Abstract: Advanced digital technologies are rapidly permeating agriculture from laboratory to field. Machine-based breeding, robotics and big data technologies have deeply transformed not only production systems but also the way scientific research is conducted. How are digital applications revolutionizing people’s jobs and skills? What are the challenges and opportunities for managing and sharing agricultural big data? This article addresses these and other questions by surveying international experts in plant biotechnology. Results show that digital innovations in the form of decision-support tools are perceived as promising. Most surveyed experts anticipate the deployment of big data analytics and artificial intelligence to boost agricultural productivity. Another key finding is that substantial physical investment, specialized human capital and effective data governance are critical to successful implementation of technological innovations associated with big data.

Keywords: artificial intelligence; big data analytics; data sharing; food security; innovation; plant breeding

1. Introduction

Technological advances in their scale and breadth are revolutionizing entire production systems, their management, and policy making and governance of all economic activities. The agri-food sector, like other industries, is becoming increasingly data-driven and data-enabled. From crop breeding to the assessment of consumer food preferences, data are generated at every stage of the agricultural value chain. Compiled and mined, massive amounts of data, so-called “big data” have already fundamentally transformed the structure and responsiveness of various agricultural production and supply systems. Alongside land, labor and capital, big data are already considered by some as a fourth production factor [1].

Agricultural data in general, and farm data in particular, have gained importance with the gradual progression towards precision and digital agriculture [2–5]. Technologies that have enabled precision agriculture to flourish, have also paved the way for management practices that use farm inputs more efficiently, thus optimizing agricultural output [4]. A significant positive externality of precision agriculture is a reduced environmental footprint [4,6]. Precision agriculture has evolved significantly since the first precision agricultural tools, including geographic information systems (GIS) in the 1960s, intelligent devices and implement components (e.g., yield monitors, on-the-go nitrogen testers) in the 1980s, global positioning systems (GPS) in the 1990s, and now the Internet of Things, decision support systems and machine-to-machine communications [4,7]. Shaped by big data and advanced analytics, precision agriculture has evolved into digital agriculture [6,8]. The latter refers to the generalized adoption of cutting-edge computational and information technologies to advance profitable and sustainable agriculture [9,10]. Digital agriculture makes use of advanced precision technologies along with intelligent networks and data management.
tools [6]. By incorporating modern technologies and applications, new agriculture routines—alongside other solutions (e.g., new products, different plant breeding techniques) aim to address increasing demand for food from a larger world population, disruptions caused by climate change, and scarcity of land and other natural resources [11,12].

The expected outcomes of digital agriculture are sometimes overstated in the scientific literature [9]. Given the relative novelty of the concept of big data, quantitative evidence on the benefits and intricacies of agricultural big data and knowledge about their impact on research and development is still emerging [13,14]. This article provides insights from a panel of international experts on how big data are currently used in their work environment, what related benefits in the medium term they expect, what issues are likely to arise in the data-sharing process. Experts were also asked about their opinions regarding the use of big data analytics in agriculture, and what impacts artificial intelligence could have on the agri-food sector. The article draws preliminary inferences on the enabling and disrupting impacts of technological innovations associated with big data in the midst of rapid technological and social change.

Following this introductory section, there is a discussion of big data and its potential use in the agriculture industry. The role of artificial intelligence on big data is then explored. Next, we describe the method and survey design, followed by results and discussion. The final section provides conclusions.

2. Big Data Ontology

Despite the ubiquity of the term “big data” in various academic and non-academic publications, a formal definition for the concept does not exist yet. According to Kacfah Emani, Cullot [15], the term became part of the academic lexicon in 2005 in order to describe large and heterogeneous data sets that are hard to manage given traditional data management tools and techniques. Since then, many scholars have set forth their own formal definitions for the concept [1,15–17].

Though definitions of big data differ slightly, underlying them all is a common ontology. That is, what characterizes and sets big data apart from what is simply referred to as data (or small data) are their volumes, varieties and their velocity—the three most common “Vs” found in the literature. Laney [18] points out that lower costs for and enhanced availability of digital infrastructure (e.g., internet, computing capacity) have encouraged firms to view data as an asset. This in turn has made them rethink how they manage their data. Instead of discarding it as it is used, the goal now is to accumulate large volumes of data. Variety of data refers to the multitude of sources from which data originate. In agriculture, not all segments of the value chain capture data in the same fashion. There is no standard by which data are captured, which makes the harmonization and compilation of data from various sources difficult. Finally, velocity, refers to the rate or frequency with which data are generated and captured.

Other authors and scholars add three other “Vs” to further characterize big data: value, veracity and visualization [15]. Veracity refers to the quality with which sensors and people record data [19]; value refers to how data are appraised by both individuals and the agricultural industry as a whole; and visualization refers to the ability to graphically depict the underlying data. Thus far, most proposed definitions for the term operate within the bounds of these six “Vs”. Given that multiple disciplines use and generate big data in different ways depending on their specific goals and many value the Vs differently, it could be some time before academia settles on a universal definition.

An estimated 95% of data are generated in an unstructured (raw) format (e.g., audio, images, video and unstructured text). The view is that they are “worthless in a vacuum” [20]. That is, data become useful when they are organized in a way that assists with decision making. Thus, new technologies are required to manage and analyze big data to generate meaningful added value. The process of extracting value from big data includes: data management, which involves processes and technologies to acquire, store, and prepare data analysis; and analytics, comprising techniques to analyze and generate intelligence.
from big data [20,21]. This extraction process requires substantial investment, coordination and expertise. The lack of skilled data analysts and data scientists is a challenge for many industries. The talent gap limits the ability to aggregate and interpret data in a way that contributes to decision-making [22]. Davenport [23] proposes to bridge the talent gap with artificial intelligence (AI) (discussed below) by automating tasks, such as through machine learning.

