

Article

Reputational Risk Associated with Big Data Research and Development: An Interdisciplinary Perspective

Cara Stitzlein ^{1,*}, Simon Fielke ² , François Waldner ³ and Todd Sanderson ⁴

¹ CSIRO Data61, Hobart 7005, Australia

² CSIRO Land and Water, Dutton Park 4102, Australia; simon.fielke@csiro.au

³ European Commission Joint Research Centre, 21027 Ispra, Italy; Franz.Waldner@ec.europa.eu

⁴ Australian Centre for International Agricultural Research, Canberra 2601, Australia;

Todd.Sanderson@aciarc.gov.au

* Correspondence: Cara.Stitzlein@csiro.au

Abstract: Many private and public actors are incentivized by the promises of big data technologies: digital tools underpinned by capabilities like artificial intelligence and machine learning. While many shared value propositions exist regarding what these technologies afford, public-facing concerns related to individual privacy, algorithm fairness, and the access to insights requires attention if the widespread use and subsequent value of these technologies are to be fully realized. Drawing from perspectives of data science, social science and technology acceptance, we present an interdisciplinary analysis that links these concerns with traditional research and development (R&D) activities. We suggest a reframing of the public R&D ‘brand’ that responds to legitimate concerns related to data collection, development, and the implementation of big data technologies. We offer as a case study Australian agriculture, which is currently undergoing such digitalization, and where concerns have been raised by landholders and the research community. With seemingly limitless possibilities, an updated account of responsible R&D in an increasingly digitalized world may accelerate the ways in which we might realize the benefits of big data and mitigate harmful social and environmental costs.

Keywords: digital transformation; research and development; reputational risk; interdisciplinary science; big data; innovation systems; digital agriculture; responsible innovation; responsible prediction



Citation: Stitzlein, C.; Fielke, S.; Waldner, F.; Sanderson, T. Reputational Risk Associated with Big Data Research and Development: An Interdisciplinary Perspective. *Sustainability* **2021**, *13*, 9280. <https://doi.org/10.3390/su13169280>

Academic Editor: Marc A. Rosen

Received: 17 May 2021

Accepted: 10 August 2021

Published: 18 August 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

It is clear that the scarcity of both nonrenewable and renewable resources requires new approaches to provide the essential services for human health and sustainable societal development [1]. Part of the solution may lie in digital applications using ‘big data’ that help stakeholders prevent further degradation of ecosystems, manage finite resources, and track their transitions to sustainability targets (e.g., carbon neutrality). Big data technologies such as artificial intelligence and machine learning provide greater insights than were previously possible [2]. By combining large datasets with advanced modeling approaches, big data technology extracts patterns, performs time-series analyses, and makes predictions that assist problem solving in new ways. Affordances of data-driven logistics and forecasting models may help to alleviate global sustainability pressures like food security, climate resilience and efficient transportation [3]. Developments in natural capital accounting propose new knowledge systems related to ecological functions and open new markets [4]. There is little doubt that advanced data science innovations will enable the measurement and valuation of environmental assets alongside existing production systems and supply chains.

Solutions leveraging big data typify much of today’s work occurring within innovation systems, i.e., collections of technology, information and institutions that are transforming the world digitally [5]. Publicly funded research and development (R&D) organizations are important actors here, providing essential domain expertise and focus on leveraging

technology for the betterment of society. The dual interest of R&D firms to develop reliable technology and implement it for the global good is needed when considering that there is potential for both positive and negative consequences once technologies are introduced [6]. In the context of big data innovation specifically, there is evidence suggesting that the traditional role of R&D is changing [7]. As individual actors, businesses and supply chains adopt new forms of data-driven support, the public entities associated with these advancements must be prepared with clear placement and purpose in a digitally enabled innovation system. There are growing levels of concern that individual freedoms and securities are potentially compromised, and thus contribute to the public's salient distrust of the technologies making these advancements possible [8,9]. At a broader societal level, applications of big data technologies may introduce power imbalances and inequities across groups of people, and may reinforce concerns about 'big brother' [10]. When considering the intrinsic role of natural resource management in agricultural productivity and livelihoods, and the influence of R&D on the technological developments within this space, we believe it is time to reconsider how public R&D operates in relation to private and public interests.

In this paper, we revisit the components of public trust and fairness associated with the ways in which traditional R&D is perceived and managed [11,12]. In parallel, we consider the propositions for a "digitalization of everything" that use the continuous measurement and monitoring of systems with increasingly powerful forms of automation supported by advances in robotics, signal processing and computer vision [13,14]. We draw from technical and social disciplines to illustrate the disconnect between the value propositions of technology and the social contexts that they are likely to affect. In doing so, we suggest that this contribution can be integrated with other views of how to support the adaptive capacity of R&D for new types of business models and efforts of cross-industry innovation [15,16]. This paper contributes to the goal of innovation in a mature, scientifically robust and responsible manner.

2. Background

First, the asset of trust is used to characterize the 'brand' of publicly funded R&D. This is brought alongside research on public perceptions related to risk and technology acceptance. Then, a mapping of big data properties to decision complexity demonstrates the potential for big data solutions to support landscape-wide assessments. The combination of public-facing tensions and value propositions of the technology form the basis for rethinking the reputational risk associated with traditional R&D in the context of technological innovation.

2.1. The Brand of 'Trust' in R&D

Historically, publicly funded R&D organizations have played an important role in delivering scientific insights as a neutral party within multi-stakeholder projects [12,17]. In order to capture the overlap between the historic brand and new forms of technology, we draw from the work of Shin and Park [18], who identified three desirable but 'thorny' properties of algorithms—Fairness, Accountability and Transparency, or FAT. These properties are inspired by the traditional trusted brand, as represented in Figure 1. First, fair behaviors are performed impartially, without favoritism or bias. Second, being accountable is consistent with dependability, and is associated with responsible conduct. Third, the principle of transparency refers to actions which are considered open and honest, and are performed with integrity. When the three principles are practiced, a reputation of trust may be built, perceived as upholding the public good, and considered a foundation for equitable policy and practice. In this way, we propose trust as both an important element in the reputation of publicly funded research and as a quality that resonates as confidence in action [19].

