Artificial intelligence and ethics within the food sector: developing a common language for technology adoption across the supply chain

Authors

Louise Manning[1], Steve Brewer[1], Peter J. Craigon[2], Jeremy Frey[3], Anabel Gutierrez[4], Naomi Jacobs[5] Samantha Kanza[3], Samuel Munday[3], Justin Sacks,[5] Simon Pearson[1]

Affiliations

1. The Lincoln Institute of Agri Food Technology, University of Lincoln, Lincoln, UK
2. Future Food Beacon of Excellence and School of Biosciences, University of Nottingham, Nottingham, UK
3. School of Chemistry, Faculty of Engineering & Physical Sciences, University of Southampton, Southampton, UK
4. School of Business and Management, Royal Holloway University of London, Egham, UK
5. Imagination Lancaster, LICA, Lancaster University, Lancaster, UK

ORCIDS

Peter Craigon https://orcid.org/0000-0001-5081-2718

Peter.craigon4@nottingham.ac.uk

Steve Brewer: https://orcid.org/0000-0002-9708-0427

SBrewer@lincoln.ac.uk

Jeremy Frey: https://orcid.org/0000-0003-0842-4302

J.G.Frey@soton.ac.uk>

Anabel Gutierrez: http://orcid.org/0000-0002-3730-8282

anabel.gutierrezmendoza@rhul.ac.uk

Naomi Jacobs: https://orcid.org/0000-0003-2079-1658

naomi.jacobs@lancaster.ac.uk

Samantha Kanza: https://orcid.org/0000-0002-4831-9489

s.kanza@soton.ac.uk

Louise Manning http://orcid.org/0000-0002-9900-7303

Samuel Munday: https://orcid.org/0000-0001-5404-6934

s.a.munday@soton.ac.uk

Simon Pearson: https://orcid.org/0000-0002-4297-4837
SPearson@lincoln.ac.uk
Justin Sacks: https://orcid.org/0000-0002-8643-2087
j.sacks@lancaster.ac.uk
Corresponding Author LManning@lincoln.ac.uk

Abstract:

Background: The use of artificial intelligence (AI) is growing in food supply chains. The ethical language associated with food supply and technology is contextualised and framed by the meaning given to it by stakeholders. Failure to differentiate between these nuanced meanings can create a barrier to technology adoption and reduce the benefit derived.

Scope and approach: The aim of this review paper is to consider the embedded ethical language used by stakeholders who collaborate in the adoption of AI in food supply chains. Ethical perspectives frame this literature review and provide structure to consider how to shape a common discourse to build trust in, and frame more considered utilisation of, AI in food supply chains to the benefit of users, and wider society.

Key findings and conclusions: Whilst the nature of data within the food system is much broader than the personal data covered by the European Union General Data Protection Regulation (GDPR), the ethical issues for computational and AI systems are similar and can be considered in terms of particular aspects: transparency, traceability, explainability, interpretability, accessibility, accountability and responsibility. The outputs of this research assist in giving a more rounded understanding of the language used, exploring the ethical interaction of aspects of AI used in food supply chains and also the management activities and actions that can be adopted to improve the applicability of AI technology, increase engagement and derive greater performance benefits. This work has implications for those developing AI governance protocols for the food supply chain as well as supply chain practitioners.

Keywords: responsibility, accessibility, explainability, accountability, interoperability, artificial intelligence,

Highlights

- AI applications are increasingly being adopted in food supply chains.
- AI empowers decision-making, but its use must be framed by ethical considerations.

- Benefits/risks of using AI are constantly evaluated in the AI development cycle.
- Improving explainability, interpretability and accessibility enables transparency.
- Responsibility and accountability relate to governance structures for use of AI.

1. Introduction

Artificial intelligence (AI) is a computational technology that seeks to mimic, to differing extents, human abilities to perceive their environment, process information, make decisions and to take steps to achieve pre-determined goals. From banking to autonomous driving, and from healthcare to farming, AI is empowering decision-making in every field and at every level. Within the agri-food space, digital technologies and information architectures are being used by farmers to maximise land use in terms of efficient yields of food commodities whilst also enhancing biodiversity (Cambra Baseca, Sendra, Lloret & Tomas, 2019; Köksal & Tekinerdogan, 2019; Mkrttchian, 2021). The collection of data and subsequent use of advanced data analytics, algorithms and AI enables the analysis of large datasets derived from multiple sources to deliver specific objectives or outcomes. This is already the case in many other domains such as medicine, but such activities must be approached cautiously to maintain trust (Durán & Jongsma, 2021).

The use of advanced data analytics, algorithms and AI can inform the wider supply chain on how a weather event, plant or animal disease, or other supply chain shock may impact, and if or when food crises are likely to happen (Kiran, Narayana Raj & Talawar, 2020). Agri-food and supply sectors and activities, where AI is being used, include smart irrigation and nutrient management, smart soil management, harvest predictions, livestock monitoring and behaviour prediction, quality and food safety assessment (Kakani et al., 2020). Data from multiple connected, and also discrete, sources can be assimilated, aggregated and translated within a smart farming approach (Wolfert, Ge, Verdouw & Bogaardt, 2017). The potential for AI to aid and address humanity's problems, such as food insecurity or climate change is also matched by concerns about the impact of indiscriminate unconsidered use and the harms that may arise. To this end, the developments in the use of AI have been concurrent with a growth in frameworks and approaches to AI-related ethics seeking to safeguard against the considerable potential for AI enabled harm whilst maximising the significant benefits of AI technologies to society (AI Ethics Guidelines Global Inventory, nd).

The ethics of food production and food consumption is already a vast field of enquiry to consider, made larger still when the ethics associated with technology and its socioeconomic and socio-political impact are considered (Mepham, 2000). Applying AI requires consideration of the ethical implications of not only the implementation of the systems proposed, but also their impact on the wider food community. This impact ranges from how the technology affects the grower/farmer, to how it affects business practices along the supply chain, to how right, or wrong is contextualised, and whether it is a requirement to encourage or empower consumers to ethically use the extra information such technology would bring. The increasing use, and interconnected nature of distributed information technology, and the ever-growing reliance upon greater volumes of big data to feed AI algorithms are raising ethical challenges across the agricultural and food industry that regulators and society are struggling to contextualise and operationalise in practice (Ahearn, Armbruster & Young, 2016).

Algorithms “sift through data sets to identify trends and make predictions” (Martin, 2019, p.835). Algorithms can vary from simple, specified transparent sets of rules (instructions) that can be followed to solve a problem or undertake a calculation or process data, to algorithms that are sophisticated self-learning processes that can self-train and adapt their analysis procedures and self-learn (Durán & Jongsma, 2021). The latter are often called black box algorithms as they cannot be interrogated by the humans that use them and are often considered opaque in terms of the outputs they produce (Setzu et al. 2021). This raises ethical concerns of hidden discrimination and bias within system design and application, and questions can arise around aspects of transparency, responsibility, accountability, auditability, trustworthiness, culpability, reliability, explainability, interpretability and accessibility (Friedman & Nissenbaum, 1996; Martin, 2019; Durán & Jongsma, 2021; Setzu et al. 2021).

Ethical considerations of AI are often centred on issues of privacy, agency and accountability, particularly in relation to the use of personal data in computational systems. This can be seen in the enactment into law of the European Union General Data Protection Regulation (GDPR, 2018) which stipulates a series of principles, definitions, rights and responsibilities for the development and use of systems that capture and process personal data (EUR-Lex, nd). Key amongst these considerations are issues of explainability, accountability, transparency (e.g., a right to an explanation) and responsibility (e.g., a right to determine responsibility for outcomes). Whilst the nature of data used within food systems is much broader than the personal data covered by the GDPR, the ethical issues for computational and AI systems are comparable.

The aim of this review paper is to consider the ethical narrative used by stakeholders when collaborating to adopt AI in food supply chains. This review has been undertaken to explore ethical perspectives to consider how to develop a common discourse to build trust in, and more considered utilisation of, AI in food supply chains. This will benefit multiple stakeholders including food scientists, policy makers and industry specialists as they collaborate and communicate about AI with each other. The authors, who come from a range of academic disciplines, organised a series of review workshops that formed a central part of the research process to explore the collective narrative and interplay of perspectives that inform the paper. These discussions and the paper itself emerged from a foundational body of literature within each discipline and were developed through a snowball academic literature review that synthesized evidence that supported and deepened the collective narrative (Kowalska & Manning, 2021; Jacobs et al., 2021). For a wider explanation of the methodology for the whole research project see Jacobs et al., (2021). The seven aspects considered in this paper have been critiqued and positioned (Table 1) in terms of the inherent characteristics and corporate and supply chain activities and mechanisms which can embed these aspects in food supply chains.