3. Agricultural Big Data Generation and Its Potential Use

The deluge of data generated at every stage of the agricultural value chain is in great part explained by progress in sensor apparatus technologies. Table 1 briefly defines data types and analytical techniques. Be they digital ledgers, or telematics equipment attached to farm machinery (capturing machine diagnostics, time and motion data etc.), technology has become versatile enough to allow for the capture of more detailed data at a faster rate. Raw materials used in agricultural production are likely to remain unchanged given emerging digital-era technologies. What big data are likely to show at every stage of the agricultural value chain are previously unseen (some previously unknowable) patterns that will yield novel sets of instructions that when followed will result in significant productivity gains. For example, enhanced sensors on farm machinery in combination with soil data from telematics equipment will allow for tailoring of nitrogen fertilization regimens, resulting in farm productivity gains, positive environmental externalities and fewer inputs required to guarantee agricultural production [24].

Table 1. Big data types and analytical methodologies.

| Data and Sensor Types | Big Data Analysis Techniques and Examples |
|----------------------|------------------------------------------|
| **Geospatial data:** refers to data about objects, events, or phenomena that have a location on the surface of the farm. | **Audio analytics:** sensors mounted on farm equipment can detect sound waves for abnormalities in equipment functioning. |
| **Meta data:** | **Predictive analytics:** weather data going as far back as possible can be used to predict weather going into the future. Weather forecasting is already done. |
| - Application dates: when, how much and at what crop development stage were pesticides, herbicides and fertilizers applied. | **Social media analytics:** social media play an increasingly important role in daily life. Information on these platforms could complement decision-making processes. Data on these platforms can be collected and analyzed with “big data” techniques. |
| - Cultivar selection: what crop varieties were planted. | **Text analytics:** AI algorithms go through large volumes of text to glean information of interest. |
| - Planting depth: depth at which crop seed was sown. | **Video analytics:** cameras mounted on equipment use AI to uncover patterns, otherwise unnoticeable, that can aid in decision making. |

Telematics data:

- Farm-equipment collected: sensors mounted on combines, tractors, sprayers, etc.
- Machinery diagnostics: sensors indicating the energy exerted and fuel requirements by machinery depending on task. Or whether the equipment needs maintenance.
- Time and motion: amount of time machinery has devoted to a task. If the machine moves within-the-farm sensors indicate the paths taken.

Based on: [19,25,26].

One challenge is that not everyone may want to share their data. Ellixson and Griffen [27] explore the specific case of farm-level big data. They posit that though farmers may place value on the data flowing from their farms, data from any one individual farm is of little, if any, value. On the other hand, the compilation of farm-level data from many farms in a specific geographic region (big data), is likely to yield insights from which new management practices could be developed. Given that the value of regional farm-level data is greater than the value of any individual farmer (i.e., the “network effect”), it is likely that the “data aggregator” will receive most of the benefits from these data (Bühlmann et al., 2016). That is, at first, the incentive to farmers may be to “sell” their data. However, once the data accumulator has secured enough big data that enables it to mine and apply this new knowledge to farm production, farmers’ incentives may shift. The incentive then may be to continue supplying farm-level data in exchange for access to system-level big
data, increase the price demanded for selling individual farm-level data, or cease to supply farm-level data all together. It remains to be seen how farm-level data sharing will evolve.

4. Artificial Intelligence

Artificial intelligence (AI) is a key area of research in computer science that can also be understood as an extension of statistics [23,28]. In order to learn new concepts and tasks and to reason and draw relevant conclusions about the world, AI has to have access to big data. However, once the AI algorithms have mined big data, the algorithms refine themselves on their own as AI interacts with its environment in an interactive learning process. AI simply learns from the past using deep learning and machine learning, enhancing how machines work [29]. Currently, there is significant research into the development of autonomous farm machinery in various institutions around the world.

Bannerjee et al. [30] conducted a comprehensive literature survey on the applications of various AI techniques in agriculture, concluding that the realm of AI—with its rigorous learning capabilities—is richer and more complex than autonomous farm equipment. There are no fewer than 16 sub-disciplines of AI [31], some of which are being applied to agriculture for yield prediction algorithms based on weather and historical yield data, image recognition algorithms to detect pests and diseases in plants, and robotics to harvest diverse types of specialty crops [32]. For instance, Xia et al. [33] describe how AI is aiding in seed viability selection, Patrício and Rieder [34] compile 25 studies that employ AI algorithms to aid in grain production and Shadrin et al. [35] present a novel way in which AI could automate sensors to aid in determining plant leaf growth dynamics.

Given that most potential AI applications in the agri-food sector are still in the development stage, challenges and impacts remain unexplored [36]. Complementarity or substitution between agricultural workforce and automation is one concern. It has been argued that robotic automation of time-consuming and repetitious tasks will allow humans to focus on the high-value task and more difficult problems, but at the cost of some employment. With progressively advancing applications of machine learning, robots are gradually performing not only monotonous activities but also some non-routine and cognitive tasks such as yield prediction, disease detection, species recognition, crop quality estimation, and soil management [37,38]. The organization and management of the breeding space will inevitably be transformed by this new technology.