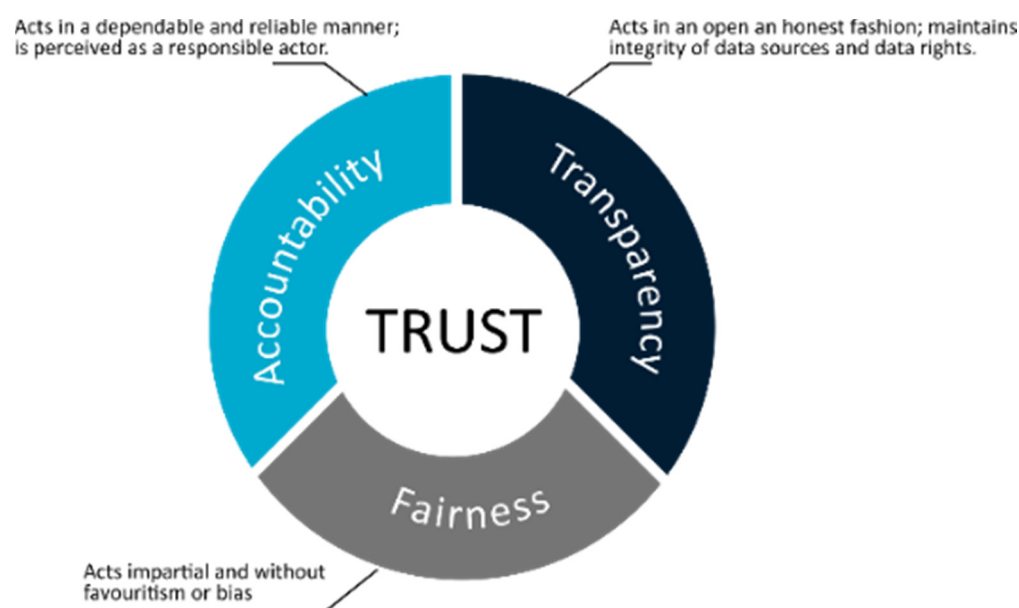


Figure 1. Notional account of ‘trust’ based on the qualities of fairness, accountability and transparency. Figure based on Shin and Park [18], modified by the authors.

2.2. Public Perceptions of Technology Risk and Science Innovation

Public perceptions affect whether an organization or an industry can maintain its social license to operate [20]. A body of supporting evidence based on behavioral research suggests that events once considered ‘extreme’ are becoming more frequently reported and common in public debate (e.g., cloning and synthetic biology) [21]. The relationship described by the behavioral research literature is a positive, linear relationship whereby public perceptions are affected by the degree of perceived uncertainty and dread risk associated with the technology [22]. The factor of dread risk refers to those risks perceived as uncontrollable, globally catastrophic, inequitable, and risky for future generations. The factor of uncertainty is based on observable risk, whether it has delayed effects or poses a new risk unknown to society. The failure to address or the active dismissal of the emotions and anxieties of the public in the face of scientific advancements that affect society can contribute to risks not being identified or mitigated appropriately [22]. The failure to acknowledge the varied interests and value systems of research partners is part of this responsibility to the public.

In the broad context of big data technologies, the negative perceptions correspond to privacy violations, a loss of autonomy, and general distrust of data-driven insights. Evidence presented in Lee [23] showed that managerial decisions performed by machine learning algorithms were perceived as less fair and less trustworthy when compared to human performance on the same task. The report of data anxieties [19] and negative emotional affect when judging algorithm performance [23] suggests that data-driven technologies are currently perceived as having a high technology risk.

The public perceptions of technology risk can be at odds with scientific innovation processes that push the boundaries of current knowledge and technological capability. Historically, there has been a societal dependency on institutions like R&D to solve issues related to sustainability, with the resulting solutions often being technocentric and underdeveloped at the intersection of science, policy and practice [24,25]. The challenge facing impact-focused R&D related to agriculture and natural resource management is to help reconcile the tensions between the perceived technology risk and business motives for innovation. Uncertain or potential dread risks would include programs of land use change with irreversible negative outcomes, or which would exacerbate disadvantaged populations (humans and otherwise), e.g., in the consideration of forest restoration projects [26], hydropower projects [27] and water market creation [28]. Turning now to a specific technol-

ogy central to digital transformation and open innovation systems—big data—we highlight linkages between the affordances of the technology and decision-making capabilities.

2.3. The Relationship between Big Data and Decisions at a Landscape Scale

Big data and the data that underpin these technologies are envisioned to support the decisions of how natural resources and land will be managed in the future. Big data is characterized by qualities of data consumption, e.g., its volume, velocity, variety and veracity [29,30]. Big data may also be characterized by the ways in which data are manipulated, through which large amounts of raw measurements are transformed into structured units using software code, machine learning algorithms and computational models.

The data–information–knowledge–wisdom (DIKW) hierarchy is commonly used to describe the transition from raw data to organized information sets to data-driven insights [29]. We apply the DIKW hierarchy to provide context for understanding the decision support possible from big data technologies. Figure 2 shows the association between the levels of the DIKW hierarchy and the types of decision support (modified from [29,30]).

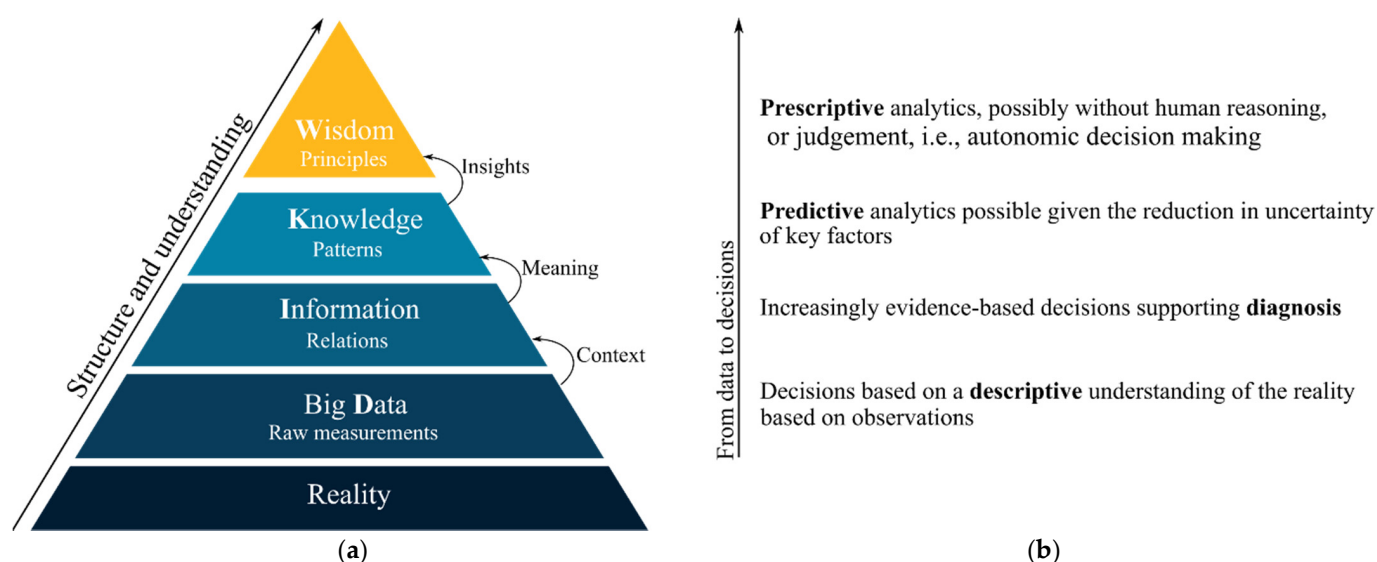


Figure 2. Characteristics of big data and decision support: (a) data–information–knowledge–wisdom hierarchy, modified from [29], and (b) the potential decision support offered, modified from [30].

