

ORIGINAL RESEARCH

Open access

Responsible AI Governance in Agri-Food Supply Chains: A Framework for Transparency, Accountability, Traceability, and Human Oversight

Victor Dupont^{1*}, Elise Morel¹

Abstract

Artificial intelligence is increasingly being embedded in agri-food supply chains through applications such as automated quality grading, risk scoring, demand forecasting, logistics optimisation, pricing analytics, and food safety alerts. These systems promise efficiency, resilience, and improved decision support, but they also reshape how decisions are made and who is affected by them. In food systems, AI governance is not only a technical issue but also a matter of safety, fairness, market access, and public trust. Existing responsible AI principles provide important normative foundations, including fairness, transparency, accountability, human oversight, and social benefit. However, these principles are often framed at a general level and do not fully address the distinctive structure of agri-food supply chains. Food chains are fragmented across farms, processors, logistics providers, retailers, technology platforms, regulators, and consumers, which makes governance more complex than in single-organisation AI deployment. This article develops a domain-specific Responsible AI Governance Framework for agri-food supply chains. The framework integrates four pillars: transparency, accountability, traceability, and human oversight. It argues that responsible AI in food must make algorithmic decisions visible, assign responsibility across distributed actors, preserve auditable records, and ensure meaningful human safeguards for high-risk decisions. The proposed framework specifies transparency requirements for different AI applications, maps accountability responsibilities across supply chain actors, connects traceability systems to auditability, and defines human oversight mechanisms for safety-critical and livelihood-sensitive decisions. Three tables operationalise the framework through transparency requirements, accountability allocation, and integrated governance mechanisms. The article concludes that responsible AI in food requires governance that is as distributed, interconnected, and safety-conscious as the supply chains it seeks to guide.

Keywords Responsible AI, Agri-food supply chain, Traceability, Governance, Accountability, Transparency

*Correspondence:

Victor Dupont

victor.dupont@outlook.com

¹ Department of Digital Agriculture and Sustainability, Faculty of Agricultural Sciences, University of Montpellier, Montpellier, France

Introduction

Artificial intelligence is moving from experimental deployment into routine decision-making across agri-food supply chains, including production planning, quality control, logistics coordination, fraud detection, and market optimisation. This transition is occurring alongside the global rise of AI ethics and governance principles that emphasise transparency, accountability, safety, human agency, and social benefit [1]. The central governance question is whether such generic principles are sufficient for food systems, where algorithmic decisions can affect public health, farmer livelihoods, consumer choice, and environmental outcomes.

Agri-food supply chains differ from many other AI deployment contexts because they are materially fragmented, temporally sensitive, and safety-critical. Decisions about contamination risk, product rejection, storage conditions, quality classification, and logistics routing often involve multiple actors who do not share the same data, incentives, or authority. Research on digital agriculture has therefore argued that responsible AI requires attention not only to model performance but also to systemic risks, externalities, and institutional arrangements [2, 3].

The need for a tailored governance approach is reinforced by the fact that food supply chains increasingly depend on data infrastructures that connect farms, processors, distributors, retailers, and consumers. Digital collaboration can improve coordination, but it also raises ethical concerns about data control, unequal bargaining power, and the visibility of automated decisions [4]. In this context, responsible AI governance must address both the technical explainability of systems and the organisational distribution of responsibility across supply chain actors.

This article proposes a domain-specific Responsible AI Governance Framework for agri-food supply chains built around four pillars: transparency, accountability, traceability, and human oversight. Transparency concerns the visibility of AI-driven decisions and the information needed by affected actors; accountability concerns responsibility allocation across the chain; traceability concerns the creation of auditable decision records; and human oversight concerns safeguards, review, and intervention when AI affects safety or livelihoods. These pillars extend general AI ethics principles into a practical governance architecture suited to multi-actor food systems [5, 6].

Responsible AI in Agri-Food Supply Chains

Responsible AI in agri-food supply chains can be defined as the design, deployment, monitoring, and governance of AI systems in ways that protect safety, fairness, sustainability, explainability, and human agency across the full food chain. Unlike narrow technical approaches that focus mainly on accuracy or optimisation, responsible AI requires attention to who benefits, who is exposed to risk, and who has the power to question or override automated decisions. In digital agriculture, this means that predictive tools must be evaluated not only by their efficiency gains but also by their consequences for farmers, processors, retailers, consumers, and ecosystems [2, 3].

Generic AI governance frameworks are valuable because they articulate broad principles such as beneficence, non-maleficence, autonomy, justice, and explicability. However, principles alone do not guarantee ethical AI because they require translation into institutional procedures, documentation duties, review mechanisms, and enforceable accountability practices [5, 6]. This translation challenge is especially acute in agri-food supply chains, where the actor that develops a system may not be the actor that uses it, and the actor affected by the output may have limited access to the underlying data or decision logic.