5. Method

In 2015, a research team at the University of Saskatchewan began conducting a multi-year survey project that investigated expert opinions regarding the application of new plant breeding techniques as one solution to ensuring global food security. Earlier surveys examined the regulatory and social barriers pertaining to novel breeding approaches using genome editing and any related risks and benefits [39–42], among other topics. Along with the new breeding methods, incorporating modern cross-industry technologies and applications (e.g., data analytics, drone technology and artificial intelligence) has also been proposed as another way to boost productivity and achieve food security [11]. To gain more insights about this, the article reports the results of an online survey conducted between April and September of 2019. The survey was designed to gather insights on how experts are coping with, and adapting to, the use of big data technologies. Specifically, the questionnaire solicits expert opinion pertaining to the opportunities and challenges of using big data and related analytics in their work environment. They were also asked about the importance of big data analytics and the impact of artificial intelligence in the agri-food sector. The survey was emailed to a panel of 466 international scientists, government officials, and agribusiness professionals involved in plant biotechnology. The expert panel enrolled in the survey project was obtained from a contact database that was created using emails of participants for several conferences on biotechnology organized by the lead researchers over the past 15 years, as well as of experts from online searches (i.e., websites of universities, research institutions, biotech companies and government agencies).
Our study was deemed exempt from full ethics review by the Behavioural Ethics Board at the University of Saskatchewan on 26 March 2019 on the basis that the participants, as experts, were not themselves the focus of the research (BEH 97) (Per the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans, December 2014, Exemption Article 2.1). Nevertheless, our survey presented participants with a standard consent statement describing the study, identifying the absence of known risks associated with participation, and a reminder that participation was voluntary, and responses would be anonymous and confidential. Upon expression of consent, participants were presented with the questionnaire.

The survey was administered in three parts (Appendix A). In the first part, experts were asked about their use of big data, foreseeable benefits, related applications and analytics that enable digital agriculture as well as potential barriers to its adoption. In the second part, participants were invited to offer opinions on the impact of artificial intelligence in the agri-food industry. The third part surveyed circumstances for data sharing. Throughout the survey, brief definitions of both big data and artificial intelligence found in the literature were provided.

6. Results and Analysis

Eighty-one participants (n = 81), including 75% males and 24% females, completed the survey; a response rate of 17.4% (only one respondent selected “I prefer not to say” when asked). Eighty percent of the respondents reside in the developed world, including 48% in North America (S and Canada), 25% in Europe, and 7% in Oceania; and the reminder lived in Africa, Asia and Central and South America (9%, 4%, and 7%, respectively). Seventy-two percent identified themselves as life scientists and 21% as social scientists (e.g., regulators, business managers). Seven percent of the sample selected “Other” when asked to identify their expertise. Given both the nature of the questions and the size of the sample, we have not sorted the responses by nature of the expertise of our respondents. A quarter of the respondents indicate they work for an industry or a private research institution, 33% for an academic institution, and 27% for a government or a public research institute. Fifteen percent of the respondents selected “Other”.

6.1. Big Data Use and Derived Benefits

Participants (n = 81) were asked whether they or their organizations acquire and manage any type of data. While 47% respondents reported that they or their organization do not gather or analyze big data, 7% reported collecting big data, 4% analyzing it, and 42% do both. Among those who collect big data (40/81 respondents representing 49% of the total sample), 22% collect only structured data (generated reports, records, tables, survey data, sensor data, customer data, etc.), 43% work mostly with unstructured data (emails, PDF texts, voice messages, web searches, social media, etc.) and 35% deal with both structured and unstructured data equally. Among those who analyze data (37/81 respondents representing 46% of the total sample), 27% analyze only structured data (generated reports, records, tables, survey data, sensor data, customer data, etc.), 38% work mostly with unstructured data (emails, PDF texts, voice messages, web searches, social media, etc.) and 35% deal with both structured and unstructured data equally. As data collection is done by only 7% of the respondents and none of them work with unstructured data solely, this reflects a shift from the era when the focus was on data volume (collection and storage) to the era of data quality or value (through real-time data processing and analysis).

Compared to peers in their field, 17% of the total sample (n = 81) believe their organization is very far ahead or ahead, 39% think it is behind or very far behind and 38% consider themselves neither ahead nor behind, with a mean confidence level of 80% (with a standard deviation of 19.55). Companies that most effectively use and control data are coming to dominate major parts of the economy, so much so that the highest valued by market capitalization are all data-driven technology firms (e.g., Apple, Amazon, Alphabet/Google and Microsoft). We asked our panel of 81 respondents whether, and how, they
deal with these firms. Thirty-six percent of respondents stated that they prefer to develop their own applications internally and 31% via formal contracts. Some 17% indicated that they generally buy applications off the shelf, while 36% were uncertain about the source or terms for the big data applications used in their work environment. Among those engaged in contracting big data applications (25/81 representing 31% of the total sample), 28% (7/25) indicated that they have contracted with leading big data companies while 40% (10/25) did not contract with these, and 32% (8/25) did not answer.

As to the benefits derived from the use of big data, a majority of respondents report that planning and forecasting (72%), the identification of problems (70%), decision-making (67%), risk management (67%) and productivity (57%) are all likely to improve in their respective organizations as they make use of big data (Table 2). Views varied as to whether big data will improve profit. This result is not surprising given that 60% of the sample work for public organizations such as universities and government organizations where profit optimization is generally not a motivating organizational goal.

| Potential Benefits                                      | Very Unlikely /Unlikely | Neither Likely nor Unlikely | Likely/Very Likely | Do Not Know |
|---------------------------------------------------------|-------------------------|-----------------------------|--------------------|-------------|
| Better planning and forecasting                         | 9                       | 13                          | 72                 | 6           |
| Better identification of root causes of problems        | 10                      | 15                          | 70                 | 5           |
| Better decision-making                                  | 9                       | 19                          | 67                 | 5           |
| Better risk management                                  | 15                      | 22                          | 55                 | 7           |
| Better targeted clients/stakeholders                    | 16                      | 16                          | 64                 | 4           |
| Increased productivity                                  | 11                      | 23                          | 57                 | 9           |
| Increased profit                                        | 22                      | 27                          | 37                 | 14          |

Table 2. Perceived benefits of big data use by surveyed experts (% of responses, n = 81).

