Received: 25 May 2021 | Accepted: 24 August 2021 | Published online: 14 November 2021

DOI: 10.1002/agj2.20873

Agronomy Journal

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

Digital agriculture platforms: Driving data-enabled agricultural innovation in a world fraught with privacy and security concerns

Bryan C. Runck[1,4,#] | **Alison Joglekar**[1,3,#] | **Kevin A. T. Silverstein**[1,2] | **Connie Chan-Kang**[1,3] | **Philip G. Pardey**[1,3] | **James C. Wilgenbusch**[1,2]

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

[2] Minnesota Supercomputing Institute, 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] Masters of Geographic Information Science Program, Univ. of Minnesota - Twin Cities, Minneapolis, MN 55455, USA

Correspondence
Bryan C. Runck, GEMS Informatics Center, Univ. of Minnesota - Twin Cities, St. Paul, MN 55108, USA.
Email: runck014@umn.edu

Assigned to Associate Editor David Clay.

[#]Co-first authors

Abstract

Digital agriculture is often heralded as the next major wave of innovation in food and agriculture. Driven by big data, the intention is that digital agriculture will transform the entire research and development pipeline across agricultural value chains throughout the developed and developing worlds. Yet, issues concerning data quality, interoperability, intellectual property ownership and data privacy present considerable challenges to this vision. Digital agriculture platforms, which support data sharing, analysis, interoperability, and public and private sector collaboration, are one approach to address this challenge, but as a new research domain, there is a lack of conceptual clarity around what constitutes a "digital agriculture platform". Here, we use a "bottom up" and "top down" analysis approach to develop a taxonomy of the digital agriculture landscape. Then, we select a set of digital agriculture platforms for in-depth analysis across a set of technical and use requirements. While digital agriculture will remain a constantly evolving landscape with varied technologies and application areas, this presents a first attempt to characterize this landscape and establish a common vocabulary for understanding digital agriculture platforms.

1 | INTRODUCTION

Many posit that digital agriculture has the potential to be the next major wave of innovation in food and agriculture in rich and poor countries alike (see, e.g., Grey et al., 2018; Trendov et al., 2019; Wolfert et al., 2017; World Bank, 2016). Here we define *digital agriculture* as technologies that "digitally collect, store, analyze, and share electronic data and/or information along the agricultural value chain" (e.g., from field to mouth) (Wikipedia 2020). Building on the mechanical, chemical, and biological innovations that emerged prior to the 21st Century, digital innovations may be transformational in addressing the substantial challenges facing agriculture, from declines in productivity performance to pressing environmental challenges (Pardey & Alston, 2021). Unlocking the big data promise for food and agriculture requires unlocking data throughout the entire research and development (R&D) pipeline; from pre-commercial to commercial and from molecule to market.

Digital agriculture platforms constitute an enabling environment for cross-sector, data-driven agricultural innovation across the entire agricultural and food technology innovation

Abbreviations: R&D, research and development; FAIR-ER, findable, accessible, interoperable, resusable, ethical, reproducible; GEMS, genetic, environmental, management, socio-economic

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2021 The Authors. *Agronomy Journal* published by Wiley Periodicals LLC on behalf of American Society of Agronomy.

14350645, 2022, 5, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/agj2.20873, Wiley Online Library on [28/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

landscape (Gustafson et al., 2017). Here we define digital platforms as a group of technologies that are used as a base upon which other applications, processes or technologies are developed (Techopedia, 2020). Ideally, such platforms allow researchers in the public and private sectors to store, clean, share, and analyze diverse types of functionally interoperable data ranging from genetics and environmental to management and socioeconomics. Management aspects are pervasive at all spatial scales, be that experimental plots, farmers' fields, or entire farming operations, to landscapes and beyond. Often explicitly spatial and temporal, such data can be difficult for scientists and others to manage, and thus the goal of digital platforms is to improve the ease and efficiency of data management throughout the entire data use and re-use life cycle.