The movement up the hierarchy begins with aggregated data that attribute meaning to information at the next level; then, insights are drawn at the level of knowledge, and so on. The levels of decision support follow a similar progression as the data-driven outputs reduce uncertainty and improve decision confidence about future states. Movement up the DIKW hierarchy (from raw data towards wisdom) reflects an increase in technical complexity, and with it the types of increasingly complex decisions that may be supported. When supported by higher-level insights (at the level of knowledge and wisdom), the decision maker can (in principle) evaluate options over longer timescales and in response to patterns that are not observable in conventional terms. The DIKW decision mapping in Figure 2 forms a basis for understanding the affordances motivating technology development and the ways in which digital innovations might mature over time [31,32].

With the potential to support decision making at a scale that was not previously possible, big data technologies may create new ways of operating within the world [31]. The propositions of the transformation of natural resource management with big data solutions have the potential to combine information and operations at a relatively small (farm-level) scale with large entities operating at a broader (landscape-level) scale. When ag-

gregated with datasets of land cover, historical productivity data and soil information, advanced modeling approaches provide data-driven outputs that support new ways of understanding the location, movement and significance of natural resources. The implications are that the technology can simultaneously probe the biological, social and economic aspects of any system, and as a result the technologies can be applied in ways to optimize decision making [32]. Because they offer insights into different temporal (past and future) and spatial (local and global) scales, there is increasing demand for these technologies from actors working in private, public and government agencies [33]. Notwithstanding, they might incentivize new types of data-driven digital economies, including green innovation [34].

In order to synthesize the technical, social, and economic perspectives presented, we present a case study of the ways in which the issues are being realized in agriculture. The domain application example illustrates the social context of how the technologies (and systems they support) reshape the realities of landholders, communities, and the physical landscape. We also look at how the risks within the domain application appear across the stages of technology development. This synthesis frames our discussion of how actors within R&D may anticipate reputational risk while at the same time supporting innovation and more responsible prediction.

3. Case Study

The digitalization of the agricultural sector demonstrates the affordances of big data solutions and the responsiveness needed to maintain public trust in R&D. Historically, agriculture has been transformed by technology and is currently going through what has been labeled the fourth revolution, culminating from a series of earlier technological disruptions: the introduction of animal power (Agriculture 1.0), the combustion engine (Agriculture 2.0), and precision agriculture (Agriculture 3.0) [25]. If realized, Agriculture 4.0 will transform the economic sustainability of agribusiness and farm productivity as it is known today.

For the optimization of farm productivity and profitability, a range of decisions are being targeted for big data solutions. Basic insights can inform land managers of potential problems, thus providing improved responsiveness to on-farm conditions [35–37]. When integrated with farm modeling systems, data-driven outputs may support seasonal decisions, including crop selection and other applications of precision agriculture [36]. Moving further up the DIKW hierarchy, predictive analytics support longer-term planning related to expansion opportunities and financial security. At a national and international scale, the predictive analytics afforded by digitalized agriculture might increase the food supply, increase logistic efficiency, and provide insights into food availability and quality [38–40]. Coordination around this knowledge could reduce food price volatility and increase the likelihood of meeting sustainable development goals [39]. The management of natural resources supported by big data technologies has the potential to assist with climate adaptation strategies and innovative agribusiness planning [40].

The immediate and aspirational application areas of big data to digital agriculture appear limitless and compelling. In 2018, it was estimated that Australian AgTech provides the digital infrastructure and insight to help earnings from agriculture almost double to \$100 billion as soon as 2030, making it a competitive and rapidly evolving sector [41]. There has been significant interest from actors across the supply chain in the value capture and exploitation of data captured from farms: applications will be picked up by actors operating in the field as well as outside the farm-gate. As a result, digital agricultural innovation systems are crowded with players (including startups and incubators), each hoping to bring a product to the market first [5]. There is little doubt that digital transformation will bring significant change to the people and practices operating within these systems [42].

3.1. Sources of Reputational Risk in Agriculture

Agriculture is a domain where public concerns are held by individual landholders and rural communities. Landholders may feel unprotected or unable to negotiate the informed consent process [43,44]. Therefore, when data collected in an R&D context are shared with parties unknown to them or without context, the trust between farmers and researchers may be threatened [45,46]. Social scientists have raised concerns that the existing consent arrangements are vague and may contribute to unfair benefit structures [45,47,48]. Relatedly, data ownership has become a contentious issue between equipment manufacturers, service providers and landholders. For example, companies can collect vast amounts of data from equipment being operated on individual farms. Companies consider themselves the owners of such data because without their equipment it would not be collected, although the farmers that pay for and operate these machines may have different views [47]. When the value associated with this data is restricted to companies and kept from farmers, traditional notions of consent are violated. There is the potential for this data to be used in products that contribute to the marginalization of farms in the same locations. Without mechanisms in place to protect farm(er)s, the technologies become threatening and are no longer perceived as a helpful innovation.

In order to maintain integrity as a trusted entity for on-farm actors and their communities, publicly funded R&D organizations must operate with the recognition that off-farm and private actors can leverage big data, exploit informational asymmetries, and increase their profit by marginalizing others. These diverging perceptions are tied to uneven power dynamics, unclear (and changing) motivations, and cultural clashes between agricultural (and broader societal) perceptions and startup communities [49,50]. Finally, while the proposition to connect smallholders into export markets is promising, this is not likely to be realized in a competitive data-driven digital economy [49]. Within agriculture where digitalization is quickly advancing, the public-facing concerns related to data ownership, integrity, and how big data solutions present a new form of competitive advantage are examples of the ways in which the trusted reputation of R&D may be at risk. With greater awareness of how issues of fairness, accountability and transparency are being addressed in agriculture, forthcoming R&D related to natural resource management can be better prepared to manage the issues arising within the public debate. In the next section, we summarize the ways in which reputational risk is currently being managed across the phases of technology development.