Take in Table 1.

After reflecting on some of the ethical aspects of the use of AI in the context of the food supply chain, we explore the aspects of the vocabulary that were commonly used in the workshop discussions: transparency, traceability, explainability, interpretability, accessibility, accountability and responsibility. We critique how this range of vocabulary is framed by different actors and relate these terms to the development and implementation of AI within the food supply chain.

2. Ethics, morality and food

Ethics is defined for the purposes of this research as a set of moral principles that inform judgements of right or wrong for a particular group or activity. As a discipline, ethics can be broadly divided into three areas of interest: firstly, moral philosophy or *meta-ethics*, which is concerned with the nature of morality, and secondly, *normative ethics* which seeks to provide structures or norms to guide ethical behaviour according to approaches such as virtue, deontological or consequentialist measures, including the notions of rights for example. The third area is *applied ethics* which seeks to adapt such normative frameworks and other consideration to guide behaviour in real life contexts according to the area of interest, for example medical ethics or bioethics (Durán & Jongsma, 2021). The intersection of applied food supply related ethics and emerging technology focused ethics is where this work is situated.

2.1 Socially and technologically determined ethics

The influence of AI and the associated ethical considerations is often viewed through the lens of the degree of agency the technology is afforded in how it influences, constrains, and produces the lived experience of the people that are subject to it. The agency associated with technology can be seen as a continuum. One perspective is that technology is an innocent value-free tool whereby it can bear no innate responsibility for its influence on the people who use it. This viewpoint suggests responsibility is *socially determined* (social determinism) and solely the responsibility of the stakeholders that interact with the technology. At the other extreme, *technological determinism*, sees technology being innately afforded responsibility and influence in shaping human behaviour and society, and cultural development through its use or other social factors (Kostina & Khorina, 2012). Martin (2019) states that the greater the degree of agency that an individual has over the operation of the algorithm the less the degree of accountability that can be attributed to the role of the algorithm itself within the decision process. Others suggest there is an interaction between accountability and answerability, where algorithms are used to inform human decision-making and this requires aspects of explanation and justification to be suitably addressed (Busuioc, 2021).

Whilst there are different perspectives on where technologies such as AI, are positioned on this socio-technological spectrum (between social and technological determinism), and on where ethical questions sit as well, it is important to consider that there will be variation in perspectives and the socio-technological aspects of interest may change over time. With the introduction of AI technologies, we may also have to address questions relating to how much responsibility can be afforded to automated systems that aid or make decisions independently. Adoption of technology in agriculture will potentially reorder or reengineer already complex animal-human-technology-plant-natural-environment relationships. For example, using automatic milking machines as a case, technological determinism will inform the design and deployment of automatic milking machines to drive optimum performance, but their adoption can fundamentally influence associated human-animal relationships (Schewe & Stuart, 2015). The reverse can also be the case in that as human-animal relationships evolve this will influence how technology is used to support those reframed human-animal relationships. Dafoe (2015) proposes that this social versus technical dichotomous argument is problematic and we ought to consider that lived reality is a more nuanced socio-technical relationship that is dynamically centred around the autonomy of technological change and the associated change of society. Further, Dafoe argues the design of technological solutions can deliver not only intended outcomes, but also unintentional outcomes especially in the event of unforeseen selective pressures. These unintentional outcomes can then shape societal norms and expectations.

The development of AI technologies in the sphere of agri-food brings data and new technological interactions into food-related socio-technical systems with the promise of greater efficiency. This both raises new ethical issues and also potentially addresses complex ethical dilemmas that already exist within the food system. Smart agriculture, climate-smart agriculture, or internet of things (IoT) based agriculture are terms that can be considered as an example of this contextualisation. These terms frame the widespread adoption of technology as having a net positive benefit, but they also reorient agricultural systems under a new reality (Lipper et al., 2014). However, there is the potential for such technologies to increase power imbalances to the commercial disadvantage of those who are unable to access or afford such technologies or the infrastructure to operate them (Long, Blok & Coninx, 2016). This is often called the "digital-divide" (Mark, 2019). New technological approaches in food supply chains mean that the digital-divide is no longer just information asymmetry and a lack of knowledge and information for some stakeholders, but also the wider ethical framing of financial and social accessibility to that data and information (Long, Blok & Coninx, 2016). The processes that have been used to package information for users, the decisions, the pre-existing and emergent biases (Friedman & Nissenbaum, 1996; Buolamwini & Gebru, 2018), which drive opacity and prevent open and free sharing of data (Martin, 2019; Durán & Jongsma, 2021), or fail to disclose the inherent value of the data collected, all impact trust in such technology (Mark, 2019). If the data produced and stored could be integrated in a mutually agreed way, e.g., in the form of a data trust (Brewer et al., 2021, Durrant et al. 2021), then this could reduce such concerns, yet there are significant barriers to achieving this (van der Burg, Wiseman & Krkeljas, 2020). Thus, applying ethical consideration is central to realising the potential organisational and individual benefits in a fair and equitable way for all the actors involved in the food system.

The use of AI requires both effective governance structures and also open collaboration between multiple stakeholders such as food businesses, traditional technology companies, and new entrant disrupters (Wolfert, Ge, Verdouw & Bogaardt, 2017). Albeit in a non-food context, studies have explored the barriers to collaboration caused by a lack of understanding of common domain expertise, an absence of shared vocabulary, or a lack of trust (Saunders & Corning, 2020). With such a variety of uses and users, the language surrounding the new technology and the inherent assumed meaning derived from given activities and operations may vary depending on the specific implementation of AI at a given step or stage in the supply chain. Each disciplinary domain defines the language surrounding their work. In food and agriculture specifically, complex meaning can develop around local and industry level vocabulary and when and how language and discourse is used, revised and refined, so specific vocabulary becomes culturally embedded over generations (Malhotra, 2001, p.7).

Addressing food supply and sustainability from a systems level perspective requires a collaborative approach from all actors with a common, mutually understood vocabulary. Ethical concerns can arise, and we can highlight some areas of primary ethical concerns identified by the Nuffield Council for Bioethics for the need to provide food in a sustainable manner (Jackson, 2018). Summarising their discussions according to the values embodied therein they identified the following areas of key interest: food and nutritional security; health and access to sufficient, safe nutritious food; fairness and equity through fair access to food, distribution of risk and treatment of farmers and others within the food system; responsibilities i.e. consideration of the roles of actors in the systems including governments, farmers, manufacturers amongst others; democracy and giving people a say in food systems and associated research; autonomy choice and diversity enabling choice to allow people to express their identities and preferences; high farm animal welfare; and environmental sustainability i.e. preserving the environment for future generations due to its intrinsic value.

Considering 'ethics' as a whole is an important first step in laying the groundworks for how we view the rest of the terms described in this paper. Without properly interrogating each of the aforementioned ethical aspects it becomes difficult to properly assess the ethical implications of any decisions that have been made to embed AI in agri-food chain applications. It is important to ethically interrogate the human-technology interaction and the ethical impact of actors (food technologists, computer programmers, farmers etc.) using differentiated meanings to frame the use of AI. Differentiated meanings are considered in this paper to represent meanings that can be enacted by different people from the same information at the same time, or when considering the same issue at different times (Malhotra, 2001). Further, Malhotra (p. 7) suggests that meaning is a critical construct to understand: "*how humans convert information into action and consequently performance, it is evident that information-processing based fields of AI and expert systems could understand how humans translate information into meanings that guide their actions.*" In summary, stakeholders need to develop sense making strategies to position a collective narrative that all disciplines can own and use and as a result reduce ambiguity and build mutual trust. The seven aspects are now considered in turn.