6.2. Challenges to Adopting and Implementing Big Data Analytics

To obtain a sense of which barriers to implementing big data analytics are significant, we asked our experts (n = 81) to identify the most important factors to successful big data implementation and the top barriers to using big data analytics in their work environment. As displayed in Table 3, big data implementation in an organization hinges on a clear strategy and means to execute it, as well as training of personnel to manage the large volumes of data. Training and talent, more than capital, are most critical for future optimal production.

Table 3. Perceived important factors to successful implementation of big data analytics by surveyed experts (% of responses, n = 81).

| Factors                                                      | Weighted Score |
|--------------------------------------------------------------|----------------|
| Supporting systems and procedures                            | 17             |
| A clear company strategy                                     | 16             |
| Training                                                     | 13             |
| Talent                                                       | 12             |
| A sound procedure for legal, ethical and reputational issues | 11             |
| Support by higher management                                 | 11             |
| Financial budget                                             | 10             |
| An organizational structure that supports multi-disciplinary projects | 7             |
| Other                                                        | 2              |

Total (%) 100

The score is a weighted sum value of the 5 ranked responses. Items ranked first were multiplied by 0.4. Ranks 2, 3, 4, and 5 were weighted 0.3, 0.15, 0.1 and 0.05, respectively.

Table 4 shows that 69% of the total sample identified the lack of qualified specialists in big data analytics and the cost of investment (61%) as the major barriers to the adoption of big data within their organizations, followed by the lack of strategic vision at the management level (40%). Less than a quarter of the panel reported data privacy and
ownership as barriers (22%). Inadequate staffing or skills for big data analytics were also found to be hurdles to implementing and improving big data analytics in reports by The Data Warehouse Institute (TDWI), e.g., [43,44]. Over half of the survey participants predict an increase in the number of big data specialists in their respective organization within the next three years. If this occurs, this could exacerbate the current shortages in human capital, especially data specialists—data analysts and data scientists—with the technology skills and deep understating of big data analytics necessary to implement AI-based innovations. This could pose an insurmountable barrier to full adoption and implementation of big data analytics.

Table 4. Experts’ perceived barriers to the adoption of big data analytics in work environments (% of responses, n = 81).

| Barriers                                                                 | %   |
|-------------------------------------------------------------------------|-----|
| Shortage of skilled experts in big data analytics within the organization | 69  |
| High overall cost of investment                                         | 61  |
| Lack of strategic vision/interest by the management (i.e., not a priority) | 40  |
| Cultural and communication barriers to the integration of new information technologies in the work environment | 38  |
| Lack of funding                                                          | 32  |
| No identifiable end users/market                                         | 25  |
| Uncertainty about data privacy/ownership                                 | 22  |

6.3. Data Sharing

Experts (n = 81) were asked whether they or their organization would be willing to share machine-generated data. Forty-three percent indicated they are willing to do so under some terms whereas 14% are not willing to do so under any terms; 43% report that their organization does not own machine data to share. As illustrated in Table 5, respondents (n = 35) are willing to share their machine data for some benefits. Some respondents indicated that they would share data for public benefit—whether directly with others (28%) or by publishing them for open access (31%). Respondents would also share data with others (32%) or publish them (26%) for professional gain (scientific merit, partnership, etc.). To maintain transparency, 20% of the respondents would share data with others and 25% would publish them.

Table 5. Surveyed expert motives for data sharing (% responses, n = 35).

| My Data Will be Shared | Sharing Data with Others | Make Data Publicly Available |
|------------------------|--------------------------|------------------------------|
| for public benefit     | 28                       | 31                           |
| for professional benefit| 32                       | 26                           |
| for transparency and re-use | 22                       | 25                           |
| because data sharing is standard practice in my field | 19                       | 15                           |
| for free               | 14                       | 11                           |
| for monetary incentives | 11                       | 9                            |

Total columns do not add up to 100% as the task was multi-response.

For 14% (11/81 participants) who are unwilling to share data, security concerns over the handling of sensitive or confidential data (9%), risk of cybercrimes (7%) and worries about intellectual property rights and expropriation of ideas (9%) are concerns that limit data sharing.

Not all partners are viewed equitably. Trust in data and its management varies significantly (Table 6). Most surveyed experts have higher trust in universities (60%), and statistical bureaus (54%) but lower confidence in sharing data with governments, agribusinesses, equipment manufacturers, consultancies and banks or other financial institutions. Clearly, experts are more attracted by win–win exchanges than by exchanges with those they view as having different motives and interests.
Table 6. Surveyed expert trust in different sources of data management (% responses, n = 81).

| Source                                      | I Do Not Trust at All | I Trust a Little | I Trust a Lot | No Opinion |
|---------------------------------------------|-----------------------|-----------------|--------------|------------|
| Universities                               | 6                     | 28              | 60           | 6          |
| Statistical bureau                         | 7                     | 29              | 54           | 10         |
| Government                                 | 20                    | 36              | 38           | 6          |
| Farmers                                    | 4                     | 48              | 36           | 12         |
| Agribusinesses                             | 12                    | 62              | 16           | 10         |
| Companies providing equipment              | 15                    | 59              | 15           | 11         |
| Consultancy agencies                       | 21                    | 52              | 15           | 12         |
| Banks/Financial institutions               | 27                    | 47              | 14           | 12         |

6.4. Big Data Analytics and Artificial Intelligence (AI) in Agriculture

When the expert panel (n = 81) was asked about enabling analytics for digital agriculture (Table 7), 58% of surveyed experts identified cross-cutting technologies such as computational decision aid tools that address system dynamics and optimization issues as the most enabling (to develop recommendations for management and optimize some farm tasks). Forty-eight percent identified sensors that gather information on the functioning of equipment and farm resources to support management decisions and 48% identified digital communication tools (Table 3). Other field-based technologies important to the future of agriculture include geographic information (46%) and geo-locating systems that collect real-time data (40%) (e.g., to use computerized mapping to aid inventory management and to make geographical crop input prescriptions such as for fertilizer).