Food-specific AI governance must also recognise that agri-food systems combine operational, commercial, regulatory, and public health objectives. Ethical reflection in the food sector has emphasised the need for a shared language around AI adoption because technology providers, producers, processors, retailers, and regulators often understand risk and responsibility differently [7]. Without such shared governance language, AI systems may be adopted as efficiency tools while their effects on transparency, contestability, and accountability remain underdeveloped.

A domain-specific governance framework is therefore needed because agri-food AI operates within a layered decision environment rather than a single organisational boundary. Existing AI ethics guidelines identify important principles, but evaluations of such guidelines show uneven treatment of implementation, enforcement, and accountability [1, 8]. For food systems, responsible AI must be operationalised through visible decision standards, actor-specific accountability, auditable traceability, and human safeguards that match the risk profile of each application.

Transparency and Decision Visibility

Transparency in agri-food AI is not simply a matter of publishing model documentation or making algorithms technically explainable. It concerns whether affected actors can understand the basis, timing, data inputs, and consequences of AI-driven decisions in forms that are meaningful for their role in the supply chain. Explainability becomes especially important when AI systems classify food quality, flag safety risks, detect fraud, allocate logistics capacity, or influence pricing outcomes [9, 10].

Different actors require different forms of decision visibility. Farmers may need to know why a shipment was downgraded or rejected, which quality variables were decisive, and whether they can appeal the decision; processors may need to understand why a batch was flagged as unsafe or anomalous; retailers may need visibility into quality scoring and shelf-life predictions; and consumers may need accessible information about provenance, sustainability, or safety claims. Research on supply chain transparency shows that technology can improve visibility, but governance must determine what information is disclosed, to whom, and under what conditions [11].

Transparency also has a power dimension because information asymmetry can reinforce dependency on dominant platforms, processors, or retailers. If a platform controls data aggregation, model design, and decision outputs, downstream or upstream actors may be governed by systems they cannot inspect or challenge. Ethical work on digital collaboration in food systems therefore suggests that transparency must include not only technical explanation but also disclosure of data access rules, decision authority, and institutional responsibilities [4].

Transparency requirements should be proportionate to the risk and impact of the AI application. Low-risk operational forecasting may require internal documentation and periodic review, while safety alerts, contamination risk scores, fraud detection, rejection decisions, or livelihood-sensitive pricing tools require stronger explanation, appeal, and audit duties. **Table 1** defines the transparency requirements for different AI applications in agri-food supply chains.

Table 1. Transparency and Decision Visibility Requirements for AI in Agri-Food Supply Chains: Application Types, Information Needs, and Disclosure Standards

AI application type	Typical agri-food decision	Primary affected actors	Required decision visibility	Minimum disclosure standard	Governance purpose
Quality grading and product classification	Assigning grade, freshness category, defect class, or accept/reject status	Farmers, processors, retailers, consumers	Key quality indicators, image or sensor features used, confidence level, rejection threshold, and comparison with accepted standards	Provide actor-specific explanation at the point of decision, including reason codes and appeal route	Prevent opaque downgrading, support contestability, and reduce unfair market exclusion
Food safety risk scoring	Flagging contamination, allergen, pathogen, spoilage, or unsafe handling risk	Processors, logistics firms, regulators, retailers, consumers	Input data sources, risk factors, severity level, uncertainty level, and required human action	Provide documented risk explanation, escalation pathway, and human review requirement before high-impact action	Protect public health while preventing unjustified disruption or panic
Logistics and routing optimisation	Prioritising shipments, rerouting goods, allocating cold-	Farmers, logistics providers,	Optimisation objective, constraints considered, affected shipment status, and	Provide operational explanation to affected business actors and retain	Make efficiency decisions accountable where they affect