The growing deluge of agriculturally-relevant data along with new computational capabilities to make sense of these data, coupled with increasing concerns about data privacy and security (see, e.g., Wilgenbusch et al., 2021), has spurred the development of a number of new digital agriculture platforms. These platforms are being created for a wide range of users and uses (agribusiness to individual farmers to store, share, and process data). While the rapid emergence of a diverse set of platforms is promising, it also introduces a new set of obstacles. Because the field of digital agriculture is young and highly interdisciplinary, the language used to describe these platforms has failed to keep pace with their development, and the notion of what constitutes a "digital agriculture platform" is currently unsettled.

Few common terms or standards exist to both characterize what role specific platforms play in the larger digital agriculture ecosystem. The term *agricultural platform* currently has a range of definitions that vary in both functionality and content. Examples include: a database of farm-level data; a website where information about digital tools is aggregated; data repositories; and a digital tool suite that links data to tools and thus supports analytics. Furthermore, the type, quantity, quality, and usability of data on agricultural platforms varies. The data in some platforms are entirely open, some are proprietary fee-for-service portals, and others seek to straddle both open and closed access. The lack of a clear vocabulary and a conceptual schema makes it difficult to describe and develop a shared understanding of the lay of this changing digital landscape. A goal of this paper is to provide a taxonomy to navigate this evolving ecosystem.

In the following, we (a) explore the changing and complex landscape defining 21st century agricultural R&D; (b) scope the field of digital agricultural technologies supporting the wide range of stakeholders within agriculture across the public, private, not-for-profit, and governmental sectors; (c) present a framework to characterize and evaluate digital agriculture platforms; and (d) apply that framework to a group of digital agricultural platforms developed over the past decade.

Core Ideas

- Digital agriculture is the next major wave of innovation in food and agriculture.
- Data quality, access, interoperability, and intellectual property present sizable challenges.
- Digital agriculture platforms are an enabling environment for innovation.
- The current field lacks clarity around the core elements that constitute a digital agriculture platform.
- We establish a common vocabulary for understanding digital agriculture platforms.

2 | BACKGROUND

2.1 | The evolving digital landscape for agricultural innovation

There are several foundational and interrelated factors reshaping the food and agricultural R&D landscape that affect the ability of agricultural technologies to deliver on their promise. In parallel with the technological advances now producing a deluge of data, there is a significant shift from the public sector to private industry in both the funding and performance of food and agriculturally related R&D (Pardey et al., 2016). This means that much of the new data in agriculture are privately owned. To advance big data in agriculture, scientists – both public and private – need access to these data. However, these data are often encumbered in formal and informal property rights, and some of the data involve legitimate questions over data privacy, security, and ownership (Herbold-Swalwell, 2018; McIntosh, 2018). As a result, a substantial portion of food and agriculture data remains siloed in a host of domain-specific repositories or within public or private research labs. This lack of data sharing and interoperability makes it costly and time consuming to access and use data across studies and domains.

In addition to the sheer volume of data created, there are varying degrees of intellectual property and privacy that need to be protected. Pre-commercial collaboration can involve data created and held either as a public or semi-public (e.g., club) good, and can involve multiple public and private sector entities. Such examples include collective action around oat (*Avena sativa* L.) breeding convened by the University of Minnesota with PepsiCo, General Mills, grain millers throughout North America, USDA-ARS, and various land grant colleges (https://oatglobal.umn.edu/). In another example, the USDA-Foundation for Food and Agriculture Research (FFAR) co-funded a consortium involving the University of Wisconsin, Inari, KWS, and Syngenta to identify genetic markers in corn (*Zea mays* L.) associated with drought

14350645, 2022, 5, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/agj2.20873, Wiley Online Library on [28/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

tolerance that will accelerate the breeding of drought-resistant varieties (USDA-FFAR, 2019). In these examples and others, commercial food and agriculture R&D requires digital technologies that incorporate stringent data security protocols.

The food and agricultural R&D space has a multitude of diverse players and is constantly evolving. Organizations driving forward agricultural technology range from Fortune 500 companies to small startups, and from Land Grants Universities with more than 150 yr of experience to Ivy League Universities that are relative newcomers to the food and agricultural sciences. Agricultural technology startups have surged recently due to an increased supply of venture capital in the economy, growth in agricultural commodity prices, and previous successful sales of companies (e.g., "exits") (Graff et al., 2020). As of 28 July 2020, CBInsights identified 694 startups as "Agricultural Technology" companies, which have received on average US$21.7 million (median $3.3 million) worth of venture capital. Graff et al. (2020) identified more than 4,500 agricultural startups located across 125 countries, and illustrate how this constantly evolving landscape is difficult to characterize.