3. Aspects of AI and algorithm application in food supply chains

3.1 Transparency and Traceability

It is important here to differentiate between transparency and traceability. Traceability is the ability to follow the history, application, movement and location of an object (product, material, unit, equipment or service) through specified stage(s) of production, processing and distribution (ISO, 22000:2018). Regulation EC/178/2002 defines traceability as the ability to trace and follow a food, feed, food-producing animal or substance intended to be, or expected to be incorporated into a food or feed, through all stages of production, processing and distribution. A traceability system is therefore a "*record-keeping and task-triggering mechanism to improve consumer confidence in food consumption and to efficiently reduce the asymmetry of information across food supply chains*" (Chen, 2015, p.70). Traceability information adds value to the product as it enables supply chain partners to meet product standards and customer expectations (Pizzuti & Mirabelli, 2015). Thus, traceability is a transactional process of tracing ingredients forward to final products and food products back to source ingredients, and yet at the same time the process creates a set of credence attributes such as consumer confidence, trust, promotion of health benefits (Anastasiadis, Apostolidou, & Michailidis, 2021), openness or transparency that add value to the product itself (Islam & Cullen, 2021).

Traceability systems also underpin reliable, cost-effective quality and safety management (Anastasiadis, Apostolidou, & Michailidis, 2021). Qian et al., (2020) suggest there has been three evolutions of traceability systems:

Traceability System 1.0 compliance and information recording in simple paper or electronic systems.

Traceability System 2.0 data integration – real-time information sensing and integration across the supply chain utilising Internet of Things (IoT) and Distributed Ledger Technology (DLT).

Traceability System 3.0 intelligent decision-making systems that improve food safety and quality management and utilise emerging technologies.

Transparency is the characteristic of being visible and open. In the food context, transparency is about the visibility and assessment of the production process and the associated disclosure activities by one actor to other actors in the supply chain (Turilli & Floridi, 2009;

Manning, 2018). Modern food supply chains with a wide range of stakeholders have become increasingly more complex (Astill et al., 2019) and there are serious potential consequences to non-transparent food supply chains such as food adulteration e.g. horsemeat substitution or seafood fraud (Leal et al., 2015), and under diagnosis during outbreaks of foodborne illnesses (Hoelzer et al., 2018). It is the nature of the disclosure mechanism, the access agreement and the purpose for access that is most important when considering transparency, and a failure to do so will drive inbuilt bias and embedded power relationships (Egels-Zanden, Hulthen & Wulff, 2015; Mol, 2015; Gardner et al. 2019). In order to monitor operational activities and mitigate supply chain risk, organisations will focus on supply chain transparency, enabling them to monitor and manage operational activities (Zhu et al. 2018). Supply chain transparency is a tool that can respond to consumer pressure to disclose information and a willingness to buy or alternatively a corporate mechanism to increase revenue and reduce costs (Egels-Zandén & Hansson, 2016). Transparency in the political context can be described as information about decisions and decision-making processes that is provided or made available to the public (de Fine Licht, 2014a). Information in this context is different to data.

Indeed, there is a difference between actual decision-making processes and public perception of decision-making processes that means perceptions of transparency also influence attitudes towards legitimacy and this in part is mediated by trust (de Fine Licht, 2014b). Legitimacy in this context is the perception that the actions of an individual or organisation are "desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions" (Suchman, 1995, p.574). Thus, the central constructs "upon which the concept of legitimacy rests are norms, values, beliefs, and morals" (Suddaby, Bitektine & Haack, 2017). de Fine Licht (2014a) suggests that there are degrees of transparency i.e., transparency can be partial or full, indeed the same can be said of personal or corporate disclosure itself. Therefore, perceptions of transparency are shaped by transparency cues, and how they are appreciated and understood by a range of stakeholders, rather than by the degree of actual transparency in information sharing in the first place (de Fine Licht, 2014b).

Transparency cues are statements provided by external sources (de Fine Licht, 2014b). In the food supply chain, for example, third-party certification (TPC) provides market signals and the opportunity for assurance that such cues are associated with a set of defined private standards that are routinely independently verified at steps in the supply chain (Rees, Tremma & Manning, 2019). In a given context, these cues can be cognitively or procedurally ordered in terms of hierarchy (rank-ordered cues) and value in order to inform decision making and can drive perceptions via positive validity or negative validity mechanisms (Kurz-Milcke, Gigerenzer & Martignon, 2008), for example the binary aspects of organic versus conventional product, geographic origin versus no claim being made and so on. This area is worthy of further research to consider the use of transparency cues in machine learning applications in food supply chains (Chao, Cakmak & Thomaz, 2010).

The process of being transparent allows autonomy, greater democracy and equity and informed decision-making in the supply chain and also drives accountability (Dingwerth & Eichinger, 2010; Mol 2015). However, it is important to share information using mechanisms that will retain the quality and quantity of information i.e., no loss, delay, distortion or noise (Hofstede et al. 2004; Wognum et al. 2011). These mechanisms also play a role in supply chain agility and response (Zhou et al. 2014). Further, the innate characteristics of the data and information itself impact on its innate transparency e.g., accuracy, relevance, reliability and timeliness (Hofstede et al. 2004; Wognum et al. 2011). The characteristics of the data and the process of translation into distinct disclosure activities influences the extent to which stakeholders believe that an organisation itself has acted in an open and transparent way (Manning, 2018). Mol (2015) states there are multiple forms of disclosure that reduce information asymmetry that can be characterised as disclosure of information ‘by’ economic actors in supply chains, regulators and certification bodies and disclosure of information ‘for’ the downstream economic actors in supply chains, regulatory, certification and inspection bodies, consumers, the public as citizens and the media. Context around the disclosure activity e.g., whether it is voluntary or not, willing or reluctant, accessible or dense will influence actors’ perceptions of whether an organisation is perceived to have been transparent (Turilli & Floridi, 2009). The quality of information disclosure therefore not only reflects the quantity of information, but also the density or richness of content (Beretta & Bozzolan, 2004).

In summary, transparency firstly depends on effective traceability i.e., collection, analysis and dissemination of data (Mol, 2015); and creating greater visibility of the findings often taking complex supply chain information and developing processes of “simplification, reduction, standardisation and disembedding” of data from its existing contexts (Gardner et al. 2019). Dissemination through reporting and disclosure can be via reports, score cards, platforms, calculators, certification, labelling and packaging cues (Egels-Zanden et al. 2015, Gardner et al. 2019). This approach in turn can drive active, timely decision-making and action. Transparency within food supply chains will then enable informed decision-making by single and multiple actors. The notion of “being transparent” at the technology level is more nuanced. Consideration at the wider socio-technical perspective, means transparency is crucial when defining both explainability and the ethical questions that surround food supply chain processes and activities. Achieving transparency across these complex supply chain/network models is not a simple task. In recent years though, a new suite of technologies such as Federated AI, DLTs (including Blockchain) and IoT have enabled significant advances that, when combined with AI and machine learning, could be used to create a new level of digital systems to enhance transparency in the food chain.

The ethical consideration of algorithmic transparency in particular has become even more important with the emergence of these new advanced technologies (Bertino, Kundu &

Sura, 2019; Larsson & Heintz, 2020). Indeed, transparency has become one of the key requirements in "Trustworthy AI" (European Commission, 2019), with a strong focus on creating transparent algorithms (Blacklaws, 2018; Boscoe, 2019; Rauber, Trasarti & Giannotti, 2019). Thus, a transparent algorithm should be visible and open in order to comply with these regulations, which would apply to any digital system for the food supply chain with respect to both the food itself, and any associated data and algorithms. A key aspect is considering how to make algorithms transparent as opposed to black box algorithms that are opaque (Martin, 2019). Making the code behind algorithms open source and therefore available to access is one approach. This outcome however, as noted by Blacklaws (2018), is often not enough by itself and is not likely to make the algorithm non-opaque due to the innate complexity, inscrutability, and lack of understandability inherent in such algorithms. Less complex interpretable algorithms are proposed as an approach instead (Busuioc, 2021). Indeed Busuioc (2021, p. 834) questions whether the use of black-box algorithms is justifiable, particularly when 'interpretable alternatives are available'.