	chain capacity, or delaying delivery	retailers, consumers	reason for priority change	routing records for later audit	freshness, waste, or contractual delivery
Dynamic pricing and market coordination	Setting procurement prices, retail prices, promotional timing, or demand-linked purchasing	Farmers, processors, retailers, consumers	Pricing variables, demand assumptions, market data inputs, and exclusion of prohibited discriminatory variables	Provide understandable explanation of pricing logic to contractual actors and maintain internal audit documentation	Reduce hidden exploitation and support fair commercial negotiation
Fraud and authenticity detection	Flagging suspected substitution, mislabelling, adulteration, or provenance inconsistency	Producers, processors, certifiers, retailers, regulators, consumers	Evidence sources, traceability records used, anomaly score, and grounds for suspicion	Provide auditable explanation before punitive action and allow evidence submission by affected parties	Link AI detection to due process and reliable enforcement
Sustainability and provenance analytics	Scoring emissions, water use, origin, certification status, or environmental claims	Farmers, processors, retailers, consumers, regulators	Data provenance, calculation assumptions, uncertainty, certification linkage, and update frequency	Provide public-facing summary for consumers and detailed documentation for auditors and supply chain partners	Prevent greenwashing and strengthen credible sustainability claims
Platform-based supplier ranking	Ranking farms, processors, or suppliers for procurement, access, or visibility	Farmers, processors, platforms, retailers	Ranking criteria, data categories, weighting logic, performance period, and review mechanism	Provide supplier-facing explanation and correction process for inaccurate data	Prevent opaque exclusion from markets and support procedural fairness
Automated compliance monitoring	Detecting violations of safety, hygiene, labelling, or contractual rules	Processors, retailers, regulators, workers, consumers	Rule basis, evidence record, system confidence, human reviewer identity, and enforcement pathway	Provide complete decision record for regulated or punitive actions	Ensure that AI-supported compliance remains legally and procedurally accountable

Accountability across Supply Chain Actors

Accountability in AI-enabled agri-food supply chains is difficult because decisions are rarely produced by one actor alone. A platform may design the algorithm, a processor may configure the thresholds, a retailer may rely on the output, and a farmer or consumer may bear the consequences. This “many hands” problem is well recognised in algorithmic accountability debates, where responsibility becomes diluted when systems are distributed across designers, deployers, operators, and institutional users [12].

In agri-food supply chains, accountability must therefore be allocated according to decision authority, control over data, ability to intervene, and proximity to harm. Technology providers should be accountable for model design, validation, documentation, and known limitations; supply chain firms should be accountable for how systems are deployed in operational decisions; and regulators should be able to inspect records when AI affects safety, labelling, compliance, or market access. Reflexive governance approaches in food technology adoption show that responsibility must be embedded before deployment rather than repaired after harm occurs [13].

Accountability must also address commercial asymmetry, because powerful platforms, processors, or retailers may impose AI-mediated decisions on smaller suppliers without giving them meaningful routes for explanation or appeal. In such cases, responsible governance requires procedural rights: notice of automated decision-making, access to reasoned explanations, correction of inaccurate data, and review by a competent human authority. Ethical analyses of AI governance emphasise that accountability must move beyond abstract commitment and become a practical chain of answerability, documentation, and remedy [8, 14].

Accountability allocation should be designed around specific decision points rather than general statements of responsibility. A product rejection decision, a contamination alert, a fraud flag, a supplier ranking, and a dynamic pricing decision each require different accountability rules because they create different harms and involve different actors. Table 2 maps accountability challenges and allocation principles across supply chain actors.

Table 2. Accountability Allocation in AI-Enabled Agri-Food Supply Chains: Actor Responsibilities, Decision Points, and Liability Principles

Supply chain actor	AI-related responsibility	Key decision points	Accountability challenge	Allocation principle	Required governance mechanism
Farmers and primary producers	Provide accurate production, quality, sustainability, and provenance data where required	Data submission, sensor validation, corrective action after quality or safety alerts	Farmers may be judged by systems they do not control and may lack power to contest automated decisions	Responsible for data accuracy within their control, but not for opaque model logic or downstream platform decisions	Data correction rights, explanation access, supplier appeal procedures, and proportional evidence duties
Processors and manufacturers	Use AI for quality control, safety screening, batch classification, and compliance monitoring	Batch acceptance, contamination response, allergen control, product recall triggers	Processors may rely on vendor systems while retaining operational responsibility for food safety	Accountable for deployment context, threshold selection, human review, and compliance consequences	Validation records, escalation rules, human sign-off for high-risk decisions, and audit logs
Logistics and cold-chain providers	Apply AI to routing, freshness prediction, temperature risk, and delivery prioritisation	Shipment rerouting, cold-chain exception handling, delivery delay decisions	AI optimisation may shift risk to producers, retailers, or consumers without visible justification	Accountable for operational decisions and for preserving transport-related decision evidence	Routing explanations, temperature record linkage, incident documentation, and service-level review