3.2. Reputational Risk across Phases of Technology Development

In the early activities of technology development, data are collected from various sources. Traditionally, a written consent agreement specifies the terms and conditions for the ways in which the data are collected from individuals, stored, and made available. Data collection and reuse is a legitimate handling of data in most standard agreements. The prevailing legal perspective in Australia, for example, prescribes that the acquisition of data requires basic consent from its owner, and that the data may be reused in any way, subject to the provisions of the Privacy Act 1988 [50]. Similar governance mechanisms are covered in the European Union's General Data Protection Regulation (GDPR) [51]. While an organization's reuse of the collected data may enjoy legal protection, it does not provide insulation from the reputational backlash of inadequately informed data contributors. The users of common data-as-payment services, relying on 'notice and choice' permissions, are rarely fully informed about the intended or potential reuse of their data and the implications for their privacy in exchange for 'free' access [52,53]. These conventional arrangements have led to questionable data reuse, such as the Facebook–Cambridge Analytica incident, in which the reputational damage was significant and may be long-lasting [52].

Technology developers have ways to prioritize assurances of individual privacy and security in the way technologies are designed, deployed and used [53–56]. Multiple technical solutions help to protect privacy, including anonymization techniques that replace

personal identifiers with random numbers. Data aggregation techniques can protect by blurring the location of a data source. A third solution includes federated systems that provide data access in a more tightly controlled fashion, such as OPAL (Open Algorithms for better decisions) [57]. Nowadays, differential privacy has become the de facto standard for preserving data privacy by mixing personal data with artificial white noise which theoretically prevents actors from identifying the details of the source data [58]. While anonymization, aggregation, federated systems and differential privacy represent tangible solutions for the protection of static data, they fail to completely anonymize data and establish transparency. There are instances where the reverse engineering of algorithms to their source data has been reported [59].

In the engineering of digital solutions, labeled data are used for model training and evaluation. While algorithms may eliminate types of human bias, they are subject to their own bias when the data are not adequately representative or contextually accurate. From minor embarrassment to more concrete negative backlash, the trust in organizations can be compromised if they are perceived to be developing or using a biased product [60]. Providing a record of the provenance of how insights are generated can demonstrate accuracy, but it can be difficult or impossible to explain how an algorithm reached a specific outcome. Issues of model transparency, i.e., how particular models work or how data-driven insights are reached, are being addressed by the Artificial Intelligence (AI) community using a concept of explainability that encompasses detailed model documentation, transparent “white-box” models, global models (finding out influential input variables), and local models (finding out reasons for the prediction of a particular outcome). There has been a shift towards developing models that allow predictions to be interpreted, and where relationships can be extracted and tested.

Issues related to transparency—in terms of promoting data collection and developing explainable algorithms—are present in the current debates in AI, specifically about how AI models manage the various sources of uncertainty. Model uncertainty might be related to noisy inputs or labels (aleatoric uncertainty, see [61], under review) or embedded in the model itself (epistemic uncertainty). As is consistent with public-facing concerns related to uncertainty in technology, there is unrealized potential in estimating and communicating uncertainty to users and developers. Accommodating for uncertainty in the design of tools could unlock new value propositions and improve usability. For instance, uncertainty estimates could lead to more principled decision making and allow models to (semi)automatically abstain on overly uncertain samples [62]. Alternatively, expressing the uncertainty could encourage a trade-off whereby humans use intelligent automation while remaining empowered to provide feedback [63]. Above all, communicating uncertainty effectively can provide a critical layer of transparency and trust, which is crucial for improved AI-assisted decision making. Therefore, models which provide a confidence level associated with the prediction based on the quality of the available data [64] as well as open-source toolboxes for uncertainty quantification [65] are good steps towards reconciling uncertainty.

There are social and behavioral factors during development that affect whether data-driven insights are considered ethically fair and unbiased when considered in the broader context of implementation [66]. There are concerns that biases within these products can (and do) exacerbate existing power asymmetries [67]. Originating from legal analysis to determine unintentional discrimination, disparate impact analysis is gaining momentum as a practical tool to discuss and handle observational fairness, i.e., how model predictions affect different groups of people [68]. Ethical AI has begun to attract attention from governmental bodies such as the European Commission [69], and commercial solutions are being brought to the market. However, designing algorithms with sensitivity to what constitutes public fairness remains a challenge.

In the context of deploying big data technologies at a scale of landscapes and shared natural resources, publicly funded R&D organizations must work with commercial, private and government entities. As these entities have varying motives, the role of R&D as a

neutral stakeholder becomes increasingly important [12]. The social sciences have long highlighted the rural–urban divide through which communities are marginalized due to a lack of literacy, access and the skills required to maximize their use of data and the Internet of Things [70]. Armed with this knowledge of undesirable social consequences, actors within R&D must carefully manage their association with external entities throughout the technology development process, recognizing the value systems and operational principles of their collaborators and clients. Principles of responsible innovation are becoming a part of the conversation of how to proceed responsibly and ethically [71]. The responsible innovation framework consists of four principles: anticipation, inclusion, reflexivity and responsiveness [72]. Participation in these behaviors is required if the transformational outcomes being sought in individual businesses, the private sector and the public sector are to be achieved [73].

4. Recommendations on Reframing the Role of R&D

The digital innovations and technologies that utilize AI systems powered by big data constitute a new normal. When such powerful technologies are introduced into our supply chains, management systems, organizations and societies, they are bound to have social and economic implications. We live in a time of rapid technological change and digital connectedness. It is important to understand how we are influenced by the possibilities of big data and analytical tools and techniques [74]. The potentially transformative nature of technological developments in big data calls for attention to the broader sociotechnical context: “innovation is equally about people as it is about products, processes, and technical systems” ([75], p. 2). Herein lies the crux of why the modus operandi of R&D organizations need updating: we need to acknowledge that in the new digital normal, there exist independent value exchange networks where actors are using digital technologies in new ways to create profit (for themselves) and mitigate business risk (for themselves and their shareholders, as well as society). These networks are, by their nature, different from the objectives of traditional publicly funded R&D focused on societal benefits and, at times, creating knowledge for its own sake.

Some implications of technological innovation may be controversial, and thus public R&D must reconcile the possible repercussions from participating in their development. We have argued that R&D investment may, at times, contribute to public harm and a weakening of trust in the institution of research. We came to this conclusion after confronting the issues arising from multiple research disciplines, recognizing that the prism of a single discipline is too narrow to cover the complex and multifaceted realities of digital transformation and its effects across the data value chain. It is through interdisciplinarity that sociotechnical issues can help identify the “temporalities of data” ([19] p. 2). Synergies resulting from working across disciplines enable dialogue about how we might update our ways of working within the digital realm, with one another in R&D processes, and in partnership with industry and the private sector [74,76].