The code behind some algorithms is only one element in the process, as machine learning algorithms will learn from the data they are trained on. Indeed, innate biases in the training data will be learned by, and eventually coded into, the algorithm (Martin, 2019). This postulates the notion that perhaps access to both data and the algorithm will infer transparency, although this still may not prove to be the case as an understanding of how the code works and weights the data would be required. Another key consideration is who the system is providing transparency for; as there will be different transparency and explainability requirements between, for example, users who want to understand why decisions are made by the AI, to incident investigators who are trying to trace the causes of a food safety or health and safety incident, and auditors who are evaluating the potential for bias in a system. In proposing a new standard for transparent autonomous systems, Winfield et al., (2021) highlight that not only do these different groups of stakeholders exist, who may have different transparency requirements, the *appropriate* level of transparency in each case may vary for each context, taking into account the specific autonomous system in question and its socio-technical context. For example, proprietary data and algorithms may need to be protected (Busuioc, 2021) and therefore are less transparent to all except auditors, and a security system which functions through its obscurity should not be transparent to the general public, though it may still be explainable. Explainability as a characteristic is now considered in more detail.

3.2 Explainability

Explainability has been linked with either being intelligent, being knowledge-based, providing meaning, creating understanding, reconciling differences (Gregor & Benbasat, 1999); or a process of communication and interpretation, "*facilitating the human user's understanding of the agent's logic*" (Rosenfeld & Richardson, 2019, p674). Setzu et al., (2021) distinguish between being explainable by design (ante-hoc) i.e., the AI or algorithm is explainable via the problem it is trying to solve, or post-hoc i.e., explaining the decisions that have been made. By using explainable design criteria, explainable processes, and explainable algorithms we can introduce transparency into the use of AI in the agri-food sector.

In the wider field of AI there has been considerable work in positioning 'explainable artificial intelligence' or XAI with the 'X' being phonetic for 'ex'plainable (Gunning et al. 2019; Royal Society, 2019). XAI allows '*users and parts of the internal system to be more transparent, providing explanations of their decisions in some level of detail*' (Gilpin et al., 2018, p. 80). The General Data Protection Regulation (GDPR); (2018) introduced, to some extent, a right of explanation for all individuals to obtain "*meaningful explanations of the logic involved*" when automated decision-making takes place. This has a profound effect not only on the ethics of systems, but on how they regulate safety and industrial reliability too. Meaningful means that the communication process is framed in a way that recognises different audiences have varied capacity to understand and interpret information and as a result supports improved understanding and accountability through detailed and individualised explanations (Suzor, West, Quodling & York, 2019). Brauneis and Goodman (2018) use the term *meaningful transparency* as the first step towards having sufficient knowledge to approve or disapprove of an algorithm's performance. They position this against *perfect transparency* where stakeholders have "complete knowledge of an algorithm's rules of operation and process of creation and validation" (p.31).

Tools are being developed that use big data to optimise food supply chains, increase food security and help with food production. Fusing these tools with XAI will ensure that there is meaningful if not perfect transparency across the sector. Indeed, the Food and Agricultural Organisation of the United Nations (FAO) has signed up to following the ethical resolution on AI (Mehmet, 2020), the so-called 'Rome Call for AI Ethics.' (Romecall, 2020) This "Call" highlights the importance of implementing a 'highly sustainable approach, which also includes the use of AI in ensuring sustainable food systems in the future. One of the key aspects in the FAO's ethical resolution on AI is that it must be explainable, though there is no definition either of what XAI is or how it relates to the food industry specifically. One working definition of explainability may be that models must be developed (ante-hoc) that are inherently easy for the user to understand (Rosenfeld & Richardson, 2019), or alternatively "*extracting some form of explanations from complex pre-developed models that are otherwise difficult (if not impossible) to understand for their users"* (Khaleghi, 2019, p.1). However, in this context, Bryson (2019, p.8) differentiates between explainability and understandability stating: "we do not need to completely understand how a machine learning algorithm works to regulate automated decision making, any more than we need to completely understand the physics of torque to regulate bicycle riding in traffic."

A further potentially more technical definition of explainability has been offered by Dhurandhar, Iyengar, Luss & Shanmugam (2017) where they define explainability relative to a target model which is applied to a task rather than a concept. In particular, explainability is defined as a process where some information is extracted from a complex model and communicated to a target model, in this case a human, to improve performance. The Dhurandhar, et al., (2017) definition does not require the target model to be a human. In practice, it can be any model e.g., a linear model or a decision tree. Another advantage of this contextualisation is that it makes it straightforward to compare different explainability methods based on the performance gain of the relative target model. If this definition of explainability is related back to the agri-food industry some constructs become clear. Firstly, the level of explainability needed will be different at each stage of the chain, and the consequences of not being able to explain a given output from a machine learning model will also differ at each stage and with different actors. This is not to suggest that different definitions of explainability are needed, rather that definitions of explainability must be able to encompass different perceptions and meanings associated with explainability at each stage by different actors i.e., it must be human agent centric.

Secondly, decisions driven by the output of an algorithm must be properly tempered with the experience and insight of human agents if they are to be generally meaningful and ‘explainable’ to users at other points in the chain. Using the previous example of automatic milking machines whilst an output from a robotic milking system may be explainable to the farmer in the context in which they are using the technology, it may not be considered as explainable by consumers who are purchasing the associated dairy products. Therefore, the ‘explainability’ of AI used in the food sector must be judged, by those with the correct expertise and understanding, for its ability to be understandable for multiple different users sometimes in different timeframes, and for users with the correct technical experience to come to the same conclusion as the AI given the same information or understand the output from the AI and how it has been derived. Rosenfeld & Richardson (2019), highlight the ethical context of the link between transparency, explainability, and interpretability. The next section will consider the characteristics of interpretability in more detail.

3.3 Interpretability

The process of information assimilation and interpretation requires data to be collated, ordered, and analysed by one or more supply chain actors who each assign a given and sometimes differentiated meaning. Whilst Lipton (2018) considers terms such as transparency, explainability, visualisability, and interpretability, the research acknowledges that interpretability still has a lack of consensus on its definition. Interpretability and visualisability of algorithms by humans have been linked by other literature (Durán & Jongsma, 2021) especially the use of visualisation tools, prototype analysis, and feature analysis as a foundation to demonstrating transparency (Rosenfeld & Richardson, 2019). Doran, Schulz and Besold (2017) define interpretability as the opposite of opacity or black box i.e., a system where users can see, study and understand how inputs are mathematically mapped to outputs. The nuances of social determinism and technological determinism have been touched on in this paper but are worthy of further research and critique in the context of the use of AI in food supply chains. Opacity and transparency in the design, development and implementation of AI applications in the food supply chains can only be assured if the factors that lead to “black box” algorithms are widely understood. Inherent in this process is the interaction between the technology and human agents at different stages of the supply chain. As a result, differentiated meaning can arise at either different steps in the supply chain, or where information asymmetry occurs affecting interpretability, explainability and transparency.

Rosenfeld and Richardson (2019) propose six approaches to generating interpretations, each with different aspects of explicitness and faithfulness, the latter which links to trust (see also Lipton, 2018). The concept of trust, especially consumer trust is not discussed in depth here, but is an underlying aspect of meaning associated with the use of AI. The six approaches are interpretability via: (a) use of a transparent machine learning algorithm, (b) design and feature selection and/or analysis of the inputs; (c) using an algorithm to create a post-hoc model tool, (d) using an algorithm to create a post-hoc outcome tool; (e) using an interpretation algorithm to create a post-hoc visualisation of the agent's logic or (f) using an interpretation algorithm to provide post-hoc support for the agent's logic via use of prototypes. Interpretation of given content will be mediated by the degree of local or content specific knowledge of the user (Suzor et al. 2019) and thus will vary between users. Accessibility relates to usability of information, tools or technology and this is now explored in the next section.

3.4 Accessibility

Accessibility can have many different meanings even within the domain of food supply chains. In the context of food, it can refer to the cognitive accessibility of information pertaining to the food, such as nutritional information to help consumers make informed choices about the food they purchase (Wellard, Glasson, Chapman & Miller, 2011). Alternatively, it can refer to physical accessibility of food itself, such as enabling access to varied, healthy and inexpensive food to aid with public health (Apparicio, Cloutier & Shearmur, 2007). In the context of digital collaboration, data sharing and use of AI in the agri-food sector, it is also important to consider the technical aspects of accessibility. In the areas of computer science and data science there are different characteristics presented by the FAIR principles (Findable, Accessible, Interoperable and Reusable) i.e., data should be accessible in a way that it can always be "obtained by machines and humans" (Wilkinson et al. 2016). This definition addresses the need for appropriate authorisation levels and protocols for data access.