Retailers and food service firms	Use AI for procurement, pricing, shelf-life management, supplier ranking, and consumer-facing claims	Supplier selection, product removal, pricing adjustment, sustainability labelling	Retailers may transform technical scores into market consequences for suppliers and consumers	Accountable for commercial use of AI outputs and for fairness of supplier-facing decisions	Supplier transparency notices, pricing governance review, provenance verification, and consumer disclosure
Technology platforms and AI vendors	Design, train, test, document, update, and monitor AI systems used across the chain	Model development, deployment configuration, data integration, system updates	Vendors may claim limited liability while shaping decisions across many organisations	Accountable for foreseeable system risks, documentation quality, bias testing, security, and model limitations	Model cards, audit trails, risk assessments, contractual accountability clauses, and post-market monitoring
Certifiers and auditors	Verify compliance, sustainability claims, provenance records, and AI-supported evidence	Certification approval, audit sampling, non-compliance reporting	Auditors may lack access to underlying model logic or complete data trails	Accountable for independent verification standards and evidence sufficiency	Audit access protocols, traceability verification, model documentation review, and conflict-of-interest controls
Regulators and public authorities	Set minimum governance standards and intervene in safety, fraud, and market fairness cases	Inspection, recall, enforcement, safety alert validation, dispute resolution	Regulators may face fragmented data systems and cross-jurisdictional chains	Accountable for oversight capacity, enforceable standards, and public-interest safeguards	Regulatory audit powers, mandatory reporting for high-risk AI, incident registries, and harmonised guidance
Consumers and civil society actors	Use disclosed information to make informed choices and raise concerns about safety, fairness, or sustainability	Purchase decisions, complaint submission, public-interest scrutiny	Consumers are affected by AI decisions but rarely participate in system design	Not liable for AI decisions, but entitled to meaningful disclosure and remedy pathways	Consumer-facing transparency, complaint mechanisms, labelling standards, and public accountability reporting

Traceability and Auditability

Traceability is the evidentiary backbone of responsible AI governance in agri-food supply chains. If transparency explains a decision and accountability assigns responsibility, traceability preserves the record needed to verify whether the explanation is accurate and whether the responsible actor acted properly. Blockchain and other distributed ledger technologies have been widely discussed as mechanisms for strengthening food chain traceability by creating tamper-resistant records of provenance, transactions, and product movement [15-18].

For AI governance, traceability must extend beyond product movement to include decision provenance. An auditable AI record should show what data were used, when the data were captured, which model version generated the output, what threshold was applied, who received the output, what action followed, and whether a human reviewed or overrode the recommendation. This broader understanding connects technological traceability with governance accountability, because blockchain-based supply chain systems can support transparency only when the recorded data are relevant, reliable, and accessible to legitimate auditors [19, 20].

Food traceability also has a trust function because consumers, retailers, regulators, and supply chain partners increasingly rely on digital records to verify safety, authenticity, sustainability, and ethical sourcing. Studies of blockchain-based food traceability show that trust depends not merely on the existence of a traceability system, but on whether users believe the information is credible, usable, and connected to meaningful oversight [21, 22]. Therefore, traceability-as-governance requires standards for data quality, interoperability, access rights, and correction procedures.

Auditability should not be reduced to technical logging, because a complete record is valuable only if it can be interpreted and acted upon. Food fraud detection, authenticity verification, and traceability systems show that AI-supported decisions often require evidence chains that connect sensor data, provenance records, expert judgement, and regulatory standards [10, 23]. In responsible AI governance, auditability means that an independent reviewer can reconstruct the decision pathway and determine whether the system acted within approved rules, whether the data were adequate, and whether human intervention was required.

Human Oversight and Decision Safeguards

Human oversight is the safeguard that prevents responsible AI governance from becoming an automated compliance exercise. In agri-food supply chains, AI can support rapid detection and coordination, but decisions involving contamination risk, allergen management, fraud accusations, supplier exclusion, or product recall should not be left to automated outputs alone. The principle of meaningful human control requires that human actors retain the capacity to understand, supervise, intervene in, and be accountable for consequential automated decisions [24].

The appropriate form of human oversight should vary by risk level. Low-risk applications such as routine demand forecasting may require periodic human monitoring, while medium-risk applications such as logistics prioritisation or shelf-life prediction may require human review when confidence is low or consequences are significant. High-risk applications involving public health, product rejection, regulatory enforcement, or supplier exclusion require stronger human-in-the-loop safeguards before irreversible action is taken [9, 25].

Human oversight must also guard against automation complacency, where decision-makers accept algorithmic outputs because the system appears objective or technically sophisticated. In food safety contexts, this risk is particularly serious because false negatives may expose consumers to harm, while false positives may create unnecessary recalls, waste, or unfair commercial penalties. Responsible AI in agriculture has therefore been framed as requiring systemic understanding of externalities rather than narrow trust in algorithmic performance alone [3].

Meaningful oversight requires institutional design, not just the presence of a human reviewer. Reviewers must have access to explanations, uncertainty information, traceability records, and authority to pause, override, escalate, or request further evidence. Variable autonomy approaches are useful here because they recognise that human control should increase when uncertainty, severity, or social impact increases, rather than remaining fixed across all AI applications [25].

