# Results and Discussion

## In this guide

In this guide

1. Food Standards Agency Science Council Report of Project ‘Artificial Intelligence Applications in Food Safety and Authenticity’
2. Authors, Acknowledgements and Declarations of Interest
3. Executive Summary
4. Recommendations to the FSA
5. Introduction
6. Methodology
7. Results and Discussion
8. Conclusions
9. Appendix A: Workshop Case Study Briefing Document
10. Appendix B: Workshop Participant List
11. Appendix C: Workshop Case Study Responses
12. References

This section is structured around the major themes that arose from the workshop and subsequent deliberations by the Project Team. Each theme leads, in turn, to a recommendation.

**Adoption of AI tools by food businesses**

Guidance to FBOs should be issued by the FSA to explain the underlying principles for the use of AI technologies within food safety and assurance processes, including minimum performance expectations, legal responsibilities, risk monitoring, documentation requirements, and clear criteria for human oversight. Ultimately all safety critical decisions should be made by humans, be explainable and traceable.

The case studies reinforced this principle across different contexts. In Case Study 1 (risk assessment of manufactured foods), participants warned that AI could lull businesses into a false sense of security, creating outputs that look convincing but are unverified. Case Study 2 (third-party certification) suggested how AI might process incomplete records, underlining the need for human judgement to challenge results and assess context. In Case Study 3 (abattoirs), the importance of human oversight was highlighted in safeguarding against drift and ensuring ambiguous/unusual cases were resolved by inspectors.

Thus, accountability and business process ownership may be even more important with the advent of AI where there may be a danger of human workers leaving tasks to AI tools without adequate critical supervision. Safety assurance process owners should be human. The individuals concerned should be explicitly identified and competent for that role. While tasks can be assigned to AI algorithms, these should be under human supervision applying appropriate validation and documentation of the process and routine verification to assure that the AI tools perform correctly.

AI systems may support decision-making through data analysis, pattern recognition, or anomaly detection, but must not replace human judgement in safety-critical contexts such as Hazard and Critical Control Point (HACCP) decision points or regulatory inspections. Best practice will include transparent system logs that distinguish between AI-generated outputs and human decisions. This ensures traceability, supports due diligence defences under the Food Safety Act 1990, and maintains public and regulatory trust in AI-augmented assurance systems.

The business relationship between FBOs and AI technology suppliers needs to be carefully managed to ensure AI tools are validated using real world business data rather than based on experimental or hypothetical examples. The onus should be on the technology supplier, in partnership with the food business, to provide tools that are validated and fit for purpose. The FBO should be aware of the applications for which the tools have been developed and any limitations.

Case study discussions repeatedly highlighted that many AI systems available today are adapted from other domains and may not have been developed with food safety in mind. In Case Study 1 (risk assessment of manufactured foods), concerns were raised that generic AI systems might “lull users into a false sense of security” if validation was inadequate, while in Case Study 2 (third-party certification), participants emphasised that systems could misinterpret documentation unless they were trained on sector-specific, high-quality records.

These examples underline the need for transparent agreements between FBOs and suppliers that define how tools are validated, the data they are trained on, and the contexts in which they can or cannot be reliably used.

There is also a broader governance issue: food businesses remain legally accountable for food safety, but they may increasingly depend on AI suppliers for technical assurance. Case Study 3 (abattoirs) showed how AI could miss rare pathologies if suppliers failed to provide diverse training datasets, while Case Study 4 (ports) highlighted the importance of ongoing updates to keep pace with regulatory change. In both examples, weaknesses in supplier responsibility could directly undermine the ability of FBOs to demonstrate compliance. This risk underscores the importance of clear contractual frameworks that assign responsibility for validation, updates and transparency in system performance.

Going forward, closer collaboration between FBOs, AI providers, and regulators will be essential to avoid fragmented responsibility and ensure shared accountability. While ultimate legal responsibility for compliance cannot shift from the FBO, suppliers must be held to account for the quality, transparency, and robustness of their systems. Establishing common expectations for supplier validation, performance disclosure, and limitation reporting would not only protect businesses but also provide greater assurance to regulators and consumers. Without such safeguards, there is a danger that AI adoption could create new vulnerabilities in food safety rather than strengthening assurance.

