmers in the United States, including a lack of trust in public and private entities receiving their data, and insufficient clarity about the direct benefits of sharing their agricultural data (Slattery et al., 2021).

A first step to addressing these concerns is to recognize them as legitimate concerns to be addressed. In our work, we strive to approach farmers as co-producers of knowledge. Yet we must navigate a fundamental tension: how to protect farmer data while also making these rich data available for research and development—particularly to develop and refine models and technological tools to expand successful cover crop adoption, and ultimately, the sustainability of agriculture. Fortunately, recommendations are available on how to build trust among data stakeholders: from practicing transparency when sharing and using agricultural data and building awareness around best practices regarding data management (Wiseman et al., 2019), to working across sectors to enact both policy and technological safeguards around farmer data and creating a culture, which prioritizes providing value to the farmer (Slattery et al., 2021). Jakku et al. (2019) further advocate working within farmers' already trusted advisors and networks. However, they do not provide guidance on how to operationalize trust via social and technical infrastructure.

Within PSA, our approach heeds the concerns and perspectives shared by farmers, researchers, and agricultural data stakeholders, takes guidance from recommendations made in both agricultural and technological communities, engages cross-sector partnerships, and constructs a practical social and technical infrastructure to support collaboration and trust.

In the context of agricultural data, we focus on cultivating trust through (a) privacy, including mechanisms for protecting data stakeholder personal information, handling geolocation information, and providing granular data access controls; and (b) agency, including negotiation of data ownership, situational control over data, informed consent for data sharing, and inclusion of data stakeholder voices. In the context of our research and development process, we cultivate trust through (a) transparency, including stakeholder inclusion, feedback tools, open-source development of tools, and open access to privacy-preserving data; and (b) accountability, including the responsibility to protect data and the rights of research participants, the traceability of our actions, and mechanisms for feedback across the network.

3.1 | Trusted with data

Agricultural data are diverse: crop growth and yield metrics; animal breed and management data; data regarding management of fertilizers, pesticides, and other inputs; location-specific characteristics of soils, climate, and water; and financial, labor, and machinery data. There is a growing tension around farm data privacy, as a range of actors–researchers, technology makers, input suppliers, and farmers themselves–begin to amass big data on farms (Rotz et al., 2019). Skyuta (2016) succinctly describes the paradox of farm data: though it is conceptually similar to commercial data, it is treated as personal data by farmers. This likely stems from concerns around how such data are going to be used, by whom, and to what consequence for farmers' livelihoods (Slattery et al., 2021; Sykuta, 2016; Wiseman, 2019). As a result, we consider preserving privacy essential, not only to protect farmers' sensitive information, but to build trust and encourage continued collaborations.

Despite our commitment to data privacy, the intangibility of data obscures notions of data ownership, as well as what it means to use or share data. It is difficult to ascribe ownership discretely. For example, when researchers measure soil moisture on a farm, who owns the data? The landowner, the field technician, the lead scientist? Agricultural data are often considered to be owned by the farmer, though this is complicated when, for instance, technology providers claim ownership of machine-collected data (e.g., yield monitors), or when data are collected on a farm via a research collaboration. Each instance of data ownership is rife with unique challenges. This is further complicated because unlike many other resources, data are not consumed upon use. This means that data can be used and reused limitlessly, making traceability of access and use a complex problem (see FAO, 2018). Ellixson and Griffin (2017) offer that the collaborative development of data sharing agreements between data creators and data users can ensure that the ownership rights for data creators are respected and the expectations of data users are clearly laid out. They broadly suggest that each actor with access to farm data consider their rights and responsibilities prior to sharing data, and more specifically offer recommendations for farmers to protect their data via a range of legal frameworks (Nondisclosure agreements, contract language, etc.) (Ellixon & Griffin, 2017).

Guidance for research data management, including agricultural data, are still evolving. Research data protections for identifiable, personal, farmer data typically falls under the purview of the US federal regulations and are subsequently monitored by a research organization's Institutional Review Board (IRB). However, as we discuss in Section 4.5, the boundary between agricultural research data collected on research partners' farms and private data is not always distinct and it is left to the researchers and their research participants to determine what constitutes acceptable use and responsible conduct.

Given the characteristics of agricultural data, the range of data protections, it is vital that data stakeholders, from creators to users, develop a shared understanding around ownership, expectations of privacy, and expected use, prior to data collection or access. Rotz et al. (2019), describe the need for a shift in power to data creators, providing them with agency over their data.

3.2 | Trusted in process

Despite recent high-profile exceptions, public trust in science has been consistently high over the past 40 years and increased between 2016 and 2019 (Krause et al., 2019; Pew Research Center, 2019). The Pew report identified two key factors that engender greater public trust in scientific findings: open availability of data to the public and review of scientific data by independent committees.