AgFunder (2019) characterized the "agri-food tech" sector as one that aims to advance state-of-the-art technologies across a wide range of fields, from "upstream" farm inputs and production system design to "downstream" consumer oriented innovations. Within this sector, they cluster companies into various groups, specifically ag biotechnology; marketplaces; bioenergy and biomaterials; farm management software, sensing and "internet of things"; farm robotics, mechanization and equipment; food supply chain technologies; novel production system generation such as indoor farming or insect protein; food innovation such as cultured meat; in-store and retail technologies; restaurant marketplaces; eGrocery; home and cooking tech; online restaurants and meal kits; along with a wide array of other technologies throughout the food and agriculture value chains (AgFunder, 2019, p. 21). Thus, while innovation requires integration both within and across these diverse data domains and market segments spanning entire supply chains in both the pre-commercial and commercial space, functionally integrating such data remains a challenge, even in cases where the relevant data management and sharing agreements are addressed.

In addition to the complex organizational landscape, many questions exist about data quality. In particular, metadata are essential for either successfully pooling data (such as genomics and field trial data; see McFarland et al., 2020) from different sources within a particular subject-matter domain or making such data functionally interoperable across domains (see Beddow et al., 2015 for such an analysis). These and related concerns have spurred guidelines to improve the management of scientifically relevant data that often lack findability, accessibility, interoperability, and reusability, though challenges remain for truly broad adoption across the public and private scientific communities (Wilkinson et al., 2016).

3 | MATERIALS AND METHODS

3.1 | Scoping digital agricultural technologies supporting agricultural innovation

Given the diverse set of public and private organizations driving agricultural technology development, our categorization of the digital agriculture landscape will inevitably be incomplete. Our intention is to provide a starting point for the broader community to begin formalizing the landscape, vocabulary, and respective unique contributions of different participants in the field of digital agriculture. Table 1 presents our taxonomy of terminology related to web-based digital agriculture research technologies widely used in food and agricultural R&D. This initial classification focuses mainly on web-based technologies, thus setting aside for now other digital agriculture technologies such as sensors, robotics, and "internet of things" technologies. Our objective was to begin demarcating the agricultural technology space into what constitutes a research platform vs. other related digital agriculture technologies (e.g., websites, repositories, archives, dashboards, and tools). We characterize these web-based digital agricultural technologies according to their portability, extensibility, security, storage, interoperability, and processing capabilities – contending that digital agricultural platforms include all six of these features, while other technologies are characterized by subsets of these features.

To generate our classification, we took both a "bottom-up" and top-down" approach by classifying 89 agricultural technology companies and products by their core technology. Table 2 summarizes the results of our classification exercise which is provided in the Supplemental Data. Of the 89 digital agricultural technologies evaluated, two-thirds are considered tools, meaning that they are primarily focused on processing or analysis of data, but do not explicitly include mechanisms to ensure interoperability between data, data protection or storage. The majority of the technologies were from private companies (89%). Private companies are classified as entities that are private for non-profit, private not-for-profit or public corporations, while public companies are considered government- or university-based operations. When classified by their user focus, roughly half (53%) were targeted to farmers. Only 17% were targeted to researchers and 5% to food and agricultural companies, while 27% had multiple user types.