The recognition that digital transformation includes aspects of both technical and social change is helping to promote principles of responsible innovation [76] and other transitions towards active social learning in research [77]. R&D managers and leaders can improve their standard operations by using support such as the Risk Mitigation Checklists published by Ethical OS [78], the rules proposed by Zook et al. [79], and the guidance proposed for the public sector by the Alan Turing Institute [80]. These provide actionable ways of anticipating sources of technology mistrust, reinforce perceptions of digital responsibility, and link to business model development and adaptation [75]. Technological advances have demonstrated the ways in which improved security measures reduce the likelihood of privacy breaches. Engineering solutions that anonymize data and control data access are becoming more widely available. The recognition of public concerns in the development of National Standards [81] further reinforces that it is possible to proactively articulate the ways in which technology risk can be identified. However, these measures will take time to both initiate and put into practice. Basic literacy about the

technology and these improved practices is just emerging in the public debate. Emerging from the social and legal perspectives is the fact that the reputation associated with the ways in which digital systems are handled requires greater transparency about data collection, reuse, consent, and custodianship. These terms need to capture both the current state and possible future technologies built to mine this source data.

The value propositions of big data rely on effective data sharing and reciprocity. In order to promote the sharing of farm data for issues related to natural resources, we recommend making those propositions more tangible, and making insights related to natural resources more accessible in order to encourage that reciprocity. Concerning the case study—big data in agriculture—guidance is becoming available through voluntary codes of practice such as the Farm Data Code published by Australia's National Farmers Federation [82]. Created in consultation with industry, the code emphasizes the ways in which both service providers and landholders can increase the mutual transparency and quality of data (sharing). The simple act of checking a box when signing up for a service may disguise the realities of a data agreement that is fundamentally more complex. The engagement of Robinson and colleagues with an Indigenous population demonstrated another example of how principles can be linked to deliberate action [83]. We consider these examples as a starting point for the questions that future research will need to address.

The role of digital technologies as drivers and facilitators of economic exchange will only increase, and will in turn require adaptive governance processes [84]. The danger of developing technologies that put groups of people at a disadvantage can be better responded to via research calling for greater attention to the social dimensions of digital innovations in and for regions [85–88]. There is some evidence that responsible innovation principles may encourage fairer and more equitable technology use [25,45,89,90]. Putting principles into practice is less straightforward. In order to uphold the traditional notion of 'research for the public good', R&D's role in innovation may need to include an explicit reorientation to what digital technologies mean and afford to different groups of people and stakeholders [17]. This includes clearly articulating a position related to data—as both a mechanism to maintain organizational value (e.g., intellectual property rights), for research transparency, and for the greater good [91]. Such considered transparency may help reduce perceptions of vulnerability and violation [92]. Further responsiveness might be achieved by data custodianship taking shape with new dedicated parties or organizations—in agriculture, independent data brokers might be required, rather than equipment, fertilizer and seed company representatives—in order to allow for the trustworthy collection of the farmers' data and to provide controlled access to aggregated datasets to all actors.

5. Conclusions

The practice of operating in a scientifically informed and neutral manner—as publicly funded R&D has historically been perceived by many to do—needs updating in order to manage a more equitable exchange of big data products and services if the lofty aims of industrial digital transformation are to be realized. The fabric of legal policies and procedures that once managed reputational risk are no longer sufficient. We know that the existing privacy measures are poorly understood and not necessarily consistent with an image of fair and transparent business activity. The negative public backlash to these arrangements requires public R&D and other trusted brokers to reorient their traditional role as data custodians, technology developers and public advocates when their engagement in big data activities can be harnessed by private interests. This balance includes responding to the ways in which people perceive risks around data sharing and the information created by powerful analytical tools to engender insights.

After presenting these interdisciplinary perspectives regarding digital transformation through big data (i.e., data science, social science and design), we assert that working in such a broad (and potentially transdisciplinary, where appropriate) manner is an opportunity to better build shared, yet still ambitious, goals for public R&D. Interdisciplinarity is a way of working where researchers are more likely to be aware of each other's activities,

can build and share similar ways of identifying problems, and are better able to co-develop appropriate interventions. In an age of digital transformations, the representation of public interests will require an integrated approach that includes interdisciplinary perspectives, renewed policies and standards of best practice, and socializing advancements in protections of privacy and algorithm bias.

Big data products are becoming part of the ubiquitous fabric of our modern world. In this paper, we sought to demystify some of the technical and social data dangers of technology research and development in the context of natural resource management, specifically reputational risk and issues of trust and privacy, particularly in agriculture. Our insights into these urgent and future-oriented issues remain valid in other application domains, particularly those that are future-focused and grand in scale.

Author Contributions: All authors contributed to the writing of the manuscript. Conceptualization, C.S.; Writing—original draft, C.S., S.F., F.W. and T.S.; Writing—review & editing, C.S., S.F., F.W. and T.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the CSIRO Digiscape Future Science Platform (Australia).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Song, M.; Cen, L.; Zheng, Z.; Fisher, R.; Liang, X.; Wang, Y.; Huisingh, D. Improving Natural Resource Management and Human Health to Ensure Sustainable Societal Development Based upon Insights Gained from Working within ‘Big Data Environments’. *J. Clean. Prod.* **2015**, *94*, 1–4. [\[CrossRef\]](#)
2. De Mauro, A.; Greco, M.; Grimaldi, M. A Formal Definition of Big Data Based on Its Essential Features. *Libr. Rev.* **2016**, *65*, 122–135. [\[CrossRef\]](#)
3. Diebold, F.X. On the Origin(s) and Development of the Term ‘Big Data’. PIER Working Paper 12–037, 2012. Available online: <https://ssrn.com/abstract=2152421> (accessed on 10 August 2021).
4. Izac, A.-M.N.; Sanchez, P.A. Towards a Natural Resource Management Paradigm for International Agriculture: The Example of Agroforestry Research. *Agric. Syst.* **2001**, *69*, 5–25. [\[CrossRef\]](#)
5. Fielke, S.J.; Garrard, R.; Jakku, E.; Fleming, A.; Wiseman, L.; Taylor, B.M. Conceptualising the DAIS: Implications of the ‘Digitalisation of Agricultural Innovation Systems’ on Technology and Policy at Multiple Levels. *NJAS Wagening. J. Life Sci.* **2019**, *90*, 100296. [\[CrossRef\]](#)
6. Hellström, T. Systemic Innovation and Risk: Technology Assessment and the Challenge of Responsible Innovation. *Technol. Soc.* **2003**, *25*, 369–384. [\[CrossRef\]](#)
7. Blackburn, M.; Alexander, J.; Legan, J.D.; Klabjan, D. Big Data and the Future of R&D Management. *Res. Manag.* **2017**, *60*, 43–51. [\[CrossRef\]](#)
8. Stilgoe, J. Machine Learning, Social Learning and the Governance of Self-Driving Cars. *Soc. Stud. Sci.* **2018**, *48*, 25–56. [\[CrossRef\]](#)
9. Carolan, M. Agro-Digital Governance and Life Itself: Food Politics at the Intersection of Code and Affect. *Sociol. Rural.* **2017**, *57*, 816–835. [\[CrossRef\]](#)
10. Coble, K.H.; Mishra, A.K.; Ferrell, S.; Griffin, T. Big Data in Agriculture: A Challenge for the Future. *Appl. Econ. Perspect. Policy* **2018**, *40*, 79–96. [\[CrossRef\]](#)
11. Pielke, R.A., Jr. *The Honest Broker: Making Sense of Science in Policy and Politics*; Cambridge University Press: Cambridge, UK, 2007.
12. Rantala, L.; Sarkki, S.; Karjalainen, T.P.; Rossi, P.M. How to Earn the Status of Honest Broker? Scientists’ Roles Facilitating the Political Water Supply Decision-Making Process. *Soc. Nat. Resour.* **2017**, *30*, 1288–1298. [\[CrossRef\]](#)
13. Fielke, S.; Taylor, B.; Jakku, E. Digitalisation of Agricultural Knowledge and Advice Networks: A State-of-the-Art Review. *Agric. Syst.* **2020**, *180*, 102763. [\[CrossRef\]](#)
14. Leviäkangas, P. Digitalisation of Finland’s Transport Sector. *Technol. Soc.* **2016**, *47*, 1–15. [\[CrossRef\]](#)
15. Enkel, E.; Gassmann, O. Creative Imitation: Exploring the Case of Cross-Industry Innovation. *R D Manag.* **2010**, *40*, 256–270. [\[CrossRef\]](#)
16. Moellers, T.; Visini, C.; Haldimann, M. Complementing Open Innovation in Multi-Business Firms: Practices for Promoting Knowledge Flows across Internal Units. *R D Manag.* **2020**, *50*, 96–115. [\[CrossRef\]](#)
17. Elias, A.A. Analysing the Stakes of Stakeholders in Research and Development Project Management: A Systems Approach. *R D Manag.* **2016**, *46*, 749–760. [\[CrossRef\]](#)