Table 7. Perceived enabling analytics for digital agriculture by surveyed experts (% of responses, n = 81).

| Production Environment | Type of Analytics | %   |
|------------------------|-------------------|-----|
| Cross-cutting          |                    |     |
| technologies           | Computational decision tools | 58  |
|                        | Sensors            | 48  |
|                        | Digital communication tools (e.g., mobile, broadband) | 48  |
|                        | The cloud          | 32  |
|                        | Robots (e.g., drones) | 27  |
| Field-based activities | Geographic information system (GIS) | 46  |
|                        | Geo-locating (e.g., GPS; global positioning system) | 40  |
|                        | Spectral reflectance sensing | 36  |
|                        | On-board computers | 35  |
|                        | Variable rate technology | 31  |
|                        | Unmanned aerial systems | 26  |
|                        | Yield monitors     | 25  |
|                        | Precision soil sampling | 25  |
|                        | Auto-steering and guidance | 19  |

Notes: Adopted from van Es and Woodard [9]. Details on the purpose and benefits of each analytics can be found in the survey (Appendix A. Question 6).

Most producers in developed countries have adopted GPS/GIS-guidance to automate farm machinery navigation (tractors, combines, sprayers and seeders, etc.) while much of Asia, Africa and Latin American have not yet adopted modern machinery [45]. Fully-autonomous farm machinery is expected to be the next extension of this type of technology. Instead of requiring human operation, these machines will require only intermittent and minimal human supervision. Indeed, the 2018 launch of SeedMaster’s autonomous Dot Power Platform in Canada marked the first commercial deployment of this type of technology in agriculture [46].

Our surveyed experts (n = 81) anticipate that deployment of AI offers moderate, large or substantial contributions to overall agriculture (35%, 33% and 23%, respectively), albeit with a somewhat lower level of confidence in their judgment (73% on average). There is a consensus among participants that AI will contribute positively to agricultural
machinery and logistics (each at 90%), market information (89%), plant breeding (81%) and risk management (71%). In a recent study that reported results of text mining of news articles from 2014–2019, Lakshmi and Corbett [36] found that, globally, AI is primarily being applied to increase productivity and efficiency and secondarily to address labor shortages and environmental sustainability concerns. They found more active adoption of AI in North America and Europe, with smaller but increasing efforts in Asia and Africa.

Generally, the view was that AI would have modest but measurable effects on the workforce. Forty-six percent of total respondents \( (n = 81) \) expect no substantial change to the number of agricultural jobs with the deployment of AI, while 26% anticipate job losses as precise technology (e.g., robots, sensors, etc.) replaces low-skilled labor; 16% foresee an increase in high-skill jobs required to leverage AI applications. Twelve percent did not know how AI will affect agricultural jobs. According to the 2019 employment outlook of the Organization for Economic Co-operation and Development (OECD), 14% of all jobs in any economy are at high risk of automation and 32% of jobs could be radically transformed [47]. The agricultural results suggest that the transition could be greater, as lower-skilled work will be replaced by automation and algorithms. The challenge may be that the new jobs will not be located in the same place as the lost jobs, as digital workers often congregate in larger urban centers and telecommute to serve their clients.

7. Conclusions

Surveyed experts anticipate benefits from big data, be they internally-created or outsourced applications. Most respondents think that the deployment of big data analytics and artificial intelligence will make a positive contribution to agricultural productivity, but these benefits may, for a time, only enhance welfare in developed countries. Developing countries, faced with various socio-economic constraints, are likely to lag in the capture of these benefits.

Experts think that successful applications of AI and big data will require substantial physical, human and institutional capital investments. Indeed, lack of qualified data specialists and the cost of investments were identified as major barriers to the adoption of big data analytics within participant organizations. Surveyed experts anticipate a surge in demand for skilled analysts and data scientists. This talent gap is a critical challenge facing most economic sectors targeted by AI and big data [23]. The particular challenge for agriculture is that effective digitization requires both general computational skills and specific knowledge of the industry; the returns on investment in agriculture are viewed as both less competitive than in other sectors and more uncertain [9,48]. Filling this gap in agriculture may therefore be a major challenge.

Another significant hurdle identified is data governance. Although most surveyed experts indicated their willingness to share their own data under certain conditions, many expressed concerns pertaining to data privacy, security, cybercrime and intellectual property protection. Without the increasingly centralized control of growing volumes of agricultural data by agribusiness and farm machinery manufacturers, there is rising concern that de facto monopolies may arise, facilitating rent-seeking by big agribusinesses which will marginalize vulnerable producers. This opens the opportunity for new institutions (both regulatory and proprietary) to align innovation to maximize positive impacts and limit adverse effects.

As most survey respondents are scientific experts, this research should be considered only as a first attempt that examines issues regarding the adoption and deployment of big data and AI technologies. Farmers and generators of primary data also need to be consulted about their opinions regarding data ownership and benefit sharing. Data from this article can be used by policy makers to begin building anticipatory guidelines that yield appropriate and scalable policy decisions to manage the diffusion of big data and AI into parts of the agricultural value-chain. With the emerging digitization of agriculture, fundamental research is required not only to address technical challenges but also to
find transformative solutions to the complex ethical, regulatory and social challenges of integrating automation and AI into our farm and food systems.