A transdisciplinary scientific team focused on agroecological objectives such as ours must foster trust and accountability (a) among researchers, (b) between researchers and our partnering farmers, and (c) between researchers and external stakeholders, including the public. We invite stakeholders, including farmers, to participate in the project governance via an advisory committee and provide longer term feedback on research design via farmer think tanks (see Section 4.2), as a first step toward promoting accountability between our researchers and partnering farmers. In Section 4.3, we describe the open-source approach to software development, as an example of how we practice transparency and accountability in the process of developing technology among all three groups: researchers, partnering farmers, and external stakeholders.

We take a process-oriented approach to trust that prioritizes communication among stakeholders, embraces disagreement as an opportunity to expand transdisciplinary understanding, and recognizes the need for collaborations to evolve over time and in response to feedback (Bennett et al., 2018). Rashid (2015) characterizes mutual accountability in teams as a continual process in which team members provide each other with feedback to support the achievement of common goals. Iversen et al. (2020) additionally highlight a culture of safety, inclusion, and trust, and consistent stakeholder engagement as the foundation of a scientific team's success.

3.3 | Operationalizing trust

Enacting our trust framework requires concrete mechanisms to mediate our complex network of actors, values, and objectives and their associated data, models, tools, and other information products. A combination of social and technical infrastructure is used to govern data sharing among internal PSA members and external stakeholders (Figure 1).

4 | TRANSLATING TRUST INTO A TECHNICAL INFRASTRUCTURE

The PSA trust framework, which emphasizes our commitment to privacy, agency, accountability, and transparency, reflects our project objectives and team values. As a transdisciplinary project, our team members—from farmers and industry partners to researchers ranging from agronomists and economists to information scientists and anthropologists—come to this project with different areas of expertise and perspectives on the problems at hand. This is an invaluable bricolage of experience, yet it requires a series of social structures to support the ongoing work of communication, negotiation, and collective decision-making.

In this section, we detail the practicalities of how we transform our dynamic, heterogeneous, "living" social infrastructure of trust and values into a technical infrastructure that is robust, flexible, and furthers the trust we have built within our network, among our partners, and with our stakeholders. Technology does not inherently ensure trust among its users, but it can mediate trust among collaborators if it is deliberately constructed to reflect the ethos of the environment it serves (Bodó, 2020). Consider a database in the PSA network, where one's access is granted based on a set of rules of engagement. An ill-defined understanding of data use and constraints, combined with a lack of transparency and consensus during the early phase of development, could lead to unnecessarily limited data accessibility—or worst case, data exposure resulting in exploitation and misuse. If technology is to effectively mediate trust between collaborating actors, a participatory approach must be employed where data stakeholders can voice their needs, goals, concerns, and stances on privacy, data ownership, and ethics that drive data-driven agricultural research.

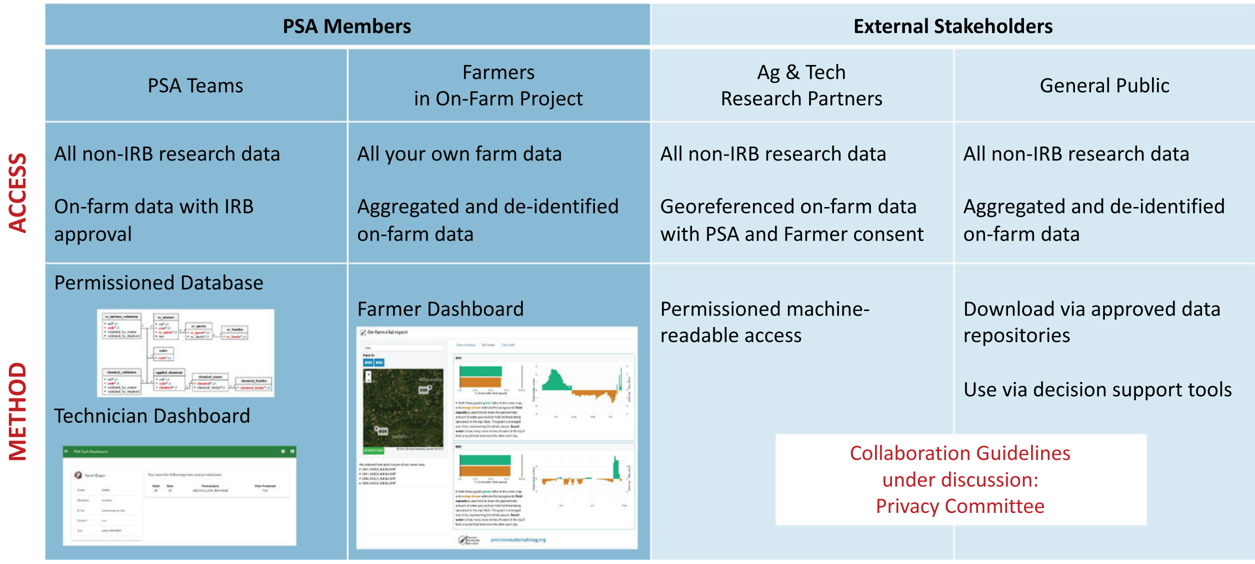

FIGURE 1 This figure summarizes internal and external data sharing processes and tools. Precision Sustainable Agriculture (PSA) research teams have access to research data via a permissioned database and technician dashboard that reflects their Institutional Review Board (IRB) approval status and farmer permissions. Research data is shared back to farmers via a farmer dashboard. De-identified data will be shared with external stakeholders via three mechanisms: a permissioned readable database for research partners, download via data repositories, and through PSA-development of decision support tools