Proposed Governance Framework

The proposed Responsible AI Governance Framework integrates transparency, accountability, traceability, and human oversight as mutually reinforcing pillars rather than separate compliance categories. Transparency makes AI decisions

visible; accountability identifies who must answer for those decisions; traceability preserves the evidence needed for audit; and human oversight provides the safeguard against inappropriate automation. This integrated architecture responds to the limitations of generic AI principles by translating them into food-chain-specific governance mechanisms [2, 7].

The framework is designed for fragmented supply chains in which multiple actors share data and consequences but do not share equal authority. Farmers, processors, logistics providers, retailers, platforms, certifiers, and regulators must each have defined governance roles that correspond to their control over data, models, decisions, and remedies. This allocation reflects broader debates in AI ethics showing that principles must be connected to concrete institutional responsibilities if they are to guide real-world deployment [1, 26].

Figure 1 illustrates how the four pillars of responsible AI governance operate across fragmented agri-food supply chain actors.

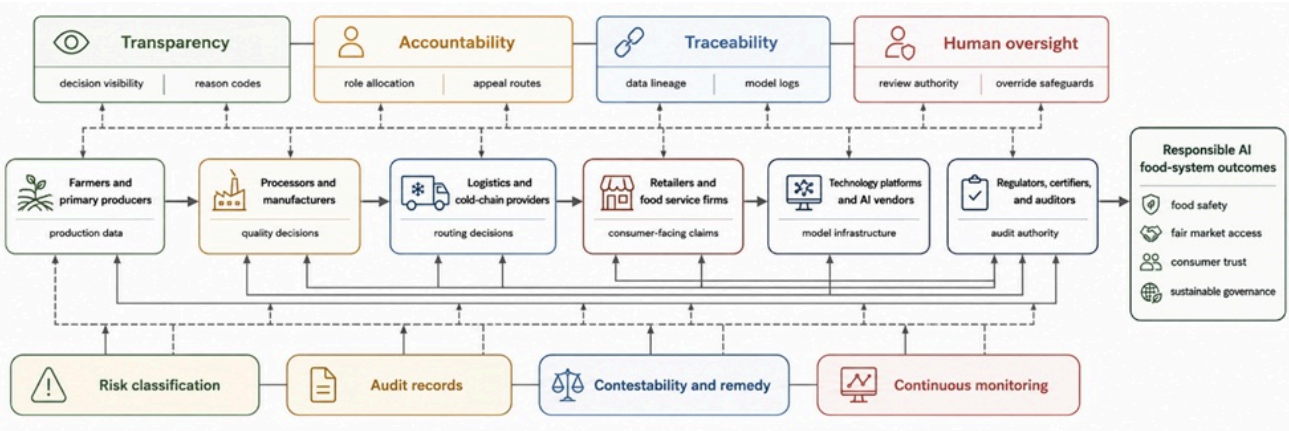


Figure 1. Integrated Responsible AI Governance Architecture for Agri-Food Supply Chains: Linking Transparency, Accountability, Traceability, and Human Oversight Across Multi-Actor Food Systems

Operationally, the framework requires risk classification of AI applications before deployment, decision visibility standards at the point of use, traceability records for auditable reconstruction, and mandatory human review for safety-critical or livelihood-sensitive decisions. It also requires governance documentation across the system life cycle, including data provenance, model validation, threshold approval, incident reporting, and appeal outcomes. **Table 3** presents the proposed Responsible AI Governance Framework integrating the four pillars.

Table 3. Proposed Responsible AI Governance Framework for Agri-Food Supply Chains: Pillars, Principles, and Operational Mechanisms

Governance pillar	Core governance principle	Agri-food-specific interpretation	Operational mechanisms	Primary responsible actors	Evidence required for audit
Transparency	Affected actors must understand consequential AI decisions in role-appropriate terms	Farmers, processors, retailers, regulators, and consumers require different forms of visibility into quality, safety,	Decision notices, reason codes, model documentation, uncertainty disclosure, threshold	AI vendors, processors, retailers, platforms, certifiers, regulators	Explanation records, disclosure logs, model documentation, communication records, and appeal notices