- **Recommendation: Publish Guidance on Responsible Use of AI to Assure Food Safety and Regulatory Compliance**

#### Diversity and speed of introduction of AI tools and applications

AI is evolving rapidly, and its application is likely to change as capabilities mature, and costs fall. The extent and rate of this evolution is difficult to predict, making AI application highly dynamic.  In addition, deployments in the food system are at an early stage, with many tools piloted in constrained settings rather than embedded into day-to-day operation and assurance. This makes real-world performance uncertain: behaviours observed in trials may not hold when systems face the variability of commercial operations, diverse datasets, and shifting standards. Against this backdrop, ongoing surveillance enables the FSA to observe how AI is actually being used, adapts over time and where new risks or opportunities emerge.  In addition, not all applications will be in the published

scientific literature; awareness of the grey literature and business activities will also be essential.  An additional target of a broader understanding of developments in AI use in food systems would be to ask if current supply chain standards are resilient to the possible use of AI to assist food fraud. The widespread availability of AI will undoubtedly attract criminals searching for ways to circumvent business food controls and regulatory checks.  AI could assist in label counterfeiting, document fraud, and many help criminals find vulnerabilities in food systems.

Evidence from the case studies showed how unintended consequences may arise once systems are deployed. In Case Study 1 (risks for manufactured foods), participants cautioned that AI outputs can create a false sense of security if accepted uncritically, and that weaker operators might use AI to generate convincing risk assessments. In Case Study 2 (third-party certification), groups highlighted that multimodal document tools could take falsified or incomplete records at face value, producing authoritative-looking evidence lacking substance. Case Study 3 (abattoirs) emphasised the risk of performance drift and gaps around rare pathologies, arguing for long term trials at scale and continuous revalidation.  Case Study 4 (ports of entry) raised concerns about large language models producing hallucinations and false positives, underscoring the need for monitoring and human challenge where results appear plausible but incorrect.

Because AI systems learn from and react to new data, one-off validation is not sufficient. Surveillance would allow the FSA to track adoption patterns (where, by whom, and for what decisions), watch key performance indicators over time (e.g. false-positive/negative rates, override and challenge frequencies, drift alerts), and identify signals of misuse or over-reliance (e.g. declining human verification, reliance on non-explained “black box” outputs without traceable evidence). It would also help the FSA spot data governance issues as they arise, such as inconsistent training data, poor provenance, or undocumented model updates, and target guidance or engagement accordingly.

- **Recommendation: Establish Ongoing Monitoring of AI Systems and Potential Impacts**

#### Data availability and quality as a major prerequisite for AI applications in food safety and authenticity assurance

High-quality, FAIR (Findable, Accessible, Interoperable, Reusable) data is fundamental to the development, exploitation and validation of AI systems in food

safety, authenticity and assurance. The FSA should intensify its efforts to promote trusted food supply chain data sharing and alignment with recognised standards, ensuring datasets are protected from bias and drift, data provenance and ownership is clear, and regulatory consistency is maintained, including terminology, document formats and traceability frameworks. This builds on Science Council's WG4 report on data usage (Wolfe et al., 2020) and current FSA support for the Defra Food Data Transparency Partnership (FDTP) program. Alignment with UK food system vocabularies and record-keeping practices is essential to support transparent, auditable and fair AI behaviour. The barriers to AI captured in the Centre for Data Ethics and Innovation (2020) report including ethical and data barriers, are still highly pertinent. High quality data sharing and transparency is especially important as a means of detecting and preventing food fraud. The utility of AI as a tool to detect anomalies in supply chain data will be determined by the quality and accessibility of raw data from multiple sources. In summary, data access, quality and veracity is a pre-requisite for the successful application of AI tools by businesses and regulators to assure food safety and regulatory compliance. Increased data sharing would benefit all stakeholders and help prevent food fraud. Harmonized use of AI methodology across businesses and regulators would support consistent decision making and would build trust. There is a growing number of cases of cyber attacks on food businesses that have shut operations. In all cases, multiple sites and multiple enterprises were affected. In applying digital tools food businesses will need to look beyond their own organisation to guarantee cyber security.