The PSA research and development activities take the form of modular, interconnected projects that vary in maturity and technology use. At research stations, PSA scientists conduct common experiments, that is, replicated experiments to examine the effect of cover crops on nitrogen and water dynamics, pests, and cash crop performance. In the On-Farm Project we collect agronomic, environmental, economic, personal, and farm management data to inform cover crop performance across varying landscapes. A high-level overview of the PSA data lifecycle is shown in Figure 2. Data are collected at university and government research sites and from our network of farmers who partner in the On-Farm Project. The PSA data are gathered using a variety of techniques: from destructive measures of plants and soil, to cyber physical systems deployed to measure a diverse array of crop and soil dynamics at the micro (i.e., point) to macro (i.e., geospatial) scale, to surveying and interviewing farmers. These primary data are supplemented with a range of external data from partner organizations and public sources. Given the range of methods, each experiment or project requires different approaches to mediating trust with data, models, and tools. It is the responsibility of our PSA Technology Teams to (a) inform our data management guidelines to ensure technical feasibility, as well as (b) adhere to and enforce these guidelines in technical implementations. Different types of data are being handled using data-appropriate management practices that result in, for instance, permissioned databases and file storage, as well as a suite of data exploration tools. Data are used in the development of both models and decision support tools that will be subsequently released to the public via extension and education efforts to ultimately increase uptake and effective management of cover crops in agricultural systems.

4.1 | Project governance

Effective project governance provides a stable set of organizational structures, processes, and decision-making frameworks to support a team moving forward in achieving its objectives. The PSA is structured around distributed governance across a series of project-based teams and issue-specific subcommittees, which meet on a weekly, bi-weekly, or monthly basis. These transdisciplinary groups (including an Executive Leadership Team composed of project-based team leads) dedicate time to discussion, consensus-building, and soliciting input from the full PSA team (and at times, external stakeholders). This structure facilitates a collaborative approach to developing research protocols, informs the design of data acquisition and data management tools, and is ultimately embodied in the products and technologies we develop.

Our publication guidelines, which were developed via subcommittee drafting and full PSA team engagement, are an exemplar of this approach to project governance. They build on the guidelines developed by the Sustainable Corn Project (Abendroth et al., 2017), and were adapted to frame our shared expectations regarding the basic "rules" for publishing data. For example, we outline responsible timelines for publication and internal data sharing to allow teams collecting primary data the opportunity to publish their individual site data or domain-specific results while also ensuring that teams