Accessible does not mean that all should be data be freely available, rather there can be degrees of accessibility especially for proprietary data where companies do not wish to release datasets into the public domain. Proprietary data may be retained as private and 'permissioned' to protect competitive advantage. Similarly, there is often unwillingness to share data between organisations, making it difficult share information across a supply chain (Brewer et al., 2021); an issue when developing and embedding traceability systems. Further, certain software can also make data and consequential information inaccessible by holding it hostage, either through the use of proprietary data formats that cannot be easily read by other pieces of software, or by a refusal to allow data to be taken out of a software package which is also known as "vendor lock-in" (Wiley & Michaels, 2004; (Gutierrez, Boukrami & Lumsden, 2015). Thus, when AI or algorithms are used in the food supply chain, accessibility for users, individually and collectively, needs to be negotiated between stakeholders.

3.5 Accountability

Accountability at government and business levels involves tracking and/or mapping how and why decisions are made, who makes those decisions and on what basis, how power is used in these processes, whose views are important and who ultimately holds decision makers to account (Kraak, Swinburn, Lawrence & Harrison, 2014). Nissenbaum (1996) positions accountability in terms of 'answerability': the obligation to give information about an action taken, explaining or justifying the taking of that action, and the obligation to make some kind of consequent action, including punishment, rectification etc. Obligation suggests a sense of duty i.e., that accountability links both to being legally required, compulsory, and also that obligation is morally framed suggesting legal liability and accountability could be driven by normative voluntary standards. Koppell (2015) suggests accountability is comprised of several dimensions: liability, controllability, responsibility and responsiveness.

Binns (2018, p. 544) considers accountability from a transactional viewpoint i.e., that "A is accountable to B with respect to conduct C, if A has an obligation to provide B with some justification for C and may face sanction if B finds [the] justification inadequate." In the food industry this could be illustrated as Business 1 is accountable to Business 2 for the material they supply being nut-free as per the specification agreed i.e., the justification C is that any presence of nuts should be prevented. Business 1 may supply assurances to Business 2, but Business 1 may face sanctions if they cannot demonstrate they have suitable protocols in place, or have not followed those protocols adequately, to provide nut-free product. In the UK, the House of Lords Select Committee report on AI (2018), states accountability is primarily framed through who is responsible if something goes wrong i.e., in terms of culpability. As a comparison, the Japanese Society for AI principles report (2017, p. 3) includes both pre-emptive and retroactive approaches to accountability, stating that: "In the event that potential danger is identified, a warning must be effectively communicated to all of society…. If misuse of AI is discovered and reported, there shall be no loss suffered by those who discover and report the misuse."

Accountability can also be considered as a policy structure or framework with associated principles (trust, inclusivity, transparency and verification), protocols and mechanisms to hold stakeholders accountable for their actions and behaviours thus making them answerable to those with a particular level of authority (Kraak et al., 2014). Diakopoulos (2015) considers the concept of accountable algorithms and how this relates to the accountability of the people who develop them or who use them. Diakopoulos (2015) suggests that an element of accountability is the development of algorithmic accountability reporting which encompass the assessment of input-output relationships, and aspects of fairness and understanding an algorithm's influence, mistakes, and/or biases; all key elements of verifying transparency. In 2019, the Institute of Electrical and Electronics Engineers (IEEE) launched the P7000 standards projects intended to create a series of new standards to address ethical issues in the design of autonomous and intelligent systems, many of which have specific focus on aspects of responsibility and responsible technology development (Peters, Vold, Robinson & Calvo, 2020). The final aspect considered in this paper is responsibility.

3.6 Responsibility

Responsibility in all areas of food production and supply underpins food safety and trust, not only with the food itself, but also in production processes. This is often considered in terms of corporate social responsibility (Maloni & Brown, 2006) i.e., voluntary action by companies above minimum legal requirements where principles include legitimacy, public responsibility and managerial discretion. Responsibility can be understood through the lens of Responsible Research and Innovation (RRI) where it is defined on a high level as an interactive focus on the societal desirability, ethical acceptability and sustainability of research and its products to allow a proper embedding of scientific and technological advances in society (Von Schomberg, 2011). There is a growing field of work seeking to define responsible AI and consider how it can be achieved in practice (Dignum, 2019). AI-based food industry applications are frequently deployed in dynamic and unpredictable real-world environments because they promise the ability to react to complex situations quickly, effectively and with precision (Yang, Feng & Whinston, 2021). However, this very flexibility means that they might react in unpredicted or unanticipated ways, which can lead to undesirable or even harmful consequences. There is also no clear consensus on what it means for AI to be responsible (Jobin, Ienca & Vayena, 2019). It is generally agreed that responsible systems must address issues such as bias, transparency, justice and non-maleficence, but Martin (2019, p. 835) seeks to question whether developers are responsible "for their algorithms later in use, what those firms are responsible for, and the normative grounding for that responsibility" and concludes that the responsibility sits with organisation unless the designer has designed the algorithm "to preclude individuals from taking responsibility within a decision, then the designer of the algorithm should be held accountable for the ethical implications of the algorithm in use." (Martin, p. 825). The responsibility for errors in decisions made by AI and machine learning algorithms also needs to be considered (Kosior, 2020).

Human decisions about how data is utilised, included or discarded in a given technological application, will be driven by pre-conceptions. When training and developing AI systems it is extremely important that the data used do not contain biases or lack representativeness of specific categories. For example, a given group or community may not have been adequately represented in the data used to train a given algorithm, and this may reverberate on the accuracy of the recommendations provided by the AI system. This applies both to cultural diversity (e.g., recommending types of food that are prohibited by specific cultures), and ethnic diversity (e.g., specific ethnicities feature particular intolerance for the specific products or ingredients, e.g., lactase deficiency (see Buolamwini & Gebru, 2018). Beyond data collection, an algorithm's design has the potential to echo any pre-existing biases its human creator may have. Even if this is not the case, there is still scope for any technical biases to influence an application due to any limitations in the computer programme, its processing power or any other constraints that there may embedded in the system. Furthermore, if an otherwise unbiased algorithm is applied in an unanticipated context an emergent bias can be present.

In 2018, the Montreal Declaration for Responsible Artificial Intelligence was released following a year of public consultation. One of the 10 key principles included was responsibility, which is defined in these terms:

1. Only human beings can be held responsible for decisions stemming from recommendations made by AI system (AIS) based applications, and the actions that proceed therefrom.
2. In all areas where a decision that affects a person's life, quality of life, or reputation must be made, where time and circumstance permit, the final decision must be taken by a human being and that decision should be free and informed.

3. The decision to kill must always be made by human beings, and responsibility for this decision must not be transferred to an AIS.
4. People who authorise AIS to commit a crime or an offense, or demonstrate negligence by allowing AIS to commit them, are responsible for this crime or offense; and
5. When damage or harm has been inflicted by an AIS, and the AIS is proven to be reliable and to have been used as intended, it is not reasonable to place blame on the people involved in its development or use.

These principles do not only encompass obvious harms such as accuracy of recommendations and predictions (for example, if an automated system failed to give appropriate notification and labelling of likely allergen contamination) or of bias (for example smaller or marginalised producers being negatively impacted for loan approvals), but more complex changes too. These five points also align with Asimov's "Three Laws of Robotics" (1984):

1. A robot may not injure a human being, or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

At a wider level, questions of beneficence and harm to humans also include concerns over system-wide technological change, for example whether sector-wide introduction of AI and automation might have impact on employment levels, and potential sustainability questions over the energy requirements of automated and computational systems. It is important to consider that responsible use of AI must protect human quality of life, and dignity, at all scales. This section has considered two aspects firstly responsibility of AI and secondly, responsible use of AI and both need to be considered in any application in food supply chains.