3.2 | A framework to evaluate digital agriculture platforms

At the highest level, digital agriculture platforms seek to solve the general problems of **f**inding, **a**ccessing, **i**nteroperating, and **r**eusing (FAIR) data (Wilkinson et al., 2016). Given the

14350645, 2022, 5, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/agj2.20873, Wiley Online Library on [28/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Web-based digital agricultural technologies taxonomy

Technology	Definition	Attributes					
		Portable	**Extensible**	**Secure**	**Storage**	**Interoperable**	**Processing**
		works on different computational infrastructure	allows addition of new capabilities & functionalities	mechanisms in place to protect data	ability to store and retrieve data without leaving the primary site	facilitates standards-based exchange and use of information	includes processing and analysis capabilities
Platform	group of technologies upon which other applications, processes, or technologies are developed	Yes	Yes	Yes	Yes	Yes	Yes
Website	collection of publicly accessible, interlinked Web pages	Yes	No	Yes	No	Maybe	No
Repository	centralized place to store and maintain data	Maybe	No	Yes	Yes	Maybe	No
Archive	long-term storage of scientific data and methods used to read or interpret it	Maybe	No	Yes	Yes	No	No
Dashboard	information management tool that visually tracks, analyzes, and displays data	Yes	Yes	Yes	No	Maybe	Yes
Tool	mechanism used to extract actionable information from dataset	Maybe	Maybe	No	No	No	Yes

14350645, 2022, 5, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/agj2.20873, Wiley Online Library on [28/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

14350645, 2022, 5, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/agj2.20873, Wiley Online Library on [28/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 Attributes of digital agricultural technologies

	Industry			User focus			
Technology type	**Public**	**Private**	**Total**	**Researcher**	**Farmer**	**Ag company manager**	**Multiple users**
	count						
Platform	6	9	14	7	0	0	8
Website	0	10	10	2	6	0	2
Repository	2	0	3	2	0	0	0
Archive	0	0	0	0	0	0	0
Dashboard	0	1	1	0	1	0	0
Tool	1	58	59	2	43	2	12
Other	1	1	2	2	0	2	2
Total	10	79	89	15	50	4	24
	%						
Platform	6.7	10.1	15.7	7.9	0.0	0.0	9.0
Website	0.0	11.2	11.2	2.2	6.7	0.0	2.2
Repository	2.2	0.0	3.4	2.2	0.0	0.0	0.0
Archive	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Dashboard	0.0	1.1	1.1	0.0	1.1	0.0	0.0
Tool	1.1	65.2	66.3	2.2	48.3	2.2	13.5
Other	1.1	1.1	2.2	2.2	0.0	2.2	2.2
Total	11.2	88.8	100.0	16.9	56.2	4.5	27.0

unique data landscape of agricultural R&D, we propose that platforms must also meet two additional specification criteria to become FAIR-ER: One, platforms must be **e**thical – whereby the intellectual property or privacy concerns of the data owner are protected; Two, platforms must support **r**eproducibility – a stricter form of reusability that encompasses data lineage and transformation after the point of collection – whereby data are tightly tied to the original algorithms and computational environment. In addition to these high-level requirements, digital agriculture platforms may meet a wide array of additional requirements specific to their core users. Within our evaluation of digital technologies, we identified 14 (15.7%) that fit this more comprehensive definition of a platform.

To characterize the broad set of requirements in the digital agriculture platform ecosystem, we developed a framework to classify different platforms (Table 3). This framework was based on the authors' collective experience working across multiple sectors on digital agricultural projects, and provides a set of potential requirements that a platform may need to meet. The framework was applied by the authors in an iterative process, whereby documentation was reviewed for each chosen platform. Then, based on weighing the respective features across all platforms considered, the framework was updated.

In particular, we paid special attention to four broad functional areas. First, data accessibility, which describes the storage, security, and repository functionality of a platform. Second, data interoperability and integration, which describes the extent to which a platform supports harmonizing data within the same domain while maintaining extensibility. Third, data domain and target users, which is another area that differentiates the various platforms. The agricultural sciences can be characterized as covering four domains [e.g. Genetic (G), Environmental (E), Management (M), and/or Socio-economic (SE)], and are also explicitly (geo)spatial and temporal in nature. Fourth, data processing, which provides insight into each platforms' functionality for data wrangling and modeling.