**Author Contributions:** R.L.: Conceptualization, data analysis, drafting. D.M.M.: conceptualization, drafting. S.J.S.: conceptualization, drafting and editing, funding acquisition. P.W.B.P.: conceptualization, reviewing and editing, funding acquisition. H.H.: conceptualization, reviewing and editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded through the Canada First Research Excellence Fund (CFREF) grant that established the Plant Phenotyping and Imaging Research Centre (P2IRC) project.

**Institutional Review Board Statement:** Our study (BEH 97) was exempt of ethics review by the Behavioral Ethics Board at the University of Saskatchewan on 26 March 2019. The requirements for exemption status are met as participants are not themselves the focus of the research per the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans, December 2014, Exemption Article 2.1.

**Informed Consent Statement:** The online survey presented participants with a standard consent statement describing the study, identifying the absence of known risks associated with participation, and a reminder that participation was voluntary, and responses would be anonymous and confidential. Upon expression of consent, participants were presented with the questionnaire (See Appendix A).

**Data Availability Statement:** The raw data supporting the conclusions of this manuscript are not publicly available because academic survey policy at the University of Saskatchewan states that all personal survey data will be protected and held confidential to ensure responder anonymity. Requests to access these datasets should be directed to Stuart Smyth at stuart.smyth@usask.ca.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

**Appendix A Questionnaire**

Consent form:

Dear Participant,

We appreciate your participation in our tenth survey centered on the integration of Big Data and Artificial Intelligence (AI) in agriculture and the food industry in a digital age. The questionnaire is part of a multiyear-year project on risk decision-making regarding new breeding technologies (NBTs). You have already completed at least one survey with us, and your responses have been invaluable in moving the project forward.

The survey project is investigating expert opinions regarding the opportunities and challenges surrounding the applications of NBTs in the agri-food industry. The lead researchers for this project are: Stuart Smyth (stuart.smyth@usask.ca, (306) 966-2929) and Peter Phillips (peter.phillips@usask.ca, (306) 966-4021). They can be contacted should you have any questions or comments. Any questions regarding your rights as a participant may be addressed to the University of Saskatchewan Research Ethics Office ethics.office@usask.ca; (306) 966-2975. Out of town participants may call toll free (888) 966-2975.

This survey is hosted by Voxco, a Canadian-owned and managed company whose data is securely stored in Canada. Please consider printing this page for your records.

There are no known risks to participating in this survey; however, as with any online activity the risk of breach of confidentiality is always possible.

In order to complete this survey, you may be required to answer certain questions; however, you are never obligated to respond and you may withdraw from the survey at any time by closing your internet browser.

By selecting next and completing this questionnaire, your free and informed consent is implied and indicates that you understand and accept the above conditions of participating in this study.

**Background:** Agriculture—like many other sectors—is becoming data-driven and data-enabled. Digital innovations (e.g., digital data collection, increased storage and com-
putational capacity, high-resolution of environmental and remotely sensed data) contribute to many agricultural improvements.

In this questionnaire, we would like to learn whether and how experts are coping with and adapting to the use of big data and artificial intelligence in the agri-food sector.

The following definition of Big Data will be available throughout this section of the survey:

Big Data refers to datasets so large and complex that they cannot be processed or analyzed with conventional software systems (Baker, 2017).

Big Data are characterized by such a high volume, velocity and variety to require specific technology and analytical methods for its transformation into value (De Mauro et al., 2016).

Q1. To your knowledge, do you or does your organization collect and/or analyze Big Data?
□ Yes—Collect only
□ Yes—Analyze only
□ Yes—Both collect and analyze
□ No—my organization does not collect or analyze Big Data

Q2A. Which type of Big Data (structured or unstructured)* do you or does your organization collect?

Structured data: includes generated reports, records, tables, survey data, sensor data, and customer data.
Unstructured data: includes emails, PDF texts, voice messages, web searches and social media (tweets, Facebook posts, videos, images, news stories).
□ Only structured data
□ Mostly structured data with some unstructured data
□ Both structured and unstructured data equally
□ Mostly unstructured data with some structured data
□ Only unstructured data

Q2B. Which type of Big Data (structured or unstructured)* do you or does your organization analyze?

Q3. Compared to peers/competitors in your field, how would you say that your organization compares on Big Data?
□ Very far ahead
□ Ahead
□ Neither ahead nor behind
□ Behind
□ Very far behind
□ I do not know

Q3A. From 0 to 100%, how confident are you in your response above? ____%

Q4. How likely or unlikely do you think your organization is to experience the following benefits from using Big Data?
| Benefits                                                                 | Very Unlikely | Unlikely | Neither Likely nor Unlikely | Likely | Very Likely | Do Not Know |
|-------------------------------------------------------------------------|---------------|----------|-----------------------------|--------|-------------|-------------|
| Better targeted clients/stakeholders                                   |               |          |                             |        |             |             |
| Better planning and forecasting                                       |               |          |                             |        |             |             |
| Better decision making                                                 |               |          |                             |        |             |             |
| Better risk management                                                 |               |          |                             |        |             |             |
| Better identification of root causes of problems                       |               |          |                             |        |             |             |
| Increased productivity                                                 |               |          |                             |        |             |             |
| Increased profit                                                       |               |          |                             |        |             |             |

Q5. Most Big Data applications require a combination of skills. In your opinion, what are the five most important factors for successful Big Data implementations?

To make your selections, please click on the factor you’d like to select and drag it into the box below so that the most important factor is at the top of the list followed by the second, third, fourth and fifth most important.