The case study on border inspections (Case Study 4) demonstrated that without shared and standardised datasets, AI systems could not complete intended tasks (assuring food safety and provenance at borders). If AI systems are to deliver reliable outputs or even function, participants stressed that models must be trained on the full diversity of paperwork encountered at ports, including multilingual, handwritten and varying regulatory formats. Without this access, systems risk producing inaccurate or biased results, undermining both efficiency and trust. Missing data post training, when any AI is in operation, could render the system ineffective. Similar concerns were raised in Case Study 1 (risk assessment of manufactured foods), where participants warned that low-quality or incomplete input data could lead to poor outcomes, regardless of how sophisticated the AI appeared. Case Study 2 (third-party certification) also showed that AI tools may accept inconsistent records at face value unless data standards are robust. Taken together, these examples show that AI in food safety will not accomplish its intended tasks, or could deliver biased, incomplete, or misleading outputs, without access to high-quality, harmonised data. This makes data assurance and

consistency not just a desirable feature but a fundamental prerequisite for meaningful AI deployment in food assurance.

At the same time, the FSA should recognise that not all FBOs, particularly SMEs, currently have equal access to the data infrastructure (internet / computer), expertise, or technical capacity required to fully benefit from AI. Generic AI tools may be accessible but of variable effectiveness for businesses due to these disparities. The FSA should develop an understanding of limitations and opportunities and ensure that efforts to improve data standards and digital capability are inclusive, enabling fair and proportionate adoption across the sector.

- **Recommendation: Promote Data Assurance, Validation, and Standards Alignment**

#### Food business vulnerability to AI products and applications that have not been rigorously tested or validated

The FSA should support the development of independent standards and validation mechanisms to ensure AI systems used in food safety are safe, reliable and fit for purpose. This may include the use of regulatory “sandboxes”, digital twins, or benchmarked synthetic data for independent testing prior to deployment. Such validation should assess key performance indicators, including false positives/negatives, hallucination risk, explainability, and consistency across environmental diverse conditions. An advantage of industry-driven standards and Codes of Practice is that best practice is likely to evolve at an accelerated pace requiring an agile process to capture new developments in a timely manner.

The case studies highlighted the need for systematic validation and shared standards and assurance frameworks to underpin trust in AI. In Case Study 2 (third-party certification), participants stressed that AI conclusions must be auditable and benchmarked against human evidence standards to be credible in regulated environments. Case Study 3 (abattoirs) reinforced this by emphasising the need for long-term efficacy trials and continuous revalidation to capture performance drift and rare pathologies. Case Study 3 also emphasised the need to quantify, as part of any standard, likelihood of false positive and negative results, these could have serious food safety consequences if not properly addressed. Whilst Case Study 4 (ports) showed that outputs must remain explainable and adaptable to changing regulatory standards if they are to be accepted by inspectors. Across these discussions, it was clear that without agreed

standards and a common code of practice, AI systems risk uneven application, variable performance and erosion of trust. A coordinated framework would give businesses clarity on expectations and provide regulators with assurance that systems meet consistent, transparent benchmarks.

In parallel, the FSA should encourage the food industry, working with standards bodies, if necessary, and cross-government partners, to develop an industry-led Code of Practice for AI in food safety contexts. While not leading this directly, the FSA can act as a convenor and advisor to ensure alignment with regulatory expectations and consumer protection. The Code could draw on existing frameworks and help set clear expectations around data quality, governance, transparency and system robustness.

Coordination with wider government initiatives on AI assurance and standards will be important to ensure coherence across sectors while addressing the unique risks and regulatory needs of the food system.

- **Recommendation: Support the Development of Standards and an Industry-Led Code of Practice for Assuring AI in Food Safety**

#### Opportunities arising from similarities and synergies across different regulatory and policy domains

Given the wider implications of AI deployment across society, the FSA should engage with other relevant regulators, such as those within the Department of Science and Technology (DSiT), and appropriate international bodies to ensure coherence in governance and ethical standards. Lessons should be drawn from parallel domains (e.g. financial services, health diagnostics) where AI is being applied. Collaboration can also support benchmarking of assurance frameworks and AI auditability standards. The applications of AI in the domains of healthcare and clinical practice are developing at pace stimulating a number of commentaries and cautions on implications for regulatory compliance and quality assurance (e.g. Ong et al., 2025; Basubrin & Basubrin, 2025). It is imperative that regulators share knowledge of risks and opportunities.