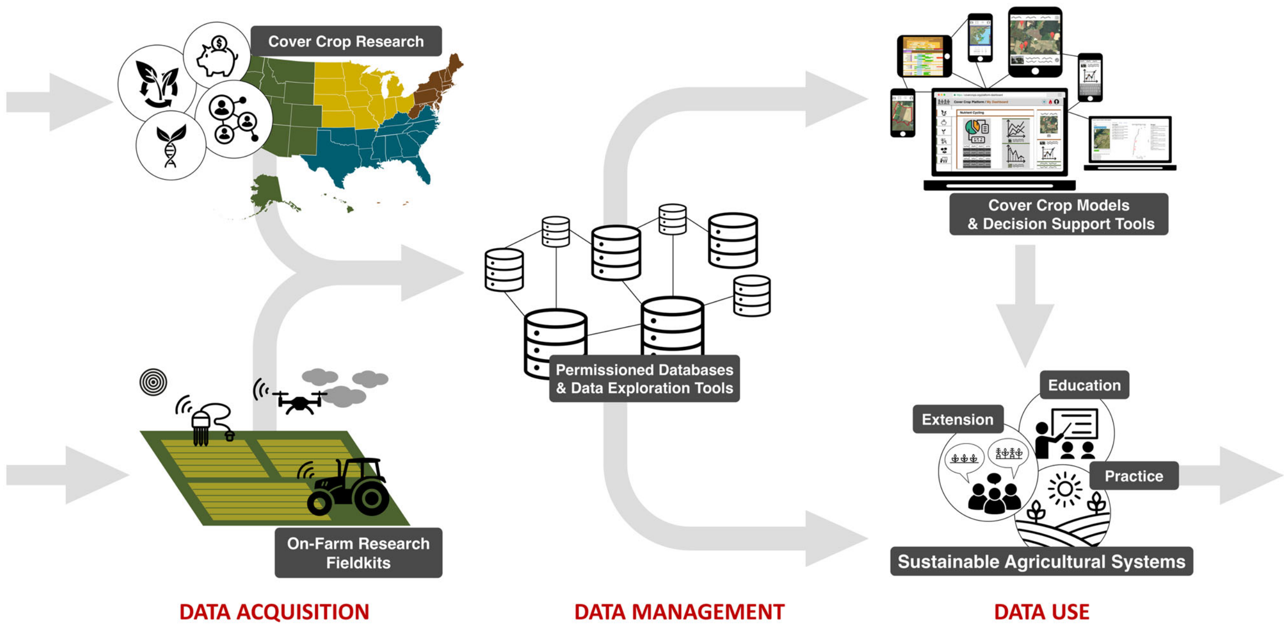

FIGURE 2 This figure provides a high-level overview of the Precision Sustainable Agriculture data lifecycle, as we produce research data as well as acquire data on farms; provide access to structured data via permissioned databases, and both use and make available our research data for use in, technology development and sustainable agriculture extension, education, and practice

aggregating data across research sites or domains (e.g., modeling and economics teams) have timely access to those data.

4.2 | Stakeholder engagement

A stakeholder-oriented approach to project governance has been shown to correlate with project success (Joslin & Müller, 2016). In collaborative research projects involving a diverse set of stakeholders, such as PSA, it is important that trust is developed and that all cooperators have agency to determine how their data are being shared, and with whom. The PSA researchers want (a) attribution for the research data they create and (b) that the data is being used responsibly by those they share it with. Farmers want assurances that their privacy is considered, particularly where it concerns financial data, or data with regulatory implications (e.g., water quality data) (Slattery, 2021). This requires both transparency and accountability about how the data are managed and shared, as well as which data are kept confidential vs. shared more widely.

Interactive information dashboards aim to close the communication loop between real-world conditions and participants in an information ecosystem. We have currently developed two approaches relevant to data suppliers: a technician dashboard supports research technicians to diagnose data flow and data quality issues related to data collection, and a farmer dashboard allows us to share the field data we collect back to farmers. Both are intended to be near real-time tools, to shorten the existing extension and research data lifecycle. The technician dashboard aggregates all PSA activities conducted in the On-Farm Project, with a unique view for each user. Technicians see all data only from sites they manage (including contact information, field histories, and scientific data like soil water content); in contrast, domain-specific "data shepherds" only see domain-relevant data (such as soil water content) across all sites. This two-tier user approach allows us to troubleshoot data quality issues quickly, within hours to weeks, as opposed to traditional research methods that may only vet data once per growing season or at the end of a multi-year experiment. The credentialed login for this tool reflects our social expectations of privacy for our partners and builds in accountability through an infrastructure that is transparent to the entire network. In turn, these expectations are enforced in the underlying technical infrastructure, that is, the database (detailed in Section 4.5).

To complement the technician dashboard, the farmer dashboard provides our farmers with summarized data collected at their farm, such as cover crop biomass, soil water budgets, or crop yield. Notably, farmer feedback has informed the design of this dashboard – from the first iteration, which simply displayed visualizations and tables of raw data, to the current version that summarizes the complex data we collect and contextualizes these data in each state by comparing them to de-identified data from other nearby participants. Thus, this tool provides near-real time data to our farmer-collaborators and creates a channel of accountability to reinforce our trust relationships with these partners. Further, through on-going dialogue with our field technicians, this tool has created an opportunity for farmers to provide informal feedback on our technology infrastructure, and insight into their comfort with data sharing as the experimental network expands.

To obtain longer-term feedback on our research efforts, our social science team is conducting farmer think tanks. At these events, farmers from each state participating in the On-Farm Project will contribute their individual impressions about the research in facilitated group discussions twice per year. These events provide strategic, mid- to long-term stakeholder participation aimed at guiding design, and promoting ownership within the research process. It also demonstrates that the PSA team respects and values farmer knowledge and expertise and provides a critical mechanism for research translation.

4.3 | The open-source way of development

The open-source development paradigm (Hauge et al., 2010) is characterized by distributed, collaborative development of software whose source code is publicly shared. This approach is attractive to our network for two reasons: we increase our efficiency by using shared community resources, and we enforce expectations that our work in turn enriches the commons. Whereas closed-source tools have a one-way relationship between developer and user, we benefit from the building and documentation of tools we use (upstream participation), and adding value to those tools through user testing, bug reports, and code contributions (downstream participation). Much of the code we write to implement aspects of the PSA infrastructure may not be directly reusable by others, but by working "in the open" on tools via platforms like GitHub, we operationalize norms around transparency and collaboration.

This open-source approach has led to the creation of tools that reach beyond their original conception in the PSA network. For example, a cyber physical system developed for soil moisture sensing has been a joint effort by founding PSA members (with development preceding the formation of PSA itself) and a private company and will be now offered to customers of the private company. Although PSA farmer data is not being shared with this company heir product development builds on our collection of farmers' field data, while we identify bugs in their hardware and software.