		pricing, provenance, and sustainability decisions	explanation, and consumer-facing summaries		
Accountability	Responsibility must follow control over data, models, deployment, and consequences	No actor should be able to hide behind the complexity of a multi-actor supply chain when AI causes harm or unfair exclusion	Contractual responsibility clauses, role-based accountability maps, liability triggers, decision ownership registers, and escalation duties	Platforms, processors, retailers, AI vendors, logistics firms, regulators	Responsibility matrix, deployment approvals, system ownership records, incident reports, and corrective action files
Traceability	AI decisions must be reconstructable through reliable data and decision records	Product provenance, data provenance, model provenance, and action provenance must be linked across the supply chain	Blockchain or interoperable ledgers, audit logs, model versioning, data lineage records, sensor validation, and access-controlled records	Technology providers, supply chain firms, auditors, certifiers, regulators	Data lineage, model version logs, transaction records, sensor records, human review records, and audit findings
Human oversight	Humans must retain meaningful authority over high-impact AI decisions	Food safety, allergen, fraud, recall, rejection, and supplier exclusion decisions require human review and override capacity	Human-in-the-loop review, on-the-loop monitoring, override protocols, appeal mechanisms, reviewer training, and escalation rules	Processors, retailers, regulators, certifiers, operational managers	Reviewer identity, review rationale, override record, escalation record, and appeal outcome
Risk proportionality	Governance intensity must match decision severity and reversibility	Safety-critical and livelihood-sensitive AI applications require stronger governance than routine forecasting or internal optimisation	Risk tiering, pre-deployment assessment, impact assessment, monitoring frequency, and mandatory review thresholds	Organisations deploying AI, regulators, auditors, governance committees	Risk classification file, assessment report, approval record, monitoring logs, and reassessment schedule
Contestability and remedy	Affected actors must be able to challenge consequential AI outcomes	Farmers and suppliers must be able to contest rejection, ranking, pricing, fraud, or compliance decisions that affect market access	Appeals process, evidence submission, data correction, independent review, and remedy timelines	Processors, retailers, platforms, certifiers, regulators	Appeal file, correction record, reviewer decision, remedy action, and communication history
Interoperability and standards	Governance must operate across organisational	Food chains cross firms, regions, certification	Common data schemas, open traceability	Standards bodies, regulators,	Standard compliance records,

	and jurisdictional boundaries	schemes, and regulatory systems, requiring shared data and audit standards	standards, audit APIs, regulatory reporting templates, and cross-border protocols	platforms, industry consortia, auditors	interoperability tests, audit access logs, and reporting records
Continuous monitoring	Responsible AI governance must continue after deployment	Models, data quality, supply chain conditions, and risk profiles change over time	Post-deployment monitoring, drift detection, incident review, periodic audit, model update governance, and stakeholder feedback	AI vendors, deploying organisations, auditors, regulators	Monitoring reports, drift alerts, incident logs, model update records, and governance review minutes

The framework’s central contribution is to treat responsible AI as a supply chain governance problem rather than a stand-alone technology ethics problem. It links AI ethics principles to existing food governance concerns such as safety, authenticity, traceability, certification, and public trust. By doing so, it creates a practical architecture in which responsible AI is embedded in everyday supply chain decision-making rather than added as a peripheral compliance label [11, 15, 27].

Implementation Risks and Managerial Recommendations

Implementation will face significant barriers because traceability infrastructure is costly, data standards remain fragmented, and powerful supply chain actors may resist transparency that reduces their discretion. Blockchain and traceability research shows that technical systems can improve visibility, but they also require coordination, interoperability, and credible data input across actors with different incentives [19, 28]. Managers should therefore begin with high-risk AI applications, such as safety alerts, product rejection, fraud detection, supplier ranking, and consumer-facing sustainability claims, where governance benefits are greatest.

A second implementation risk is regulatory fragmentation across jurisdictions, certification systems, food categories, and platform business models. Agri-food supply chains often cross national borders, while AI governance principles and food safety obligations may be interpreted differently by regulators, auditors, and firms. Managers should respond by building governance processes that exceed minimum compliance, including documentation, audit readiness, appeal procedures, and human review standards aligned with broad responsible AI expectations [1, 26].

Figure 2 shows how an AI-enabled agri-food decision can be made transparent, accountable, traceable, and subject to human review.