18. Shin, D.; Park, Y.J. Role of Fairness, Accountability, and Transparency in Algorithmic Affordance. *Comput. Hum. Behav.* **2019**, *98*, 277–284. [\[CrossRef\]](#)
19. Pink, S.; Lanzeni, D.; Horst, H. Data Anxieties: Finding Trust in Everyday Digital Mess. *Big Data Soc.* **2018**, *5*, 2053951718756685. [\[CrossRef\]](#)
20. Parsons, R.; Lacey, J.; Moffat, K. Maintaining Legitimacy of a Contested Practice: How the Minerals Industry Understands Its “Social Licence to Operate”. *Resour. Policy* **2014**, *41*, 83–90. [\[CrossRef\]](#)
21. Slovic, P.; Weber, E.U. Perception of Risk Posed by Extreme Events. In Proceedings of the Risk Management strategies in an Uncertain World, Palisades, NY, USA, 12–13 April 2002.
22. Slovic, P. Perception of Risk. *Science* **1987**, *236*, 280–285. [\[CrossRef\]](#)
23. Lee, M.K. Understanding Perception of Algorithmic Decisions: Fairness, Trust, and Emotion in Response to Algorithmic Management. *Big Data Soc.* **2018**, *5*. [\[CrossRef\]](#)
24. Leith, P.; O’Toole, K.; Haward, M.; Coffey, B. *Enhancing Science Impact*; CSIRO Publishing: Sydney, Australia, 2019. [\[CrossRef\]](#)
25. Rose, D.C.; Chilvers, J. Agriculture 4.0: Broadening Responsible Innovation in an Era of Smart Farming. *Front. Sustain. Food Syst.* **2018**, *2*, 87. [\[CrossRef\]](#)
26. Calder, I.R. Forests and Hydrological Services: Reconciling Public and Science Perceptions. *Land Use Water Resour. Res.* **2002**, *2*, 1–12.
27. Mayeda, A.M.; Boyd, A.D. Factors Influencing Public Perceptions of Hydropower Projects: A Systematic Literature Review. *Renew. Sustain. Energy Rev.* **2020**, *121*, 109713. [\[CrossRef\]](#)
28. Keenan, S.P.; Krannich, R.S.; Walker, M.S. Public Perceptions of Water Transfers and Markets: Describing Differences in Water Use Communities. *Soc. Nat. Resour.* **1999**, *12*, 279–292. [\[CrossRef\]](#)
29. Rowley, J. The Wisdom Hierarchy: Representations of the DIKW Hierarchy. *J. Inf. Sci.* **2007**, *33*, 163–180. [\[CrossRef\]](#)
30. Lokers, R.; Knapen, R.; Janssen, S.; van Randen, Y.; Jansen, J. Analysis of Big Data Technologies for Use in Agro-Environmental Science. *Environ. Model. Softw.* **2016**, *84*, 494–504. [\[CrossRef\]](#)
31. Klerkx, L.; Rose, D. Dealing with the Game-Changing Technologies of Agriculture 4.0: How Do We Manage Diversity and Responsibility in Food System Transition Pathways? *Glob. Food Secur.* **2020**, *2*, 100347. [\[CrossRef\]](#)
32. Sanderson, T.; Reeson, A.; Box, P. *Cultivating Trust: Towards an Australian Agricultural Data Market*; CSIRO Publishing: Sydney, Australia, 2017. [\[CrossRef\]](#)
33. Löfgren, K.; Webster, C.W.R. The Value of Big Data in Government: The Case of ‘Smart Cities’. *Big Data Soc.* **2020**, *7*, 2053951720912775. [\[CrossRef\]](#)
34. Schiederig, T.; Tietze, F.; Herstatt, C. Green Innovation in Technology and Innovation Management—An Exploratory Literature Review. *R D Manag.* **2012**, *42*, 180–192. [\[CrossRef\]](#)
35. Bramley, R.G.V.; Ouzman, J. Farmer Attitudes to the Use of Sensors and Automation in Fertilizer Decision-Making: Nitrogen Fertilization in the Australian Grains Sector. *Precis. Agric.* **2019**, *20*, 157–175. [\[CrossRef\]](#)
36. Miles, C. The Combine Will Tell the Truth: On Precision Agriculture and Algorithmic Rationality. *Big Data Soc.* **2019**, *6*, 2053951719849444. [\[CrossRef\]](#)
37. Chen, M.; Mao, S.; Liu, Y. Big Data: A Survey. *Mob. Netw. Appl.* **2014**, *19*, 171–209. [\[CrossRef\]](#)
38. Gilpin, L. How Big Data Is Going to Help Feed Nine Billion People by 2050. TechRepublic. 2015. Available online: <https://fii.institute/2014/11/11/how-big-data-is-going-to-help-feed-nine-billion-people-by-2050/> (accessed on 10 August 2021).
39. Antwi-Agyei, P.; Dougill, A.J.; Agyekum, T.P.; Stringer, L.C. Alignment between Nationally Determined Contributions and the Sustainable Development Goals for West Africa. *Clim. Policy* **2018**, *18*, 1296–1312. [\[CrossRef\]](#)
40. Nyasimi, M.; Kimeli, P.; Sayula, G.; Radeny, M.; Kinyangi, J.; Mungai, C. Adoption and Dissemination Pathways for Climate-Smart Agriculture Technologies and Practices for Climate-Resilient Livelihoods in Lushoto, Northeast Tanzania. *Climate* **2017**, *5*, 63. [\[CrossRef\]](#)
41. KPMG. Talking 2030: Growing Agriculture into a \$100 Billion Industry. 2018. Available online: <https://home.kpmg/au/en/home/insights/2018/03/talking-2030-growing-australian-agriculture-industry.html> (accessed on 10 August 2021).
42. Duncan, R.; Robson-Williams, M.; Nicholas, G.; Turner, J.A.; Smith, R.; Diprose, D. Transformation Is “Experienced, Not Delivered”: Insights from Grounding the Discourse in Practice to Inform Policy and Theory. *Sustainability* **2018**, *10*, 3177. [\[CrossRef\]](#)
43. Jakku, E.; Taylor, B.; Fleming, A.; Mason, C.; Fielke, S.; Sounness, C.; Thorburn, P. “If They Don’t Tell Us What They Do with It, Why Would We Trust Them?” Trust, Transparency and Benefit-Sharing in Smart Farming. *NJAS Wagening. J. Life Sci.* **2019**, *90*, 100285. [\[CrossRef\]](#)
44. Wiseman, L.; Sanderson, J.; Zhang, A.; Jakku, E. Farmers and Their Data: An Examination of Farmers’ Reluctance to Share Their Data through the Lens of the Laws Impacting Smart Farming. *NJAS Wagening. J. Life Sci.* **2019**, *90*, 100301. [\[CrossRef\]](#)
45. Regan, Á. ‘Smart Farming’ in Ireland: A Risk Perception Study with Key Governance Actors. *NJAS Wagening. J. Life Sci.* **2019**, *90*, 100292. [\[CrossRef\]](#)
46. Rotz, S.; Duncan, E.; Small, M.; Botschner, J.; Dara, R.; Mosby, I.; Reed, M.; Fraser, E.D.G. The Politics of Digital Agricultural Technologies: A Preliminary Review. *Sociol. Rural.* **2019**, *59*, 203–229. [\[CrossRef\]](#)
47. Carolan, M. ‘Smart’ Farming Techniques as Political Ontology: Access, Sovereignty and the Performance of Neoliberal and Not-So-Neoliberal Worlds. *Sociol. Rural.* **2018**, *58*, 745–764. [\[CrossRef\]](#)