These social norms of openness around our software development process and tool usage extend to our group values of transparency in other domains of documentation: how we use data, how we provide services, and how we track iterations of each experiment through version control. An open-source ethos is nevertheless grounded in a spectrum of behaviors, and each aspect of the network engages this in different ways. Some tools, such as the source code for the technician dashboard, are fully public, with code contributions from each user traceable line-by-line. Other tools, such as experimental protocols, are written collaboratively, but only published in major releases by the project manager within our network. Our long-term vision includes sharing these protocols publicly for use by non-PSA researchers and others. They signal our commitment to maintaining transparency and accountability to build trust.

4.4 | A maturity model to govern data acquisition

As we have described above, the PSA project has two types of data acquisition sites: research stations, where we have frequent access and greater control over the fields, and farms, where we have periodic access and minimal control over field management. Given this dichotomy, we have developed a maturity model (Figure 3) to guide how PSA protocols and technologies are rolled out across sites.

Early-stage research protocols and data acquisition prototypes (alpha tech) are developed and tested on research stations to ensure that critical data management processes and tools are in place before we collect data on farms. The PSA Technology Team develops cyber physical and software tools for each in-field data acquisition protocol, also beginning with prototyping and testing at a limited number of research sites.

Younger protocols, more prone to change, use web-based spreadsheets that enable collaboration, flexibility, and capacity for rapid protocol refinement, with minimal technical know-how. A member of the PSA Technology Team works with the research team to glean user requirements, identify appropriate technologies, and as necessary, conceptualize potential alpha technologies. As protocols and the associated tools mature, they are beta tested on a small number of farms. From there, a protocol is then either used by a larger portion of our PSA scientists or, as appropriate, rolled out to the On-Farm Project.

Inconsistent performance of technology and data collection systems can rapidly cause fatigue in inexperienced users. Only the more robust, field-tested data acquisition technologies that have the highest potential for trustworthiness are deployed in our farmers' fields. For instance, a team at North Carolina State University has been developing a "StressCam" (Ramos-Giraldo et al., 2020), a low-cost, open-source, cyber physical system consisting of a camera and machine-learning algorithms trained to utilize leaf imagery to detect drought stress in a field crop. In 2020, it was categorized as an alpha technology, with a focus on developing a robust prototype, collection of training data for machine learning. In 2021, 40 StressCam units were distributed to eight states for beta testing. Users were asked to evaluate pairs of StressCams in on-farm settings to test field installation to the user experience with the technician dashboard, with developers learning and resolve issues such as device communication challenges and

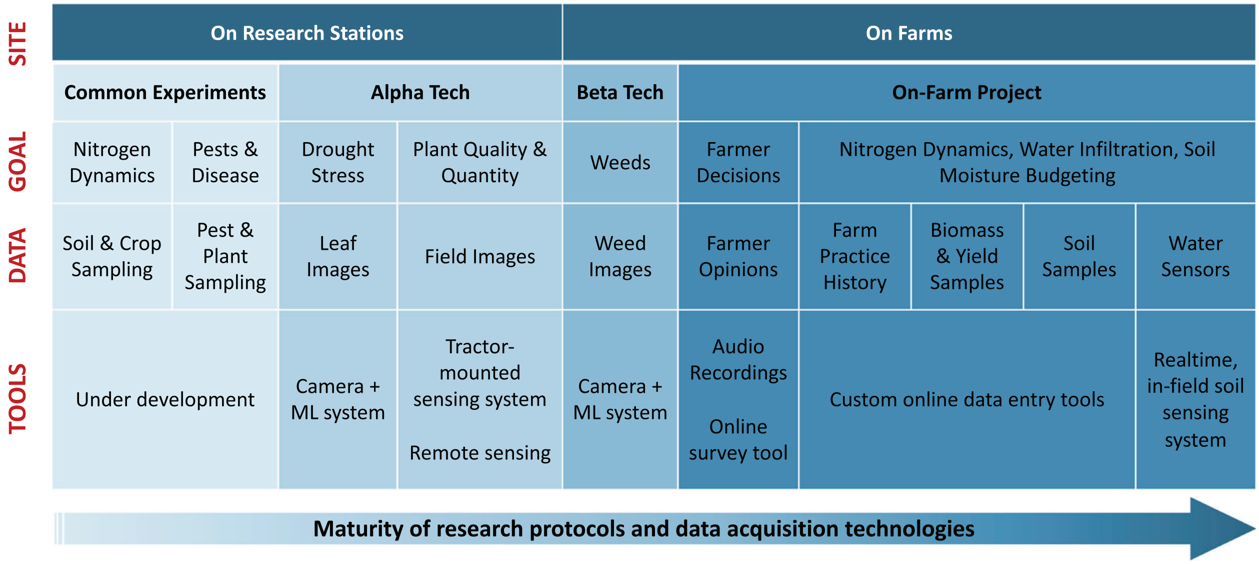

FIGURE 3 This figure overviews the Precision Sustainable Agriculture maturity model with example protocols and technologies. Each vertical is an example of a where the data is collected (either on a research station or on a partnering farm), what the research goal is, the associated data collected, and the tools used. Maturity of research protocols and data acquisition technologies increases from left to right, with examples of our most mature tool, the water sensing system in the right most column

time zone glitches. We expect to scale up to approximately 80 cameras in the next year and move to a professional manufacturing to ensure device reliability.