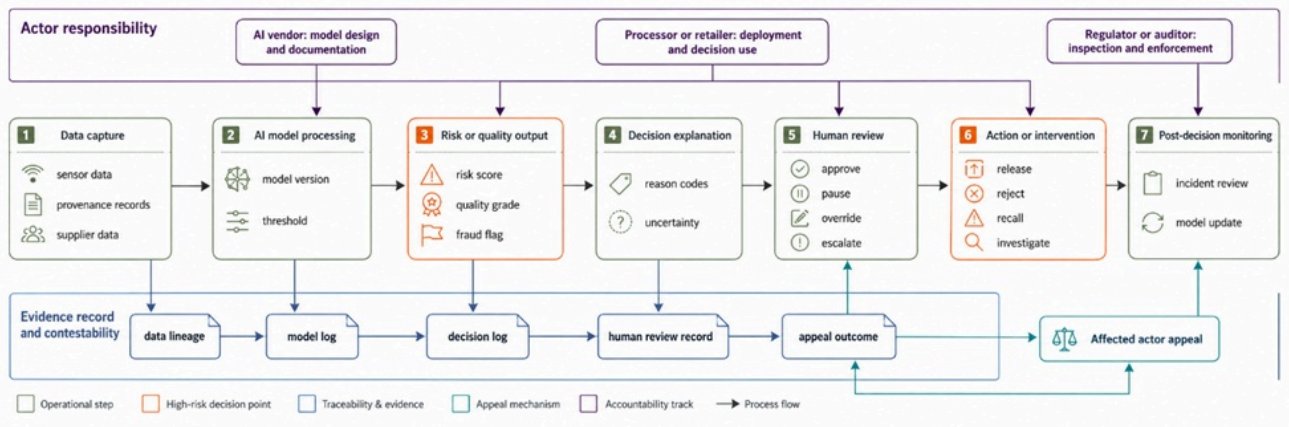


Figure 2. Governance Workflow for High-Risk AI Decisions in Agri-Food Supply Chains: From Data Capture and Algorithmic Output to Audit, Appeal, and Human Oversight

A third risk is that governance becomes symbolic rather than operational, especially if organisations adopt ethics language without changing decision routines. To avoid this, firms should establish multi-stakeholder governance bodies that include technical experts, food safety professionals, supply chain managers, legal specialists, farmer or supplier representatives, and consumer-facing accountability roles. Responsible innovation in food requires reflexive governance that continually reviews impacts, redistributes learning, and updates safeguards as technologies and supply chain conditions change [4, 13].

Conclusion

This article has proposed a Responsible AI Governance Framework tailored to the distinctive conditions of agri-food supply chains. Its central argument is that generic AI governance principles are necessary but insufficient where decisions are distributed across fragmented, multi-actor, safety-critical, and commercially unequal food systems. A domain-specific framework is required because food-related AI affects not only operational efficiency but also health, livelihoods, fairness, sustainability, and trust.

The framework contributes by integrating four pillars: transparency, accountability, traceability, and human oversight. Transparency ensures that affected actors can understand consequential decisions; accountability assigns responsibility across the chain; traceability supplies the evidence base for audit; and human oversight protects against inappropriate automation in high-risk contexts. Together, these pillars create a governance model that matches the distributed structure of the food systems it seeks to regulate.

Future work should pilot the framework in real supply chain settings, including quality grading, safety monitoring, fraud detection, supplier ranking, and sustainability claims. Regulators, industry bodies, technology providers, and supply chain actors should collaborate on open standards, audit protocols, and appeal mechanisms. Responsible AI should become a normal condition of participation in agri-food systems, not an exceptional feature added after harm occurs.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

References

- Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. *Nat Mach Intell*. 2019;1(9):389-99.
<https://doi.org/10.1038/s42256-019-0088-2>.
- Dara R, Hazrati Fard SM, Kaur J. Recommendations for ethical and responsible use of artificial intelligence in digital agriculture. *Front Artif Intell*. 2022;5:884192.
<https://doi.org/10.3389/frai.2022.884192>.
- Tzachor A, Devare M, King B, Avin S, Ó hÉigeartaigh S. Responsible artificial intelligence in agriculture requires systemic understanding of risks and externalities. *Nat Mach Intell*. 2022;4(2):104-9.
<https://doi.org/10.1038/s42256-022-00440-4>.
- Jacobs N, Brewer S, Craigon P, Frey J, Gutierrez A, Kanza S, et al. Considering the ethical implications of digital collaboration in the food sector. *Patterns (N Y)*. 2021;2(11):100335.
<https://doi.org/10.1016/j.patter.2021.100335>.
- Floridi L, Cowls J, Beltrametti M, Chatila R, Chazerand P, Dignum V, et al. AI4People—an ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds Mach (Dordr)*. 2018;28(4):689-707.
<https://doi.org/10.1007/s11023-018-9482-5>.
- Mittelstadt B. Principles alone cannot guarantee ethical AI. *Nat Mach Intell*. 2019;1(11):501-7.
<https://doi.org/10.1038/s42256-019-0114-4>.
- Manning L, Brewer S, Craigon PJ, Frey J, Gutierrez A, Jacobs N, et al. Artificial intelligence and ethics within the food sector: Developing a common language for technology adoption across the supply chain. *Trends Food Sci Technol*. 2022;125:33-42.
<https://doi.org/10.1016/j.tifs.2022.04.025>.
- Hagendorff T. The ethics of AI ethics: An evaluation of guidelines. *Minds Mach (Dordr)*. 2020;30(1):99-120.
<https://doi.org/10.1007/s11023-020-09517-8>.
- Kudashkina K, Corradini MG, Thirunathan P, Yada RY, Fraser EDG. Artificial intelligence technology in food safety: A behavioral approach. *Trends Food Sci Technol*. 2022;123:376-81.
<https://doi.org/10.1016/j.tifs.2022.03.021>.
- Buyuktepe O, Catal C, Kar G, Bouzembrak Y, Marvin H, Gavai A. Food fraud detection using explainable artificial intelligence. *Expert Syst*. 2025;42(1):e13387.
<https://doi.org/10.1111/exsy.13387>.
- Astill J, Dara RA, Campbell M, Farber JM, Fraser EDG, Sharif S, et al. Transparency in food supply chains: A review of enabling technology solutions. *Trends Food Sci Technol*. 2019;91:240-7.
<https://doi.org/10.1016/j.tifs.2019.07.024>.
- Martin K. Ethical implications and accountability of algorithms. *J Bus Ethics*. 2019;160(4):835-50.
<https://doi.org/10.1007/s10551-018-3921-3>.