48. Berthet, E.T.; Hickey, G.M.; Klerkx, L. Opening Design and Innovation Processes in Agriculture: Insights from Design and Management Sciences and Future Directions. *Agric. Syst.* **2018**, *165*, 111–115. [CrossRef]
49. Glover, D.; Sumberg, J.; Ton, G.; Andersson, J.; Badstue, L. Rethinking Technological Change in Smallholder Agriculture. *Outlook Agric.* **2019**, *48*, 169–180. [CrossRef]
50. Australia *Privacy Act 1988*, Federal Register of Legislation; Attorney-General's Department. 2017. Available online: <https://www.legislation.gov.au/Details/C2016C00979> (accessed on 10 August 2021).
51. Otto, M. Regulation (EU) 2016/679 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation—GDPR). *Int. Eur. Labour Law* **2018**, *2014*, 958–981. [CrossRef]
52. Democracy, Data and Dirty Tricks, 2018. Four Corners Episode. Available online: <http://www.abc.net.au/4corners/democracy-data-and-dirty-tricks-cambridge/9642090> (accessed on 10 August 2021).
53. Lesser, A. Big Data and Big Agriculture. 2014. Available online: <https://gigaom.com/report/big-data-and-big-agriculture/> (accessed on 10 August 2021).
54. Sonka, S. Big Data and the Ag Sector: More than Lots of Numbers. *Int. Food Agribus. Manag. Rev.* **2014**, *17*, 163351.
55. Orts, E.; Spigonardo, J. *Sustainability in the Age of Big Data*; IGEL/Wharton, University of Pennsylvania: Philadelphia, PA, USA, 2014. Available online: <http://d1c25a6gwz7q5e.cloudfront.net/reports/2014-09-12-Sustainability-in-the-Age-of-Big-Data.pdf> (accessed on 10 August 2021).
56. Darnell, R.; Robertson, M.; Brown, J.; Moore, A.; Barry, S.; Bramley, R.; Grundy, M.; George, A. The Current and Future State of Australian Agricultural Data. *Farm Policy J.* **2018**, *15*, 41–49.
57. Open Algorithms for Better Decisions. Available online: www.opalproject.org/about-opal (accessed on 20 June 2021).
58. Calude, C.S.; Dinneen, M.J.; Păun, G.; Pérez Jiménez, M.J.; Rozenberg, G. Lecture Notes in Computer Science: Preface. In *Lecture Notes in Computer Science*; Springer: London, UK, 2005.
59. De Montjoye, Y.A.; Hidalgo, C.A.; Verleysen, M.; Blondel, V.D. Unique in the Crowd: The Privacy Bounds of Human Mobility. *Sci. Rep.* **2013**, *3*, 1376. [CrossRef] [PubMed]
60. Green, B.P. Ethical Reflections on Artificial Intelligence. *Sci. Fides* **2018**, *6*, 9–31. [CrossRef]
61. Song, H.; Kim, M.; Park, D.; Shin, Y.; Lee, J.-G. Learning from Noisy Labels with Deep Neural Networks: A Survey. *arXiv* **2020**, arXiv:2007.08199.
62. Kompa, B.; Snoek, J.; Beam, A.L. Second Opinion Needed: Communicating Uncertainty in Medical Machine Learning. *Npj. Digit. Med.* **2021**, *4*, 4. [CrossRef]
63. Amershi, S.; Cakmak, M.; Knox, W.B.; Kulesza, T. Power to the People: The Role of Humans in Interactive Machine Learning. *Artif. Intell. Mag.* **2014**, *35*, 105–120. [CrossRef]
64. Kendall, A.; Badrinarayanan, V.; Cipolla, R. Bayesian SegNet: Model Uncertainty in Deep Convolutional Encoder-Decoder Architectures for Scene Understanding. 2017. Available online: <http://www.bmva.org/bmvc/2017/papers/paper057/paper057.pdf> (accessed on 10 August 2021).
65. Ghosh, S.; Liao, Q.V.; Ramamurthy, K.N.; Navratil, J.; Sattigeri, P.; Varshney, K.R.; Zhang, Y. Uncertainty Quantification 360: A Holistic Toolkit for Quantifying and Communicating the Uncertainty of AI. *arXiv* **2021**, arXiv:2106.01410.
66. Crawford, K. Can an Algorithm Be Agonistic? Ten Scenes from Life in Calculated Publics. *Sci. Technol. Hum. Values* **2015**, *41*, 77–92. [CrossRef]
67. Carbonell, I.M. The Ethics of Big Data in Big Agriculture. *Internet Policy Rev.* **2016**, *5*, 1–13. [CrossRef]
68. Barocas, S.; Selbst, A.D. Big Data's Disparate Impact. *Calif. Law Rev.* **2016**, *104*, 671. [CrossRef]
69. Veale, M.; Van Kleek, M.; Binns, R. Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making. *Conf. Hum. Factors Comput. Syst. Proc.* **2018**, *2018*, 440. [CrossRef]
70. Wolf, S.A.; Buttel, F.H. The Political Economy of Precision Farming. *Am. J. Agric. Econ.* **1996**, *78*, 1269–1274. [CrossRef]
71. Eastwood, C.; Klerkx, L.; Ayre, M.; Dela Rue, B. Managing Socio-Ethical Challenges in the Development of Smart Farming: From a Fragmented to a Comprehensive Approach for Responsible Research and Innovation. *J. Agric. Environ. Ethics* **2019**, *32*, 741–768. [CrossRef]
72. Owen, R.; Bessant, J.; Heintz, M. *Responsible Innovation*; John Wiley & Sons, Ltd.: West Sussex, UK, 2013. [CrossRef]
73. Gurzawska, A.; Mäkinen, M.; Brey, P. Implementation of Responsible Research and Innovation (RRI) Practices in Industry: Providing the Right Incentives. *Sustainability* **2017**, *9*, 1759. [CrossRef]
74. Lioutas, E.D.; Charatsari, C.; La Rocca, G.; De Rosa, M. Key Questions on the Use of Big Data in Farming: An Activity Theory Approach. *NJAS Wagening. J. Life Sci.* **2019**, *90*, 100297. [CrossRef]
75. Gorissen, L.; Vrancken, K.; Manshoven, S. Transition Thinking and Business Model Innovation-towards a Transformative Business Model and New Role for the Reuse Centers of Limburg, Belgium. *Sustainability* **2016**, *8*, 112. [CrossRef]
76. Owen, R.; Stilgoe, J.; Macnaghten, P.; Gorman, M.; Fisher, E.; Guston, D. A Framework for Responsible Innovation. *Responsible Innov. Manag. Responsible Emerg. Sci. Innov. Soc.* **2013**, *42*, 27–50. [CrossRef]
77. Beers, P.J.; Hermans, F.; Veldkamp, T.; Hinssen, J. Social Learning inside and Outside Transition Projects: Playing Free Jazz for a Heavy Metal Audience. *NJAS Wagening. J. Life Sci.* **2014**, *69*, 5–13. [CrossRef]
78. OS, E. Risk Mitigation Checklist. 2018. Available online: https://ethicalos.org/wp-content/uploads/2018/08/EthicalOS_Check-List_080618.pdf (accessed on 10 August 2021).