4.5 | Handling data with varying sensitivity

Each of the PSA projects pose distinct challenges for data management, privacy and ownership, rights and responsibilities, and the ethical conduct of data-intensive research at scale. Thus, one of the PSA Technology Teams is the “Data Flow Team”, which builds and administers a technical infrastructure that enables standardized data collection, protects privacy through security, and will train cooperators on the best trust-enhancing data management practices during a variety of PSA meetings. The broader set of PSA Technology Teams are working toward software services that provide agricultural datasets for use by both our internal and external partners. As our team gains experience as confidence in information processing, we are refining our data flow systems to avoid loss of data, models, and protocols. For example, in the On-Farm Project we have moved from shared web-based spreadsheets that were accessible to all data collectors (and thus easily corrupted), to databases with role-based permissions for users and groups, logical structuring and separation of data, as well as controlled access to machine-readable data.

The level of security required to preserve privacy depends on data sensitivity, with stricter measures required for more sensitive data. We have established a preliminary schema to classify data according to its level of sensitivity (Figure 4), and our social science team is working with farmers to determine their perspectives on data sharing with different audiences.

Data collected in the On-Farm Project are either sensitive (e.g., GPS coordinates, personal information) or not sensitive (e.g., field observations, quantitative and qualitative field measurements, sensor data). Both the sensitive and nonsensitive On-Farm data are stored in one password protected location, and PSA team members have permissioned access to On-Farm data, as we describe above. The nonsensitive data collected across the PSA On-Farm network will be publicly available once we have established a protocol for data release. In contrast, human-subjects’ data, such as those collected by our social scientists and economists, are considered very sensitive. These data, subject to IRB regulations, are isolated from On-Farm field data and other personally identifiable information and are only available to researchers listed on approved IRB protocols. This modified data management strategy provides data administrators, data collectors (e.g., technicians), and data users (e.g., modelers) limited access privileges according to each groups’ needs, and results in improved data security, quality, and uniformity.

4.6 | Managing and anticipating data use potentialities

The digital agricultural revolution involves accelerated adoption of technologies such as low-cost sensors and data acquisition platforms that result in a growing availability of rich

	On–farm		Human-Subject	
DATA	• Field observations • Quantitative and qualitative measurements • Sensor data	• Personal Information • GPS locations	Anonymized, aggregated data	Personally Identifiable Information
IRB	Not subjected	Not subjected	Subjected	Subjected
USERS	• Data administrators (all data) • Modelers (all data) • On-farm team (all data, by institution)	• Data administrators (all data) • Modelers (only GPS locations) • On-farm team (all data, l by institution)	• Social Scientists • Economists	Researchers named on the IRB
ACCESS	• Data administrators: read, write, edit • Modelers: read, edit (limited) • On-farm team: read, write, edit (limited)	• Data administrators: read, write, edit • Modelers: read • On-farm team: read, write, edit	• Social Scientists: read, write, edit (limited) • Economists: read, write, edit (limited)	Researchers named on the IRB: read, write, edit
STORAGE	Database 1	Database 1	Database 2	Database 3
	Not Sensitive	Sensitive	Very Sensitive	Very Sensitive

FIGURE 4 This figure overviews data sensitivity levels for data collected in the Precision Sustainable Agriculture On-Farm Project. Sensitivity of the data increases from left to right. Each column provides an example of the types of data collected, whether they are subject to Institutional Review Board (IRB) regulations, who the primary users of the data are, how these data are accessed, and how the data storage is split up

agricultural data. Farmers increasingly look to information management and decision support tools, driven by agricultural data and machine learning to glean novel insights into productivity, profitability, and sustainability of their farms. At the same time, agricultural research increasingly depends on "big data" analyses that draw on a multitude of agronomic, genetic, environmental, and socio-economic data. For instance, questions about the effect of climate change on cropping systems, and the validation of technology for climate change adaptation, require researchers to process and synthesize large amounts of weather and agronomic data to make sense of the complex system dynamics and develop feasible recommendations to these urgent problems.

We are using PSA data to refine, redesign, and develop cover crop models and decision support tools to enable farmers and researchers to manage cover crops to meet their agronomic and ecological goals. These value-added products may use farmer data directly (e.g., via incorporation into water infiltration models) or indirectly (e.g., refinement of soil water sensing systems). Tools such as the farmer dashboard provides our farmers with short-term insights about their fields (e.g., soil moisture dynamics), while web-based applications built using these data and models will provide valuable decision support for agricultural practitioners (e.g., cover crop selection tool, nitrogen availability calculator).