- Training
- Financial budget
- An organizational structure that supports multi-disciplinary projects
- A sound procedure for legal, ethical and reputational issues
- A clear company strategy
- Support by higher management
- Supporting systems and procedures
- Talent
- Other (please specify in box below)

Q6. Advanced analytic techniques that enable digital agriculture are multiple and varied. Among the types of technology below, which five do you feel are most important to the future of agriculture (Pick the top 5).

| Type of Technology                                                                 | Purpose and Benefits                                                                 |
|------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| Computational decision tools                                                       | Use data to develop recommendations for management and optimize multitudes of farm tasks |
| The cloud                                                                          | Provide efficient, inexpensive, and centralized data storage, computation, and communication to support farm management |
| Sensors                                                                            | Gather information on the functioning of equipment and farm resources to support management decisions |
| Robots                                                                             | Implement tasks with efficiency and minimal human labour                                |
| Digital communication tools (e.g., mobile, broadband)                              | Allow frequent, real-time communication between farm resources, workers, managers, and computational resources in support of management |
| Geo-locationing (e.g., GPS: Global positioning system)                              | Provide precise location of farm resources (field equipment, animals, etc.), often combined with measurements (yield, etc.), or used to steer equipment to locations |
| Geographic information systems                                                    | Use computerized mapping to aid inventory management and to make geographical crop input prescriptions (fertilizer, etc.) |
| Yield monitors                                                                     | Employ sensors and GPS on harvesters to continually measure harvest rate and make yield maps that allow for identification of local yield variability |
| Type of Technology                        | Purpose and Benefits                                                                                                                                 |
|------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------|
| Precision soil sampling                  | Sample soil at high spatial resolution (in zones) to detect and manage fertility patterns in fields                                               |
| Unmanned aerial systems (e.g., drones)   | Use small, readily deployed remote-control aerial vehicles to monitor farm resources using imaging UAS                                            |
| Spectral reflectance sensing (proximal and remote) | Measure light reflectance of soil or crop using satellite, airplane, or UAS, imaging, or field equipment–mounted sensors, to make determinations on soil patterns, crop, or animal performance, or on nutrient/pest problems |
| Auto-steering and guidance               | Reduce labour or fatigue with self-driving technology for farm equipment (including robots); can also precisely guide equipment in fields to enable highly accurate crop input placement and management |
| Variable rate technology                 | Allow continuous adjustment of application rates to precisely match localized crop needs in field areas with field applicators for crop inputs (chemicals, seed, etc.) |
| On-board computers                       | Collect and process field data with specialized computer hardware and software on tractors, harvesters, etc., often connected to sensors or controllers |
| Other (Please Fill in)                   |                                                                                                                                                     |

Q7. How does your organization develop Big Data applications? (Select all relevant responses)
- We develop our applications internally
- We contract with others to develop our applications
- We buy applications off the shelf
- Do not know

Q7A. Do you contract with any of the top leading Big Data companies (e.g., Apple, Microsoft, Amazon, Google, IBM, etc.)?
- Yes
- No
- Prefer not to say

Q8. What do you predict will happen to the number of Big Data specialists in your organization in the next three years?
- It will decrease (−20% or more)
- It will remain stable (±20%)
- It will increase (21–100%)
- It will more than double (>100%)
- Do not know

Q8A. From 0 to 100%, how confident are you in your response above? ____%

Q9. What are the top three potential barriers to the adoption of Big Data analytics in your work environment? (Select the top three)
- High overall cost of investment (e.g., big data storage equipment, training, software license)
- Lack of funding
- Shortage of skilled experts in Big Data analytics within the organization
- No identifiable end users/market
- Uncertainty about data privacy/ownership
- Cultural and communication barriers to the integration of new information technologies (IT) in the work environment
- Lack of strategic vision/interest by the management (i.e., not a priority)
- Other (Please specify)

The questions below focus on Artificial Intelligence (AI). The following definition of Artificial Intelligence will be available throughout this section of the survey:
Artificial Intelligence (AI) refers to the ability of computers to process data, with little to no human oversight, to make connections in complex relationships between datasets. AI technologies (e.g., machine learning, deep learning) can perform some tasks faster and with more accuracy than humans can.

Most AI examples that you hear about today—from chess-playing computers to self-driving cars—rely heavily on deep learning and natural language processing. *https://medium.com/@JoshMangus/big-data-is-not-artificial-intelligence-8f7e219f32a* (accessed on 28 March 2019).

Q10. How much contribution do you think Artificial Intelligence will make to agriculture (including primary production, processing and supply chains)?

□ No contribution at all
□ Little contribution
□ Moderate contribution
□ Large contribution
□ Substantial contribution
□ Do not know

Q10A. From 0 to 100%, how confident are you in your response above? ____%

Q11. How do you think that applying Artificial Intelligence to agriculture will affect the total number of jobs available (including primary production, processing and supply chains)?

□ It will decrease the number of jobs.
□ There will be no substantial change to the number of agricultural jobs
□ It will increase the number of jobs.
□ I do not know.

Q12. Overall, what impact do you think Artificial Intelligence will have on each of the following food sector segments?

| Category                          | Positive (Gain) | Neither Positive or Negative | Negative (Loss) |
|----------------------------------|-----------------|-------------------------------|-----------------|
| Plant breeding                   |                 |                               |                 |
| Agricultural machinery (involved in any aspect of agricultural production and processing) |                 |                               |                 |
| Logistics                        |                 |                               |                 |
| Finance                          |                 |                               |                 |
| Risk management                  |                 |                               |                 |
| End products                     |                 |                               |                 |
| Commercialization                |                 |                               |                 |
| Market information               |                 |                               |                 |

Q13. Assuming you own machine-generated data (e.g., sensor data, network data) within your organization, are there circumstances where you would be willing to share this machine data with others or make them publicly available?

□ Yes
□ No
□ Not applicable—my organization does not own machine-generated data

Q13B_1. Under which circumstances would you be willing to share your machine data with others (select all that apply)?

My data would be shared . . .
Q13B_2. Under which circumstances would you be willing to make your machine data publicly available (select all that apply)?

My data would be made publicly available...