To apply these categories, we surveyed each platform's primary point of contact, and inspected published information on their privacy and security policies, interoperability support, and data domain aspects of data holdings. If a non-negligible proportion (>5%) of the criteria of interest for the datasets/tools within each of the platforms in our review were present, the box was checked 'Y' below; otherwise it was scored with 'N'. For three criteria (privacy, security, and interoperability), with very few exceptions we ranked platforms according to the (cumulative) presence of relevant characteristics along a 0–5 scale, where a 5 means a

TABLE 3 Criteria used for evaluation of platforms

Criteria	Description
Data accessibility and security	
Privacy[a]	Technical and legal. 0: Data is private; 1: Data may be embargoed, but otherwise open; 2: Metadata may be shared separately from actual data; 3: Ability to share data/tools with selected users; 4: Ag data exempt from Freedom of Information Act-type requests; 5: Anonymizing/fuzzing data enabled;
Security[a]	Technical and legal. 0: Minimal protections on data. 1: secure remote backups; 2: alarmed no-access datacenter; 3: no direct access to common filesystem with other's data (e.g., with containers or software access layer) 4: data encrypted in flight and at rest; 5: All staff have Health Insurance Portability and Accountability Act (HIPPA) or other equivalent security & privacy training
Repository	Data can be accessed directly from platform
Data interoperability and integration	
Interoperability[a]	0: Data freely deposited in repository with no alteration; 1: Platform is supported by open source software; 2: Data matched to standardized vocabularies and ontologies; 3: Data matched to international metadata standards; 4: Data cleaning upon ingest for spelling errors, numerical outliers, null values; 5: Raw data and marked up data linked and version controlled
Ontologies	Match to standardized ontologies. Metadata markup.
Data model utilized	Extract, transform, and load procedure employed to populate a relational, graph, or object database governed by an explicit data model.
Row-level queries	Ability to query for individual records within data sets (this requires an explicit data model across all cataloged domains)
Extensibility	A software engineering and design principle designed to allow the addition of new capabilities and functionality.
Spatial	Emphasizes geospatial data sets in tools and/or data holdings
Temporal	Emphasizes data across a wide range of time scales
Data processing	
Cleaning and analytical tools	Provides tools for user to clean text, numeric, and geospatial outliers, as well as conduct unique analyses
User-ready models	Provides ready-to-use analytic agricultural models
Data domains and target users	
Genetics	Emphasizes genetic variety/marker data in tools and/or data holdings
Environment	Emphasizes environmental (e.g., soil, water, weather) data in tools and/or data holdings
Management	Emphasizes experimental or farmer management data in tools and/or data holdings
Socio-economics	Emphasizes socio-economic (e.g., fertilizer adoption, population density, poverty levels) data in tools and/or data holdings
Sectoral consideration	Platform cuts across agricultural production, agribusiness, science & policy, etc.
Geographic focus	Focus on data sets beyond a single country or the developed world.

[a]Rankings in these categories include all of the characteristics of lower ranked selections. For example, a ranking of 2 indicates all of the characteristics present in 0, 1, and 2.

technology is more aligned with an idealized definition of that specific attribute.

In sum, considering these factors we scanned product literature, performed metadata searches among datasets, and inspected public lists of data sources to evaluate each platform according to these above four categories. We prioritized a framework that provided differentiability among platforms and thus highlighted the relative strengths and weaknesses of each platform for specific use cases in order to illustrate the potential unique capabilities of each platform in the broader digital agriculture ecosystem.

4 | RESULTS AND DISCUSSION

4.1 | Framework application to platforms

We applied the framework to a subset of broadly conceived digital agriculture platforms developed across the public, private, and non-profit sectors (Table 4). In total across all the functional areas we evaluated, we found that the platforms included 74% of the criteria, on average. These findings represent both the well-developed nature of each respective platform and also illustrate opportunities for interoperability between platforms. In the following, we provide a more

14350645, 2022, 5, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/agj2.20873, Wiley Online Library on [28/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 4 Evaluation marks for platforms across a variety of criteria