Yet our reliance on farmer-generated data poses two critical challenges going forward: "How do we structure data protection in the present while anticipating emergent opportunities for data use in the future?" and "How do we avoid algorithmic bias as we codify decisions in technologies?" Despite the clear opportunities that On-Farm data present for ongoing research and development, it is vital that farmers are informed about the potential use of data collected by agricultural technologies. While farmers may consent to the use of their data in one context, they choose to withdraw from sharing data in other contexts. Consider the previously discussed example of the "StressCam" (Ramos-Giraldo et al., 2020). Imagery data are obtained by the camera, images, temperature, and light, processed locally via embedded machine-learning algorithms in the device, and simultaneously sent to the cloud through cellular connection. While the StressCam is in the *alpha* and *beta tech* level of maturity, all data is retained for algorithms research and development. Ultimately, StressCam data will only be collected for verification of operation and improvement of the algorithms if the end-user (e.g., farmer) authorizes the sharing of raw and processed data. Such systems present significant challenges related to communicating changing data sharing and use within software. Nevertheless, these tools offer an opportunity to gather big data with minimal labor and provide actionable, real-time insights for farmers.

5 | CULTIVATING TRUST: A WORK IN PROGRESS

Our goal is to use the power of the sheer geographic and numerical scope of our On-Farm and on-station research to synthesize site-specific recommendations and create the precision management technologies and recommendation tools that farmers need to effectively implement sustainable agricultural practices, like cover cropping. Our ability to inform research and build useful models relies on big data collected under the real-world conditions found on farms. Inherent in our partnership with farmers is the need for farmers to trust both researchers and our external collaborators, and the technologies we use. This sets the scene for a complex trust landscape with heterogeneous needs for data ownership, privacy, access, and sharing constraints. The PSA trust framework has two pillars: first, we aim to be trusted with data, through implementations of privacy preserving technologies as well as making sure that our tools and processes support data stakeholder agency. Second, we aim to practice transparency and accountability through our project governance, development methods, and data management practices. By implementing our trust framework, we anticipate securing and supporting a rich and broad community of cooperators and stakeholders. As a result, we aim to enhance cover crop adoption, integrating precision and sustainable agriculture, to enable more productive, profitable, and resilient agricultural systems.

ACKNOWLEDGMENTS

This research is supported by the National Institute of Food and Agriculture, U.S. Department of Agriculture, under award no. 2019-68012-29818, "Enhancing the Sustainability of US Cropping Systems through Cover Crops and an Innovative Information and Technology Network".

AUTHOR CONTRIBUTIONS

Ankita Raturi: Conceptualization; Data curation; Funding acquisition; Investigation; Methodology; Project administration; Resources; Software; Supervision; Validation; Visualization; Writing-original draft; Writing-review & editing. Jennifer J. Thompson: Conceptualization; Data curation; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Validation; Writing-original draft; Writing-review & editing. Carlene A. Chase: Conceptualization; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Validation; Writing-original draft; Writing-review & editing. Robert Myers: Conceptualization; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Validation; Writing-review & editing. Aurelie Poncet: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Software; Validation; Visualization; Writing-original draft; Writing-review & editing. Paula Ramos: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Software; Validation; Writing-review & editing. Chris Reberg Horton: Conceptualization; Data curation; Funding acquisition; Investigation; Methodology; Project administration; Resources; Software; Supervision; Validation; Writing-review & editing. Roderick Rejesus: Conceptualization; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Validation; Writing-review & editing. Alison Robertson: Conceptualization; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Validation; Writing-review & editing. Sarah Seehaver-Eagen: Conceptualization; Data curation; Investigation; Methodology; Project administration; Software; Supervision; Validation; Writing-review & editing. Steven Mirsky: Conceptualization; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Validation; Writing-original draft; Writing-review & editing.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Ankita Raturi https://orcid.org/0000-0003-0637-8541
Alison Robertson https://orcid.org/0000-0003-4281-6350
Matthew D. Ruark https://orcid.org/0000-0002-8678-6133

REFERENCES

Abendroth, L., Herzmann, D., Lokhande, S., Chighladze, G., Schwaegler, K., & Morton, L. W. (2017). *Project and research management: Integrating systems, data, and people in multidisciplinary work (Vol. 5)* (Publication No. CSCAP-0198-2017; Observations and recommendations of the USDA-NIFA funded Climate and Corn-Based Cropping Systems Coordinated Agricultural Project.) https://dr.lib.iastate.edu/entities/publication/25e5564e-eb2f-4f78-8a5e-55c6339af0e3

Bennett, L. M., Gadlin, H., & Marchand, C. (2018). *Collaboration and team science field guide*. National Cancer Institute.

Bodó, B. (2020). Mediated trust: A theoretical framework to address the trustworthiness of technological trust mediators. *New Media & Society*, *23*, 2668–2690. https://doi.org/10.1177/1461444820939922

Clark, A. (Ed.). (2008). *Managing cover crops profitably*. Diane Publishing.

de Beer, J. (2016). *Ownership of open data: Governance options for agriculture and nutrition*. Global Open Data for Agricultural and Nutrition.