- for free
- for professional benefit (partnership/collaboration, scientific merit)
- for monetary incentives
- for public benefit
- for transparency and re-use
- because data sharing is standard practice within my field
- for another reason (please specify)
- Unsure/No answer
- I would not be willing to do this

Q13C. Why wouldn’t you be willing to share your machine data with others (select all that apply)?

- Security concerns over the handling of sensitive or confidential data
- Risks of cybercrimes (identity theft)
- Concerns about intellectual property and scooping of ideas
- Inappropriate use of the shared data (Falsification, fabrication)
- Data sharing is prohibited by formal agreement in my work
- Potential lack of recognition/acknowledgment
- Lack of monetary incentives to data provision
- Other (Please specify)
- Unsure/No answer

Q14. In an increasingly integrated world, who do you trust to manage generated information?

|                          | I Do Not Trust at All | I Trust a Little | I Trust a lot | No Opinion |
|--------------------------|-----------------------|------------------|--------------|------------|
| Farmers                  |                       |                  |              |            |
| Agribusinesses           |                       |                  |              |            |
| Companies providing      |                       |                  |              |            |
| equipment                |                       |                  |              |            |
| Statistical bureau       |                       |                  |              |            |
| Consultancy agencies     |                       |                  |              |            |
| Banks/Financial institutions |                     |                  |              |            |
| Universities             |                       |                  |              |            |
| Government               |                       |                  |              |            |
| Other (Please specify)   |                       |                  |              |            |
Q15. A recent survey of companies that have adopted data driven decision-making shows those companies report 5–6% higher output and productivity than would be expected given their other investments and level of information technology usage. Do you think this result justifies investment in personnel and business adaptation in the digital agriculture?

☐ Yes
☐ No

Q15A. Why do you feel this way? Please briefly explain.

Finally, please tell us a bit about yourself . . .

Q16. Do you identify yourself as:

☐ Male
☐ Female
☐ Other
☐ Prefer not to say

Q17. Where do you currently reside?

☐ Africa
☐ Asia
☐ Europe
☐ Central & South America
☐ North America
☐ Oceania

Q18. Do you identify yourself as:

☐ A Life scientist (biologist, ecologist, etc)
☐ A social scientist (economist, lawyer, etc.)
☐ Other

Q19. Do you work for:

☐ Industry/private research institution
☐ Academic institution
☐ Government/Public research institutes
☐ Other

Q20. Do you have any feedback on this survey?

References

1. Manyika, J. Big Data: The Next Frontier for Innovation, Competition, and Productivity. Available online: http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation (accessed on 26 June 2019).
2. Griffin, T.W.; Shockley, J.M.; Mark, T.B. Economics of Precision Farming. In Precision Agriculture Basics; Shannon, D.K., Clay, D.E., Kitchen, N.R., Eds.; American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America, Inc.: Madison, WI, USA, 2018; pp. 221–230.
3. Miller, N.J.; Griffin, T.W.; Ciampitti, I.A.; Sharda, A. Farm adoption of embodied knowledge and information intensive precision agriculture technology bundles. Precis. Agric. 2019, 20, 348–361. [CrossRef]
4. Mulla, D.J. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. Biosyst. Eng. 2013, 114, 358–371. [CrossRef]
5. Giesler, S. Digitisation in Agriculture—From Precision Farming to Farming 4.0. 2018, Bioconomy BW. Available online: https://www.biooekonomie-bw.de/en/articles/dossiers/digitisation-in-agriculture-from-precision-farming-to-farming-40 (accessed on 12 December 2019).
6. CEMA-European Agricultural Machinery. Digital Farming: What Does It Really Mean? And What Is the Vision of Europe’s Farm Machinery Industry for Digital Farming? CEMA: Brussels, Belgium, 2017.
7. Brase, T. Precision Agriculture; Delmar Publishers Inc.: Salem, OR, USA, 2015; p. 288.
8. Magnin, C. How Big Data will Revolutionize the Global Food Chain. 2016. Available online: https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/how-big-data-will-revolutionize-the-global-food-chain (accessed on 1 March 2020).
9. van Es, H.; Woodward, J. Innovation in Agriculture and Food Systems in the Digital Age. In The Global Innovation Index 2017: Innovation Feeding the World; Dutta, S., Larvin, B., Wunsch-Vincent, S., Eds.; Cornell University: Ithaca, NY, USA; INSEAD: Fontainebleau, France; WIPO: Geneva, Switzerland, 2017; pp. 97–104.
43. Stodder, D. Best Practices Report | BI and Analytics in the Age of AI and Big Data. 2018, The Data Warehouse Institute. Available online: https://tdwi.org/articles/2018/12/21/adv-all-executive-summary-bpr-bi-analytics-age-of-ai-and-big-data.aspx (accessed on 25 March 2020).

44. Halper, F. Executive Summary | Data Science and Big Data: Enterprise Paths to Success; The Data Warehouse Institute: Washington, DC, USA, 2016.

45. Rehman, A.; Jingdong, L.; Khatoon, R.; Hussain, I. Modern Agricultural Technology Adoption: Its Importance, Role and Usage for the Improvement of Agriculture. Am. Eurasian J. Agric. Environ. Sci. 2016, 16, 284–288.

46. Relf-Eckstein, J.E.; Ballantyne, A.T.; Phillips, P.W.B. Farming Reimagined: A case study of autonomous farm equipment and creating an innovation opportunity space for broadacre smart farming. NJAS Wagening. J. Life Sci. 2019, 100307. [CrossRef]

47. OECD. OECD Employment Outlook 2019: The Future of Work; OECD: Paris, France, 2019; p. 345.

48. Tozer, P.R. Uncertainty and investment in precision agriculture—Is it worth the money? Agric. Syst. 2009, 100, 80–87. [CrossRef]