	Private[a]							Public[a]				
Criteria	Atlas AI	Descartes Lab[b]	FarmBeats	Gro Intelligence[b]	IBM Watson[b]	KDDart	OpenTEAM	CyVerse	GARDIAN	GEMS	TERRA-REF	ALL
Data accessibility and security												
Privacy	0	NA	4	NA	NA	5	4	3	4	5	3	3.5
Security	4	NA	5	NA	5	3	3	5	NA	5	3	4.1
Repository	N	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	0.82
Data interoperability and integration												
Interoperability	5	5	1	5	0	4	4	4	3	5	5	3.7
Ontologies	Y	N	N	N	N	N	Y	N	Y	Y	Y	0.45
Data model utilized	N	NA	Y	Y	Y	Y	Y	N	N	N	Y	0.60
Row-level queries	Y	Y	Y	Y	Y	Y	Y	Y	N	N	Y	0.82
Extensibility	N	Y	Y	Y	NA	Y	Y	Y	Y	Y	Y	0.90
Spatial	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	0.82
Temporal	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	0.91
Data processing												
Cleaning and analytical tools	N	Y	Y	NA	NA	Y	Y	Y	Y	Y	Y	0.78
User-ready models	N	Y	Y	Y	Y	N	Y	N	N	N	Y	0.55
Data domains and target users												
Genetics	N	N	N	N	N	Y	Y	Y	Y	Y	Y	0.55
Environment	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	1.00
Management	Y	N	Y	Y	Y	Y	Y	N	Y	Y	Y	0.82
Socio-economics	Y	N	N	Y	N	N	Y	N	Y	Y	N	0.45
Sectoral consideration	Y	Y	Y	Y	Y	N	Y	N	Y	Y	N	0.73
Geographic focus	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	0.91
Overall score	0.60	0.73	0.78	0.87	0.73	0.69	0.90	0.47	0.73	0.83	0.84	0.74

Note. NA, no information was found.

[a]Private indicates entities that are private for-profit, private not-for-profit or public benefit corporations. Public indicates entities that are government- or university-based operations.

[b]Representatives for these platforms explicitly declined to answer our survey questions on their platform so we consulted their website and published literature. Note: To interpret scores in this table, see Table 3. To calculate the average criteria, score across all platforms or the overall score by platform, we first standardized all scoring along a 0–1 scale – the No/Yes indicator was transformed to a binary 0/1 indicator. Total scores represent the sum of criteria relative to the total for each criteria or platform, respectively. Missing values were removed before calculations.

14350645, 2022, 5, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/agj2.20873, Wiley Online Library on [28/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

detailed look within each of the four major categories we evaluated.

1. *Data Accessibility and Security*. All but 2 of the 11 platforms we evaluated supported a data repository, and although five of the eight platforms for which we could discern relevant information have strong data security practices in place, three of the platforms (i.e., those scoring 3) did not encrypt their data nor require all staff to have data security training. Only two of the platforms (KDDart and GEMS) scored a 5 on the privacy criteria. Many of the others, including open platforms such as TERRA-REF and Cyverse as well as private platforms such as FarmBeats, did not support data anonymization or fuzzing, while GEMS was the only public platform in the evaluation that was legally exempt from information requests, including Freedom of Information Act (FOIA) requests. Finally, the privacy and security requirements of purely open platforms such as TERRA-REF and Cyverse are quite different from completely commercial platforms such as Azure FarmBeats and IBM Watson Decision Ag Platform.
2. *Data Interoperability and Integration*. Some platforms make a concerted effort to map user input to accepted agricultural concept schemes (Baker et al., 2019), ontologies (Cooper et al., 2018) and metadata standards (https://www.dublincore.org/, https://www.fgdc.gov/metadata/csdgm/) such as GEMS, OpenTEAM, and TERRA-REF, but the others do not. In fact less than half (45%) mapped user inputs directly to supported ontologies, which are critical for achieving findable, interoperable and reproducible data standards (Wilkinson et al., 2016).
3. *Data Processing*. Most (78%) of the platforms included tools that facilitated the cleaning of text, numeric or geospatial outliers, and/or enabled users to conduct unique analyses. Fewer (55%) of the platforms provided users with ready-to-use analytical agricultural models.
4. *Data Domains and Target Users*. In terms of data domains and target users, almost all platforms supported environment and management data with 100 and 82% of platforms, respectively. Genetics and socioeconomic considerations were less represented with 55 and 45% of platforms encompassing such factors, respectively. Only GEMS, Gardian, and OpenTeam covered the entire 'G', 'E', 'M', and 'SE' domain spectrum across time and space, but AtlasAI, GroIntelligence, and KDDArT each cover all these aspects except one ('G', 'G', and 'SE', respectively).