Delgado, J. A., Dillon, M. A., Sparks, R. T., & Essah, S. Y. C. (2007). A decade of advances in cover crops. *Journal of Soil and Water Conservation*, *62*(5), 110A–117A.

Ekbia, H., Mattioli, M., Kouper, I., Arave, G., Ghazinejad, A., Bowman, T., Suri, V. R., Tsou, A., Weingart, S., & Sugimoto, C. R. (2015). Big data, bigger dilemmas: A critical review. *Journal of the Association for Information Science and Technology*, *66*(8), 1523–1545. https://doi.org/10.1002/asi.23294

Ellixson, A., & Griffin, T. (2017). *Farm data: Ownership and protections*. University of Maryland Extension.

FAO. (2018). *EU code of conduct on agricultural data sharing by contractual agreement*. FAO.

Github. (2021). *Github*. https://github.com

Hauge, Ø., Ayala, C., & Conradi, R. (2010). Adoption of open-source software in software-intensive organizations – A systematic literature review. *Information and Software Technology*, *52*(11), 1133–1154. https://doi.org/10.1016/j.infsof.2010.05.008

Iversen, C. M., Bolton, W. R., Rogers, A., Wilson, C. J., & Wullschleger, S. D. (2020). Building a culture of safety and trust in team science. *Eos*, 101. https://eos.org/opinions/building-a-culture-of-safety-and-trust-in-team-science

Jakku, E., Taylor, B., Fleming, A., Mason, C., Fielke, S., Sounness, C., & Thorburn, P. (2019). If they don't tell us what they do with it, why would we trust them?" Trust, transparency and benefit-sharing in Smart Farming. *NJAS - Wageningen Journal of Life Sciences*, *90–91*, 100285. https://doi.org/10.1016/j.njas.2018.11.002

Joslin, R., & Müller, R. (2016). The relationship between project governance and project success. *International Journal of Project Management*, *34*(4), 613–626. https://doi.org/10.1016/j.ijproman.2016.01.008

Kaye, J. P., & Quemada, M. (2017). Using cover crops to mitigate and adapt to climate change. A review. *Agronomy for Sustainable Development*, *37*(1), article no. 4. https://doi.org/10.1007/s13593-016-0410-x

Krause, N. M., Brossard, D., Scheufele, D. A., Xenos, M. A., & Franke, K. (2019). Trends—Americans' trust in science and scientists. *Public Opinion Quarterly*, *83*(4), 817–836.

Lewis, J. D., & Weigert, A. (1985). Trust as a social reality. *Social Forces*, *63*(4), 967–985. https://doi.org/10.2307/2578601

Pew Research Center. (2019). *Trust and mistrust in Americans' views of scientific experts*. Pew Research Center. https://www.pewresearch.org/science/2019/08/02/trust-and-mistrust-in-americans-views-of-scientific-experts/

Ramos-Giraldo, P., Reberg-Horton, C., Locke, A. M., Mirsky, S., & Lobaton, E. (2020). Drought stress detection using low-cost computer vision systems and machine learning techniques. *IT Professional*, *22*(3), 27–29. https://doi.org/10.1109/MITP.2020.2986103

Rashid, F. (2015). *Mutual accountability and its influence on team performance*. Harvard University. https://dash.harvard.edu/bitstream/handle/1/14226095/RASHID-DISSERTATION-2015.pdf?sequence=4

Rotz, S., Duncan, E., Small, M., Botschner, J., Dara, R., Mosby, I., Reed, M., & Fraser, E. D. G. (2019). The politics of digital agricultural technologies: A preliminary review. *Sociologia Ruralis*, *59*(2), 203–229. https://doi.org/10.1111/soru.12233

Slattery, D., Rayburn, K., & Slay, C. M. (2021). *Farmer perspectives on data*. Trust in Food, a Farm Journal Initiative & The Sustainability Consortium. https://www.trustinfood.com/wp-content/uploads/2021/05/Farmer-Perspectives-on-Data-2021.pdf

Sykuta, M. E. (2016). Big data in agriculture: Property rights, privacy and competition in ag data services. *International Food and Agribusiness Management Review*, *19*, 57–74.

USDA ARS. (2017). *Agricultural Collaborative Research Outcomes System (AgCROS)*. USDA ARS. https://data.nal.usda.gov/dataset/agricultural-collaborative-research-outcomes-system-agcros

USDA National Agricultural Library (NAL). (2021). *Ag data commons | Providing central access to USDA's open research data*. USDA NAL. https://data.nal.usda.gov/

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

How to cite this article: Raturi, A., Thompson, J. J., Ackroyd, V., Chase, C. A., Davis, B. W., Myers, R., Poncet, A., Ramos-Giraldo, P., Reberg-Horton, C., Rejesus, R., Robertson, A., Ruark, M. D., Seehaver-Eagen, S., & Mirsky, S. (2022). Cultivating trust in technology-mediated sustainable agricultural research. *Agronomy Journal*, 1–12. https://doi.org/10.1002/agj2.20974