Manning L, Brewer S, Craigon P, Frey J, Gutierrez A, Jacobs N, et al. Reflexive governance architectures: Considering the ethical implications of autonomous technology adoption in food supply chains. *Trends Food Sci Technol.* 2023;133:114-26.
<https://doi.org/10.1016/j.tifs.2023.01.015>.

Mökander J, Morley J, Taddeo M, Floridi L. Ethics-based auditing of automated decision-making systems: Nature, scope, and limitations. *Sci Eng Ethics.* 2021;27(4):44.
<https://doi.org/10.1007/s11948-021-00319-4>.

Kamilaris A, Fonts A, Prenafeta-Boldú FX. The rise of blockchain technology in agriculture and food supply chains. *Trends Food Sci Technol.* 2019;91:640-52.
<https://doi.org/10.1016/j.tifs.2019.07.034>.

Rossi S, Gemma S, Borghini F, Perini M, Butini S, Carullo G, et al. Agri-food traceability today: Advancing innovation towards efficiency, sustainability, ethical sourcing, and safety in food supply chains. *Trends Food Sci Technol.* 2025;163:105154.
<https://doi.org/10.1016/j.tifs.2025.105154>.

Kamble SS, Gunasekaran A, Sharma R. Modeling the blockchain enabled traceability in agriculture supply chain. *Int J Inf Manage.* 2020;52:101967.
<https://doi.org/10.1016/j.ijinfomgt.2019.05.023>.

Iftekhar A, Cui X, Hassan M, Afrianto E. Blockchain-based traceability system that ensures food safety measures to protect consumer safety and COVID-19 free supply chains. *Foods.* 2021;10(6):1289.
<https://doi.org/10.3390/foods10061289>.

Saberi S, Kouhizadeh M, Sarkis J, Shen L. Blockchain technology and its relationships to sustainable supply chain management. *Int J Prod Res.* 2019;57(7):2117-35.
<https://doi.org/10.1080/00207543.2018.1533261>.

Treiblmaier H. The impact of the blockchain on the supply chain: A theory-based research framework and a call for action. *Supply Chain Manag.* 2018;23(6):545-59.
<https://doi.org/10.1108/SCM-01-2018-0029>.

Garaus M, Treiblmaier H. The influence of blockchain-based food traceability on retailer choice: The mediating role of trust. *Food Control.* 2021;129:108082.
<https://doi.org/10.1016/j.foodcont.2021.108082>.

Sander F, Semeijn J, Mahr D. The acceptance of blockchain technology in meat traceability and transparency. *Br Food J.* 2018;120(9):2066-79.
<https://doi.org/10.1108/BFJ-07-2017-0365>.

Biglia A, Barge P, Tortia C, Comba L, Aimonino DR, Gay P. Artificial intelligence to boost traceability systems for fraud prevention in the meat industry. *J Agric Eng.* 2022;53(4):1328.
<https://doi.org/10.4081/jae.2022.1328>.

Santoni de Sio F, van den Hoven J. Meaningful human control over autonomous systems: A philosophical account. *Front Robot AI.* 2018;5:15.
<https://doi.org/10.3389/frobt.2018.00015>.

Methnani L, Aler Tubella A, Dignum V, Theodorou A. Let me take over: Variable autonomy for meaningful human control. *Front Artif Intell.* 2021;4:737072.
<https://doi.org/10.3389/frai.2021.737072>.

Cath C, Wachter S, Mittelstadt B, Taddeo M, Floridi L. Artificial intelligence and the “Good Society”: The US, EU, and UK approach. *Sci Eng Ethics*. 2018;24(2):505-28.
<https://doi.org/10.1007/s11948-017-9901-7>.

Francisco K, Swanson D. The supply chain has no clothes: Technology adoption of blockchain for supply chain transparency. *Logistics*. 2018;2(1):2.
<https://doi.org/10.3390/logistics2010002>.

Pandey V, Pant M, Snasel V. Blockchain technology in food supply chains: Review and bibliometric analysis. *Technol Soc*. 2022;69:101954.
<https://doi.org/10.1016/j.techsoc.2022.101954>.