In sum, different platforms have different areas of specialization. Our focus here has been on factors within a single platform. In reality, over time it is likely there will be more, not fewer, platforms. Going forward, a major research and development challenge will be to further clarify these relative strengths and weaknesses and define standards for data exchange and interoperability among platforms. This presents an opportunity for additional cross-organization and sectoral collaboration where each can emphasize areas of respective strength while also supporting deeper integration across platforms.

5 | CONCLUSION

In our conception, digital agriculture platforms are those technology environments that are portable across computational infrastructures, extensible to new capabilities, secure and protect data, facilitate standards-based interoperability, and enable data processing and analysis (Table 1). Here, we have provided a preliminary framework to provide a vocabulary and conceptual schema to support the agricultural community in navigating the expanding ecosystem of digital agricultural technologies and providing guidance where further innovation is needed. Digital agriculture platforms intend to solve the general problems of finding, accessing, interoperating, and reusing data, while also providing the baseline technologies for ethical and reproducible science. As these digital technologies continue to evolve, there is sure to be value in expanding and revising our initial efforts at technology classification, hopefully building off the version 1.0 schema we have developed and deployed here.

ACKNOWLEDGMENTS

The authors wish to thank the following platform creators and developers for answering a set of questions on their platforms which were instrumental in filling out Table 4: George Azzari for AtlasAI, Eric Lyons for CyVerse, Biju Venugopal for FarmBeats, Andrzej Kilian for KDDArT, Dorn Cox for OpenTEAM, and David LeBauer for TERRA-REF.

AUTHOR CONTRIBUTIONS

Bryan C. Runck: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Validation; Visualization; Writing-original draft; Writing-review & editing. Alison Joglekar: Conceptualization; Data curation; Formal analysis; Methodology; Validation; Visualization; Writing-original draft; Writing-review & editing. Kevin A. T. Silverstein: Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Validation; Visualization; Writing-original draft; Writing-review & editing. Connie Chan-Kang: Data curation; Formal analysis; Investigation; Validation; Writing-review & editing. Philip G. Pardey: Conceptualization; Funding acquisition; Writing-review & editing. James C. Wilgenbusch: Conceptualization; Funding acquisition; Writing-review & editing.

14350645, 2022, 5, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/agj2.20873, Wiley Online Library on [28/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

DATA AVAILABILITY STATEMENT

Supplemental data (Runck et al., 2021) are available at https://datadryad.org/stash/share/sNeu72Aw8Gp_bkTfG7SOMfF-PQRaiDfCzef0BENcf88.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Bryan C. Runck https://orcid.org/0000-0002-7015-1539
Alison Joglekar https://orcid.org/0000-0003-3924-9877
Kevin A. T. Silverstein https://orcid.org/0000-0002-4955-3218
Philip G. Pardey https://orcid.org/0000-0002-8012-1341
James C. Wilgenbusch https://orcid.org/0000-0001-9464-1578

REFERENCES

AgFunder (2019). AgFunder Agri-FoodTech investing report. AgFunder. https://agfunder.com/research/agfunder-agrifood-tech-investing-report-2019/

Baker, T., Whitehead, B., Musker, R., & Keizer, J. (2019). Global agricultural concept space: Lightweight semantics for pragmatic interoperability. *npj Science of Food*, *3*, 16. https://doi.org/10.1038/s41538-019-0048-6

Beddow, J. M., Pardey, P. G., Chai, Y., Hurley, T. M., Kriticos, D. J., Braun, H. -. J., Park, R. F., Cuddy, W. S., & Yonow, T. (2015). Research investment implications of shifts in the global geography of wheat stripe rust. *Nature Plants*, *1*(10), 1–5. https://doi.org/10.1038/nplants.2015.132

Cooper, L., Meier, A., Laporte, M.-A., Elser, J. L., Mungall, C., Sinn, B. T., Cavaliere, D., Carbon, S., Dunn, N. A., Smith, B., Qu, B., Preece, J., Zhang, E., Todorovic, S., Gkoutos, G., Doonan, J. H., Stevenson, D. W., Arnaud, E., & Jaiswal, P. (2018). The Planteome database: An integrated resource for reference ontologies, plant genomics and phenomics. *Nucleic Acids Research*, *46*(D1), D1168–D1180. https://doi.org/10.1093/nar/gkx1152

Graff, G. D., de Figueiredo Silva, F., & Zilberman, D. (2020). Venture Capital and the Transformation of Private R&D for Agriculture. In P. Moser (Ed.), *Economics of research and innovation in agriculture*. University of Chicago Press for NBER.