There have been a wide range of guidelines, recommendations and other materials from industry and the public sector which attempt to build ethical and responsible practices into the use of these technologies. For example, the Japanese Society for Artificial Intelligence (JSAI) set out ethical guidelines in 2017 to be applied by its members, consisting of 9 guidelines or principles, one of which is accountability and also social responsibility. Jobin, Ienca & Vayena (2019, p. 395) note in their survey of the related literature that: "*very different actors are named as being responsible and accountable for AI's actions and decisions: AI developers, designers, institutions or industry*". They note that there is an outstanding debate over "whether AI should be held accountable in a human-like manner or whether humans should always be the only actors who are ultimately responsible for technological artifacts." It is not clear what holding an AI accountable would necessarily entail in terms of current technology, however, as discussed in the Montreal Declaration for Responsible Artificial Intelligence (2020), questions of AI systems themselves being held accountable can be a distraction from necessary consideration of human rights and harms that may be done to humans by the inconsiderate use of AI. If an AI driven allergen alert system fails to upload information in the timeframe required to prevent highly vulnerable individuals from experiencing anaphylactic shock where does the responsibility for harm lie? Does it lie with the developer who produced the application, the organisation that has sold the application and/or the user because they are ultimately responsible for their own safety and should not rely totally on such applications or the manufacturer who has incorrectly labelled the food? These ethical questions lie at the heart of considerations around responsibility. As shown by Busuioc (2021), the nature of accountability with regards to the use of AI is complex and subject to varied intertwined technical and human factors. It is not a question therefore of holding technology or human responsible but instead considering how responsibility and accountability is changed within a (food) system involving the use of AI.

4. Concluding thoughts

The emergence of the use of AI and algorithms in food supply chains brings with it a new vocabulary and context. The aim of this review paper is to consider the embedded ethical language used by stakeholders who collaborate in the adoption of AI in food supply chains. Ethical perspectives frame this review and provide structure to consider how to shape a common discourse to build trust in, and more considered utilisation of, AI in food supply chains to the benefit of users, and wider society. The seven aspects of use of AI considered in this paper were critiqued and positioned in terms of their characteristics, corporate activities and mechanisms which can embed these aspects in food supply chains. Supply chain examples are included in Table 1 to explore the aspects in a practical context.

By structuring and synergising the vocabulary in this way, we are able to begin the process of considering how these ethical perspectives can be translated into practice in the use of AI in food supply chains. Greater supply chain transparency will require the industry to reduce information asymmetry, improve legitimacy and ensure decision-making is less opaque. Having a framework within which to discuss ethical aspects of technology implementation in the food supply chain will facilitate the consideration of complex ethical challenges such as algorithmic bias, which could lead to the privileging of one group in the food supply chain over another or compromise the efficacy of AI supported decision-making. This challenge of bias is worthy of further consideration in future research.

The drawing together of the narrative in this paper makes a contribution to existing literature by supporting a more rounded understanding of the ethical interaction of aspects of AI use in food supply chains and also the management activities and actions that can be adopted to improve the applicability of AI technology, increase engagement and derive greater performance benefits. This work has implications for those developing AI governance protocols for the food supply chain as well as supply chain practitioners.

The nuances of the social-technological determinism spectrum have been touched on in this paper but are worthy of further research and critique in the context of real-case use of AI in food supply chains. The varied interpretation of aspects of AI adoption in food supply chains e.g., considerations of transparency, accountability, responsibility has implications for different stakeholders to consider as they work together to develop technological applications. Stakeholders developing a mutual understanding of language use and a shared vocabulary will catalyse consideration of the ethical complexities of the use of AI within the food system. The outputs of this research assist in giving a more rounded understanding of the language used, exploring the ethical interaction of aspects of AI used in food supply chains and also the management activities and actions that can be adopted to improve the applicability of AI technology, increase engagement and derive greater performance benefits across the food supply chain. The development of ethical frameworks for the consideration of normative ethics and applied ethics can inform and guide behaviour in real life contexts. This work has implications for those developing AI governance protocols and ethical frameworks for regulation, private standards for the food supply chain as well as supply chain practitioners.

Acknowledgements

This research augments the Food Standards Agency (FSA) funded work led by the University of Lincoln to create a Data Trust Framework related to food safety (Brewer et al., 2021). The working group that contributed to this paper is joint funded by the Internet of Food Things Network+ (Grant Number: EP/R045127/1) and the Artificial Intelligence and Augmented Intelligence for Automated Investigation for Scientific Discovery Network+ (AI3SD) (Grant Number: EP/S000356/1).

References

Ahearn, M. C., Armbruster, W., & Young, R. (2016). Big data's potential to improve food supply chain environmental sustainability and food safety. *International Food and Agribusiness Management Review*, *19*(1030-2016-83146), 155-171.

AI Ethics Guidelines Global Inventory. Available at: https://inventory.algorithmwatch.org (accessed 13.12.2021)

Anastasiadis, F., Apostolidou, I., & Michailidis, A. (2021). Food Traceability: A Consumer-Centric Supply Chain Approach on Sustainable Tomato. *Foods*, *10*(3), 543.

Apparicio, P., Cloutier, M. S., & Shearmur, R. (2007). The case of Montreal's missing food deserts: evaluation of accessibility to food supermarkets. *International journal of health geographics*, *6*(1), 4. https://doi.org/10.1186/1476-072X-6-4

Asimov, I., (1984). The Bicentennial Man. *Philosophy and Science Fiction* (Philips, M., ed), pp. 183-216, Prometheus Books, Buffalo, New York.

Astill, J., Dara, R.A., Campbell, M., Farber, J.M., Fraser, E.D.G., Sharif, S., & Yada, R.Y., (2019). Transparency in food supply chains: A review of enabling technology solutions. *Trends in Food Science & Technology*, 91, 240–247. https://doi.org/10.1016/j.tifs.2019.07.024

Beretta, S., & Bozzolan, S., (2004). A framework for the analysis of firm risk communication, *The International Journal of Accounting*, 39(1), 265-288

Bertino, E., Kundu, A., & Sura, Z. (2019). Data transparency with blockchain and AI ethics. *Journal of Data and Information Quality (JDIQ)*, *11*(4), 1-8. https://doi.org/10.1145/3312750

Binns, R. (2018). Algorithmic accountability and public reason. *Philosophy & technology*, *31*(4), 543-556.

Blacklaws, C. (2018). Algorithms: transparency and accountability. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*(2128), 20170351. https://doi.org/10.1098/rsta.2017.0351

Boscoe, B. (2019). Creating transparency in algorithmic processes. *Delphi*, *2*, 12-22

Brauneis, R., & Goodman, E. P. (2018). Algorithmic transparency for the smart city. *Yale JL & Tech.*, *20*, 103.

Brewer, S., Pearson, S., Maull, R., Godsiff, P., Frey, J.G., Zisman, A., Parr, G., McMillan, A., Cameron, S., Blackmore, H. and Manning, L., 2021. A trust framework for digital food systems. *Nature Food*, *2*(8), pp.543-545. https://doi.org/10.1038/s43016-021-00346-1

Bryson JJ. (2019). The Artificial Intelligence of the Ethics of Artificial Intelligence: An Introductory Overview for Law and Regulation. The Oxford Handbook of Ethics of Artificial Intelligence.

Buolamwini, J., & Gebru, T. (2018, January). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency* (pp. 77-91). PMLR

Busuioc, M. (2021). Accountable artificial intelligence: Holding algorithms to account. *Public Administration Review, 81*(5), 825-836.

Cambra Baseca, C., Sendra, S., Lloret, J., & Tomas, J. (2019). A smart decision system for digital farming. *Agronomy*, *9*(5), 216.

Chao, C., Cakmak, M., & Thomaz, A. L. (2010, March). Transparent active learning for robots. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 317-324). IEEE

Chen, R-Y. (2015), Autonomous tracing system for backward design in food supply chain, *Food Control*, 51, 70-84

Dafoe, A. (2015). On technological determinism: a typology, scope conditions, and a mechanism. *Science, Technology, & Human Values*, *40*(6), 1047-1076.

de Fine Licht, J. (2014a). Policy area as a potential moderator of transparency effects: An experiment. *Public Administration Review*, *74*(3), 361-371.

de Fine Licht, J. (2014b). Transparency actually: how transparency affects public perceptions of political decision-making. *European Political Science Review: EPSR*, *6*(2), 309.

Dhurandhar, A., Iyengar, V., Luss, R., & Shanmugam, K. (2017). A formal framework to characterize interpretability of procedures. *arXiv preprint arXiv:1707.03886*.

Dignum. V. (2019). Responsible Artificial Intelligence. Springer.

Dingwerth, K., & Eichinger, M., (2010). Tamed transparency: How information disclosure under the Global Reporting Initiative fails to empower. Global Environmental Politics, 10(3), 74-96.

Doran, D., Schulz, S., & Besold, T. R. (2017). What does explainable AI really mean? A new conceptualization of perspectives. *arXiv preprint arXiv:1710.00794*.