79. Zook, M.; Barocas, S.; Boyd, D.; Crawford, K.; Keller, E.; Gangadharan, S.P.; Goodman, A.; Hollander, R.; Koenig, B.A.; Metcalf, J.; et al. Ten Simple Rules for Responsible Big Data Research. *PLoS Comput. Biol.* **2017**, *13*, e1005399. [[CrossRef](#)]
80. Leslie, D. Understanding Artificial Intelligence Ethics and Safety: A Guide for the Responsible Design and Implementation of AI Systems in the Public Sector. *SSRN* **2019**. [[CrossRef](#)]
81. Standards Australia. An Artificial Intelligence Standards Roadmap: Making Australia's Voice Heard. 2020. Available online: https://www.standards.org.au/getmedia/ede81912-55a2-4d8e-849f-9844993c3b9d/R_1515-An-Artificial-Intelligence-Standards-Roadmap-soft.aspx (accessed on 10 August 2021).
82. Farm Data Code (Edition 1). National Farmers' Federation 2020. Available online: https://nff.org.au/wp-content/uploads/2020/02/Farm_Data_Code_Edition_1_WEB_FINAL.pdf (accessed on 10 August 2021).
83. Robinson, C.J.; Kong, T.; Coates, R.; Watson, I.; Stokes, C.; Pert, P.; McConnell, A.; Chen, C. Caring for Indigenous Data to Evaluate the Benefits of Indigenous Environmental Programs. *Environ. Manag.* **2021**, *68*, 160–169. [[CrossRef](#)] [[PubMed](#)]
84. Linkov, I.; Trump, B.D.; Poinssatte-Jones, K.; Florin, M.V. Governance Strategies for a Sustainable Digital World. *Sustainability* **2018**, *10*, 440. [[CrossRef](#)]
85. Bronson, K.; Knezevic, I. Look Twice at the Digital Agricultural Revolution. *Policy Options* **2017**. Available online: <https://policyoptions.irpp.org/magazines/september-2017/look-twice-at-the-digital-agricultural-revolution/> (accessed on 10 August 2021).
86. Bronson, K. Looking through a Responsible Innovation Lens at Uneven Engagements with Digital Farming. *NJAS Wagening. J. Life Sci.* **2019**, *90*, 100294. [[CrossRef](#)]
87. Pant, L.P.; Hambly Odame, H. Broadband for a Sustainable Digital Future of Rural Communities: A Reflexive Interactive Assessment. *J. Rural Stud.* **2017**, *54*, 435–450. [[CrossRef](#)]
88. Roberts, E.; Anderson, B.A.; Skerratt, S.; Farrington, J. A Review of the Rural-Digital Policy Agenda from a Community Resilience Perspective. *J. Rural Stud.* **2017**, *54*, 372–385. [[CrossRef](#)]
89. Klerkx, L.; Jakku, E.; Labarthe, P. A Review of Social Science on Digital Agriculture, Smart Farming and Agriculture 4.0: New Contributions and a Future Research Agenda. *NJAS Wagening. J. Life Sci.* **2019**, *90*, 100315. [[CrossRef](#)]
90. Gremmen, B.; Blok, V.; Bovenkerk, B. Responsible Innovation for Life: Five Challenges Agriculture Offers for Responsible Innovation in Agriculture and Food, and the Necessity of an Ethics of Innovation. *J. Agric. Environ. Ethics* **2019**, *32*, 673–679. [[CrossRef](#)]
91. Wiseman, L.; Sanderson, J.; Robb, L. Rethinking Ag Data Ownership. *Farm Policy J.* **2018**, *15*, 71–77.
92. Blundell-Wignall, A.; Wehinger, G.; Slovik, P. The Elephant in the Room: The Need to Deal with What Banks Do. *Cap. Mark. Reform Asia Towar. Dev. Integr. Mark. Times Chang.* **2012**, *2009*, 231–268. [[CrossRef](#)]