Gustafson, A., Erdmann, J., Milligan, M., Onsongo, G., Pardey, P., Prather, T., & Zhang, Y. (2017). A platform for computationally advanced collaborative agroInformatics data discovery and analysis. In *Proceedings of the ractice and experience in advanced research computing 2017 on sustainability, success and impact* (Vol. Part F1287, pp. 1–4). ACM. https://doi.org/10.1145/3093338.3093376

Gray, B., Babcock, L., Tobias, L., McCord, M., Herrera, A., Osei, C., & Cadavid, R. (2018). *Digital farmer profiles: Reimagining smallholder agriculture*. USAID, Feed the Future.

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

Mcfarland, B. A., Alkhalifah, N., Bohn, M., Bubert, J., Buckler, E. S., Ciampitti, I., Edwards, J., Ertl, D., Gage, J. L., Falcon, C. M., Flint-Garcia, S., Gore, M. A., Graham, C., Hirsch, C. N., Holland, J. B., Hood, E., Hooker, D., Jarquin, D., Kaeppler, S. M., …, & De Leon, N. (2020). Maize genomes to fields (G2F): 2014–2017 field seasons: Genotype, phenotype, climatic, soil, and inbred ear image datasets. *BMC Research Notes*, *13*(1), 1–6. https://doi.org/10.1186/s13104-020-4922-8

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

Pardey, P. G., & Alston, J. M. (2021). Unpacking the Agricultural Black Box: The Rise and Fall of American Farm Productivity Growth. *The Journal of Economic History*, *81*(1), 114–155. http://doi.org/10.1017/s0022050720000649

Pardey, P. G., Chan-Kang, C., Dehmer, S. P., & Beddow, J. M. (2016). Agricultural R&D is on the move. *Nature*, *15*(537), 301–303.

Techopedia (2020). *Platform*. https://www.techopedia.com/definition/3411/platform-computing

Runck, B. C., Joglekar, K. B., Silverstein, K. A. T., Chan-Kang, C., Pardey, P. G., & Wilgenbusch, J. B. (2021). Digital agriculture platforms supplemental data, dryad, dataset. https://doi.org/10.5061/dryad.bvq83bk93

Trendov, N. M., Varas, S., & Zeng, M. (2019). *Digital technologies in agriculture and rural areas*. Food and Agriculture Organization of the United Nations.

USDA-FFAR (2019). FFAR Grant maps corn drought tolerance genes. https://foundationfar.org/2019/09/11/ffar-grant-maps-corn-drought-tolerance-genes/

Wikipedia (2020). Digital agriculture. Wikipedia. https://en.wikipedia.org/wiki/Digital_agriculture

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*. https://doi.org/10.1038/sdata.2016.18

Wilgenbusch, J., Lynch, B., Hospodarsky, N., & Pardey, P. (2021). Addressing the new data privacy and security realities affecting agricultural R&D. *Agronomy Journal*, (this issue).

Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M.-J. (2017). Big data in smart farming – A review. *Agricultural Systems*, *153*, 69–80. https://doi.org/10.1016/j.agsy.2017.01.023

World Bank. (2016). *Digital dividends: World development report 2016*. The World Bank.

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: Runck, B. C., Joglekar, A., Silverstein, K., Chan-Kang, C., Pardey, P., & Wilgenbusch, J. C. (2022). Digital agriculture platforms: Driving data-enabled agricultural innovation in a world fraught with privacy and security concerns. *Agronomy Journal*, *114*, 2635–2643. https://doi.org/10.1002/agj2.20873

14350645, 2022, 5, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/agj2.20873, Wiley Online Library on [28/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License