Durán, J. M., & Jongsma, K. R. (2021). Who is afraid of black box algorithms? On the epistemological and ethical basis of trust in medical AI. *Journal of Medical Ethics*.

Durrant, A., Markovic, M., Matthews, D., May, D., Leontidis, G., & Enright, J. (2021). How might technology rise to the challenge of data sharing in agri-food?. *Global Food Security*, *28*, 100493.

Egels-Zandén, N., & Hansson, N. (2016). Supply chain transparency as a consumer or corporate tool: The case of Nudie Jeans Co. *Journal of Consumer Policy*, *39*(4), 377-395.

Egels-Zanden, N., Hulthen, K., & Wulff, G. (2015). Trade-offs in supply chain transparency: The case of Nudie Jeans Co. *Journal of Cleaner Production*, 107, 95–104. https://doi.org/10.1016/j.jclepro.2014.04.074.

EUR-Lex (nd). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) Available at: https://eur-lex.europa.eu/eli/reg/2016/679/oj [Accessed 30 November 2021]

European Commission, (2019). Ethics guidelines for trustworthy AI [WWW Document]. Shap. Eur. Digit. Future - Eur. Comm. URL Available at: https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai [Accessed 30th March 2021).

Friedman, B., & Nissenbaum, H. (1996). Bias in computer systems. *ACM Transactions on Information Systems (TOIS)*, *14*(3), 330-347.

Gardner, T.A., Benzie, M., Börner, J., Dawkins, E., Fick, S., Garrett, R., Godar, J., Grimard, A., Lake, S., Larsen, R.K. & Mardas, N., (2019). Transparency and sustainability in global commodity supply chains. *World Development*, *121*,163-177.

Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., & Kagal, L. (2018, October). Explaining explanations: An overview of interpretability of machine learning. In *2018 IEEE 5th International Conference on data science and advanced analytics (DSAA)* (pp. 80-89). IEEE.

Gregor, S., & Benbasat, I. (1999). Explanations from intelligent systems: Theoretical foundations and implications for practice. *MIS quarterly*, 497-530.

Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G. Z. (2019). XAI—Explainable artificial intelligence. *Science Robotics*, *4*(37). DOI: 10.1126/scirobotics.aay7120

Gutierrez, A., Boukrami, E. and Lumsden, R. (2015) Technological, Organisational and Environmental factors influencing managers’ decision to adopt cloud computing in the UK. *Journal of Enterprise Information Management*, 28(6), pp. 788-807. Available at: https://doi.org/10.1108/JEIM-01-2015-0001

Hoelzer, K., Moreno Switt, A.I., Wiedmann, M., & Boor, K.J., (2018). Emerging needs and opportunities in foodborne disease detection and prevention: From tools to people. *Food Microbiol., New Tools to Detect and Prevent Foodborne Outbreaks from Farm to Fork* 75, 65–71. https://doi.org/10.1016/j.fm.2017.07.006

Hofstede, G.J. Spaans, H., Schepers, H., Trienekens, J.H., & Beulens A.J.M., (2004). Hide or Confide: the Dilemma of Transparency, Reed Business Information, The Netherlands, 2004.

House of Lords Select Committee. (2018). AI in the UK: ready, willing and able." House of Lords 36 (2018). London

Islam, S., & Cullen, J. M. (2021). Food traceability: A generic theoretical framework. *Food Control*, 107848.

ISO 22000: 2018 Food safety management systems — Requirements for any organisation in the food chain. Available at: https://www.iso.org/obp/ui/#iso:std:iso:22000:ed-2:v1:en

Jackson, R. (2018). Unpacking the ethics of food sustainability: health, harmony and beyond. Available at: https://www.nuffieldbioethics.org/blog/unpacking-ethics-food-sustainability-health-harmony [Accessed 04 April 2021]

Jacobs, N., Brewer, S., Craigon, P.J., Frey, J., Gutierrez, A., Kanza, S., Manning, L., Munday, S., Pearson, S. & Sacks, J., (2021). Considering the ethical implications of digital collaboration in the Food Sector. *Patterns*, *2*(11), p.100335.

Japanese Society for Artificial Intelligence (JSAI) Ethical Guidelines (2017). Available at: http://ai-elsi.org/wp-content/uploads/2017/05/JSAI-Ethical-Guidelines-1.pdf

Jobin, A., Ienca, M. & Vayena, E. (2019). The global landscape of AI ethics guidelines, *Nature Machine Intelligence*, 1(9), 389-399, https://doi.org/10.1038/s42256-019-0088-2

Kakani, V., Nguyen, V. H., Kumar, B. P., Kim, H., & Pasupuleti, V. R. (2020). A critical review on computer vision and artificial intelligence in food industry. *Journal of Agriculture and Food Research*, *2*, 100033.

Khaleghi, B. (2019). The What of Explainable AI. Available at: https://www.elementai.com/news/2019/the-what-of-explainable-ai

Kiran, A., Narayana Raj, G., & Talawar, M. B. (2020). Food Crisis-How Artificial Intelligence and Machine Learning Are Solving Humanity's Greatest Challenge—A Review. *Journal of Computational and Theoretical Nanoscience*, *17*(9-10), 3839-3843.

Köksal, Ö., & Tekinerdogan, B. (2019). Architecture design approach for IoT-based farm management information systems. *Precision Agriculture*, *20*(5), 926-958.

Koppell, J. G. (2005). Pathologies of accountability: ICANN and the challenge of "multiple accountabilities disorder". *Public administration review*, *65*(1), 94-108.

Kostina, A. V., & Khorina, G. P. (2012). Information Culture in the Concepts of Information Societies. *Philosophy and Culture*, (4), 14-20.

Kosior, K. (2020). Economic, Ethical and Legal Aspects of Digitalization in the Agri-Food Sector. *Zagadnienia Ekonomiki Rolnej/Problems of Agricultural Economics*. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3712102 [Accessed on 05.04.2021]

Kowalska, A., & Manning, L. (2021). Using the rapid alert system for food and feed: potential benefits and problems on data interpretation. *Critical reviews in food science and nutrition*, 6, 906-919

Kraak, V. I., Swinburn, B., Lawrence, M., & Harrison, P. (2014). An accountability framework to promote healthy food environments. *Public health nutrition*, *17*(11), 2467-2483.

Kurz-Milcke, E., Gigerenzer, G., & Martignon, L. (2008). Transparency in risk communication: graphical and analog tools. In *Annals of the New York Academy of Sciences* (pp. 18-28). Blackwell

Larsson, S., & Heintz, F., (2020). Transparency in artificial intelligence. *Internet Policy Rev*. 9

Leal, M.C., Pimentel, T., Ricardo, F., Rosa, R., & Calado, R., (2015). Seafood traceability: current needs, available tools, and biotechnological challenges for origin certification. *Trends in Biotechnology*. 33, 331–336. https://doi.org/10.1016/j.tibtech.2015.03.003

Lipper, L., Thornton, P., Campbell, B.M., Baedeker, T., Braimoh, A., Bwalya, M., Caron, P., Cattaneo, A., Garrity, D., Henry, K. & Hottle, R., (2014). Climate-smart agriculture for food security. *Nature climate change*, *4*(12), 1068-1072.

Lipton, Z. C. (2018). The mythos of model interpretability. *Queue*, *16*(3), 31-57.

Long, T. B., Blok, V., & Coninx, I. (2016). Barriers to the adoption and diffusion of technological innovations for climate-smart agriculture in Europe: evidence from the Netherlands, France, Switzerland and Italy. *Journal of Cleaner Production*, *112*, 9-21.

Malhotra, Y. (2001). Expert systems for knowledge management: crossing the chasm between information processing and sense making. *Expert Systems with Applications*, *20*(1), 7-16

Maloni, M. J., & Brown, M. E. (2006). Corporate social responsibility in the supply chain: an application in the food industry. *Journal of business ethics*, 68(1), 35-52.

Manning, L. (2018). Systems for sustainability and transparency of food supply chains. In *Sustainable food systems from agriculture to industry* (pp. 153-187). Academic Press

Martin, K. (2019). Ethical implications and accountability of algorithms. *Journal of Business Ethics*, *160*(4), 835-850.

Mark, R. (2019). Ethics of using AI and big data in agriculture: The case of a large agriculture multinational. *The ORBIT Journal*, *2*(2), 1-27

Mepham, T.B. (2000). The role of food ethics in food policy. *Proceedings of the Nutrition Society*, *59*(4), 609-618.