The case studies made clear that AI in food safety is still in the early stages of adoption, with many tools at prototype or pilot level rather than scaled deployment. Case Study 4 (ports of entry) highlighted that without harmonisation of documentation standards across jurisdictions, AI could not deliver its intended function. Similarly, Case Study 3 (abattoirs) showed that validation of pathology

detection tools requires not just technical testing but alignment with certification frameworks and inspector practice. These examples point to a broader context: many of the challenges facing the FSA mirror those in other domains, where regulators are facing similar questions of explainability, accountability and bias. It is therefore likely that the FSA can benefit from and contribute to this wider regulatory conversation rather than seeking to resolve these issues alone.

Engagement is also essential because AI applications are evolving rapidly, with new forms such as large language models and agentic AI emerging far faster than traditional regulatory processes can adapt. The case studies underscored the risk of unintended consequences: automation creep in abattoirs (Case Study 3), over-reliance on AI-generated certification packs (Case Study 2), or misplaced trust in unexplained outputs (Case Study 1). These risks highlight the importance of ensuring that the food system is aligned with cross-sector governance approaches that are developing in real time. A siloed approach could leave the FSA unprepared for the rapid diffusion of tools into food assurance that were originally designed for other industries.

Finally, engaging with broader regulatory and policy perspectives will help the FSA anticipate the legal and ethical shifts that are already beginning to shape AI deployment. The Law Commission (2025) has warned of liability gaps in autonomous and adaptive AI “where no natural or legal person is liable for the harms caused”, while international precedents such as the EU AI Act (2024) are setting new benchmarks for risk-based regulation. By working with other regulators and government departments, the FSA can ensure that food-specific concerns such as traceability, authenticity, and public health, are not overlooked in these wider frameworks. At the same time, this collaboration will give FBOs greater legal certainty and ensure that consumer trust in the UK food system is not undermined by inconsistent or fragmented approaches to AI governance.

- **Recommendation: Engage with Broader Regulatory and Policy Perspectives**

#### Human user understanding and application of AI outputs compared with those from conventional tools and advisors

To fully understand the potential risks and opportunities of AI deployment in the food system, the FSA should consider commissioning or supporting behavioural research, either directly by engaging the Advisory Committee for Social Science ( ACSS) or with other funding agencies. While much attention has been given to the

technical assurance of AI systems, less is understood about how human behaviours, such as cognitive bias, overreliance on automation, or misplaced trust in AI-generated outputs may impact food safety outcomes. While AI may lead to less demand for some human-actuated tasks, there may be an opportunity for assurance staff and regulators to focus on higher value activities such as targeted interventions during an inspection or audit. There may also be scope for skills development of the human workforce to enable earlier interventions to prevent food safety incidents leading to better consumer protection and lower business risks. However, vigilance may be needed to prevent essential skills deterioration particularly in small organisations.

The case studies revealed that the most significant risks may arise not from the technology itself, but from how people choose to work with it. In Case Study 1 (risk assessment of manufactured foods), participants noted that staff could become complacent, assuming AI-generated outputs were correct without carrying out independent checks, which risks embedding errors into safety plans. In Case Study 2 (third-party certification), auditors were concerned that over-reliance on AI-compiled assurance packs could discourage challenge. Case Study 3 (abattoirs) raised the possibility of inspectors gradually transferring too much responsibility to automated systems, leading to “automation creep” and even loss of critical skills over time. At ports of entry (Case Study 4), officials stressed that some users might avoid using complex AI tools altogether if they lacked training or trust, while others might even seek to game the system. These examples highlight the spectrum of human behaviours, from over-trust and passivity to avoidance and opportunistic misuse, that must be recognised to ensure AI supports, rather than undermines, food safety.

Research should explore how individuals across the food system (e.g. FBOs, inspectors, and consumers) engage with AI-generated information, including how trust is formed, when human oversight may weaken and how shortcuts may impact risk perceptions. Findings from sectors such as medicine and clinical decision support, where human-AI interaction is more advanced, could offer valuable insights.

Understanding these behavioural dynamics will help the FSA to develop more effective guidance, training, and governance approaches that account not just for the capabilities of AI systems, but also for the realities of human behaviour in operational settings.

- **Recommendation: Commission Research on Human Behaviour and Interaction with AI in Food Safety Contexts**