Mehmet, S. (2020). FAO signs ethical resolution on AI-food applications. Available at: https://www.newfoodmagazine.com/news/106712/fao-signs-ethical-resolution-on-ai-food-applications

Mkrttchian, V. (2021). Artificial and Natural Intelligence Techniques as IoP-and IoT-Based Technologies for Sustainable Farming and Smart Agriculture. In *Artificial Intelligence and IoT-Based Technologies for Sustainable Farming and Smart Agriculture* (pp. 40-53). IGI Global.

Mol, A.P.J. (2015). Transparency and value chain sustainability, *Journal of Cleaner Production*, 107, 154-161

Montreal Declaration (2018). Available at: https://www.montrealdeclaration-responsibleai.com/the-declaration

Nissenbaum, H. (1996). Accountability in a computerized society. *Science and engineering ethics*, *2*(1), 25-42.

Peters, D., Vold, K., Robinson, D., & Calvo, R. A. (2020). Responsible AI—Two Frameworks for Ethical Design Practice, in IEEE Transactions on Technology and Society, 1(1) , 34-47, doi: 10.1109/TTS.2020.2974991.

Pizzuti, T., & Mirabelli, G., (2015). The Global Track and Trace System for food: General framework and functioning principles, *Journal of Food Engineering*, 159, 16-35

Qian, J., Ruiz-Garcia, L., Fan, B., Villalba, J.I.R., McCarthy, U., Zhang, B., Yu, Q. & Wu, W., (2020). Food traceability system from governmental, corporate, and consumer perspectives in the European Union and China: A comparative review. *Trends in Food Science & Technology*, *99*, 402-412.

Rees, W., Tremma, O., & Manning, L. (2019). Sustainability cues on packaging: the influence of recognition on purchasing behavior. *Journal of Cleaner Production*, *235*, 841-853.

Regulation EC/178/2002 laying down the general principles and requirements of food safety law, establishing the European Food Standards Agency and laying down procedures in matters of food safety OJ L/31 1.2.2002 pp. 001 – 024

Romecall (2020) Rome Call. AI Ethics. Available at: https://romecall.org/

Rosenfeld, A., & Richardson, A. (2019). Explainability in human–agent systems. *Autonomous Agents and Multi-Agent Systems*, *33*(6), 673-705.

Royal Society (2019). Explainable AI: The basics. Policy Briefing. The Royal Society. ISBN: 978-1-78252-433-5 London https://royalsociety.org/-/media/policy/projects/explainable-ai/AI-and-interpretability-policy-briefing.pdf

Saunders, L., & Corning, S. (2020). From cooperation to collaboration: Toward a framework for deepening library partnerships. *Journal of Library Administration*, *60*(5), 453-469.

Schewe, R. L., & Stuart, D. (2015). Diversity in agricultural technology adoption: How are automatic milking systems used and to what end?. *Agriculture and human values*, *32*(2), 199-213.

Setzu, M., Guidotti, R., Monreale, A., Turini, F., Pedreschi, D., & Giannotti, F. (2021). GLocalX-From Local to Global Explanations of Black Box AI Models. *Artificial Intelligence*, *294*, 103457.

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of management review*, *20*(3), 571-610.

Suddaby, R., Bitektine, A., & Haack, P. (2017). Legitimacy. *Academy of Management Annals*, *11*(1), 451-478.

Suzor, N. P., West, S. M., Quodling, A., & York, J. (2019). What do we mean when we talk about transparency? Toward meaningful transparency in commercial content moderation. *International Journal of Communication*, *13*, 18.

The Guide to the General Data Protection Regulation (GDPR) (2018) Available at: https://www.gov.uk/government/publications/guide-to-the-general-data-protection-regulation

Turilli, M., & Floridi, L. (2009). The ethics of information transparency, *Ethics and Information Technology,* 11(2), 105-112 van der Burg, S., Wiseman, L., & Krkeljas, J. (2020). Trust in farm data sharing: reflections on the EU code of conduct for agricultural data sharing. *Ethics and Information Technology*, 1-14.

Von Schomberg, R. (Ed.). (2011). *Towards Responsible Research and Innovation in the Information and Communication Technologies and Security Technologies Fields*. p9 Luxembourg: Publication Office of the European Union. Retrieved from http://ec.europa.eu/research/science-society/document_library/pdf_06/mep-rapport-2011_en.pdf

Wellard, L., Glasson, C., Chapman, K., & Miller, C., (2011). Fast Facts: the availability and accessibility of nutrition information in fast food chains. Health Promot. J. Austr. 22, 184–188. https://doi.org/10.1071/he11184

Wiley, H.S., & Michaels, G.S., (2004). Should software hold data hostage? Nat. Biotechnol. 22, 1037–1038. https://doi.org/10.1038/nbt0804-1037

Wilkinson, M.D., Dumontier, M., Aalbersberg, Ij.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L.B., Bourne, P.E., … Mons, B., (2016). The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3, 160018. https://doi.org/10.1038/sdata.2016.18

Winfield, A. F., Booth, S., Dennis, L. A., Egawa, T., Hastie, H., Jacobs, N., ... & Watson, E. (2021). IEEE P7001: a proposed standard on transparency. Frontiers in Robotics and AI, 225.

Wognum, P. N., Bremmers, H., Trienekens, J. H., van der Vorst, J. G., & Bloemhof, J. M., (2011). Systems for sustainability and transparency of food supply chains–Current status and challenges. Advanced Engineering Informatics, 25(1), 65-76.

Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. (2017). Big data in smart farming–a review. *Agricultural systems*, *153*, 69-80.

Yang, C., Feng, Y., & Whinston, A. (2021). Dynamic Pricing and Information Disclosure for Fresh Produce: An Artificial Intelligence Approach. *Production and Operations Management*. https://doi.org/10.1111/poms.13525

Zhou, H., Shou, Y., Zhai, X., Li, L., Wood, C., & Wu, X., (2014). Supply chain practice and information quality: A supply chain strategy study, International Journal of Production Economics, 147, 624-633

Zhu, S., Song, J., Hazen, B. T., Lee, K., & Cegielski, C. (2018). How supply chain analytics enables operational supply chain transparency. *International Journal of Physical Distribution & Logistics Management*. 48(1), 47-68

Table 1. Aspects of AI use in the food supply chain: innate characteristics of aspects and corporate/supply chain mechanisms and activities to address these aspects.

Aspect	Inherent characteristics	Corporate mechanisms/ activities	Supply Chain Example	Section of the discourse where the aspect is considered
Transparency	Visible, open (opacity), accurate, relevant, reliable, timely.	Data/information disclosure; transparency cues; validity mechanisms; standardisation, simplification, reduction, dissemination, certification.	Open sharing of data in the supply chain to develop an ‘end-to-end’ allergen control system.	Section 3.1
Traceability	Identity, movement, location, transactional, information loss	Tracing, following, tracking, record keeping.	The use of a scanning system and barcodes on pack to trace an ingredient from source (farm) to a factory.	Section 3.1
Explainability	XAI, knowledge based.	Giving meaning, creating understanding, reconciling differences.	The ability to explain the technology to a range of stakeholders so they understand how it is operating in practice e.g., yield prediction software in orchards.	Section 3.2
Interpretability	Answerability, explicit, visible.	Information assimilation and interpretation, use of tools, prototype analysis, feature analysis.	The ability to interpret the output of the technology so that it can inform decision-making; for example, being able to use scanning technology and translating the output into information on the level of lameness in a dairy herd.	Section 3.3
Accessibility	Usable, findable, reusable, interoperable, private (protected access), public. (open access)	Information provision, authorisation protocols, privacy protocols, human (inclusive) accessibility protocols.	The development of access rights with robotic milking machines on farm so that the farmer, veterinarians, machine manufacturers, dairy customer have appropriate access to data collected.	Section 3.4
Accountability	Duty, obligation, liability, controllability, responsiveness.	Corporate justification, governance; accountability protocols.	The development of a data governance protocol that identifies the uses of data by different stakeholders and defines who specific data can and cannot be shared with, for example data associated with workers in a food factory.	Section 3.5
Responsibility	Trust, legitimacy.	Corporate social responsibility. AI design protocols that define roles and responsibility.	The use of a food safety management tool that has an inbuilt alert system according to the level of responsibility in the factory.	Section 3.6