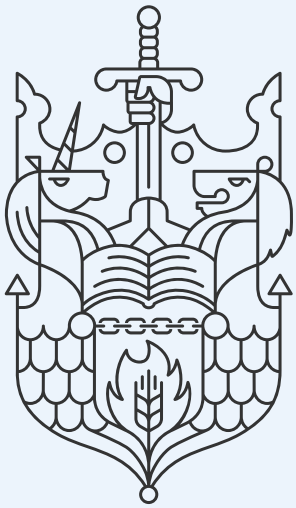

Chartered
Insurance
Institute

Standards. Professionalism. Trust.

Artificial Intelligence and vulnerable customers

Roundtable summary report

cii.co.uk

Contents

As a chartered body with a public interest mandate, the Chartered Insurance Institute provides a forum where stakeholders can collaborate on shared challenges. Our independence enables honest dialogue, facilitating the development of sector-wide guidance and recommendations that strengthen professional standards and deliver better customer outcomes.

Executive summary

The CII convened regulatory, sector, ethics, technology and lived experience experts to examine how AI is being used to identify and support vulnerable customers, exploring the opportunities, risks, and ethical considerations that emerge from its deployment.

The Financial Conduct Authority (FCA) explained that Artificial Intelligence (AI) related risks can be mitigated within existing legislative and regulatory frameworks and does not intend to introduce new rules on AI at this time. The regulator has adopted a principles-based, "tech-positive" approach, and is aligned with the 'responsible AI' principles suggested by the UK government.

AI shows promise in proactive identification through predictive analytics detecting early warning signs of vulnerability, and also in reactive approaches that analyse conversational cues, emotional signals, and language patterns that humans may miss during customer interactions.

The discussion surfaced tensions around transparency, consent and agency; capabilities required for humans to meaningfully challenge AI-generated outputs; the imperative to establish robust vulnerability data infrastructure before deploying AI solutions; and the need to combine proactive and reactive identification strategies.

Participants agreed AI should augment rather than replace human judgment, and that firms must prioritise consumer outcomes over efficiency, undertake thorough vendor scrutiny, pilot solutions, implement transparent decision-logging, and carry out outcome monitoring to prove AI delivers good outcomes for vulnerable customers.

The CII remains committed to facilitating the co-creation of practical resources and examining the ethical tensions in this evolving space.

Introduction

The FCA's Consumer Duty (FG22/5) and strengthened guidance on vulnerable customer treatment (FG21/1), require firms to evidence consistent delivery of good outcomes for all customers, including those in vulnerable circumstances. Concurrently, the rapid emergence of AI solutions is leading firms to increasingly view AI as a potential solution to enhance the identification of customer vulnerabilities and potential support.

This paper summarises the discussion from a roundtable of sector, regulatory, technology and academic experts as well as from people with lived experience of vulnerability. It examines the promising applications, inherent risks, and ethical dilemmas that firms face.

The FCA's position on AI

The roundtable discussion began with the FCA outlining its position on AI in financial services.

The FCA maintains that AI-related risks can be mitigated within existing legislative and regulatory frameworks, particularly through Consumer Duty requirements and vulnerability guidance, as well as UK GDPR and the Equalities Act. The regulator does not intend to introduce new, prescriptive rules on AI at this time, viewing doing so a potential barrier to innovation.

The regulator is taking a principle-based approach to AI, aligned with the five cross-economy 'responsible AI' principles suggested by the UK government (as outlined in the FCA's AI update):

- **Safety, security, and robustness:** the FCA's work on operational resilience, outsourcing and Critical Third Parties (CTPs) are relevant to ensuring systems function in a secure, safe and robust way, with risks carefully managed.

- **Transparency and explainability:** firms may want to consider how AI outputs are generated, i.e. 'black box' decision making. Firms must act in good faith and meet the information needs of retail customers, equipping them with the information to make decisions that are effective, timely and properly informed.

- **Fairness:** systems should not undermine or discriminate unfairly against individuals across demographics, testing for biases in risk assessments, profiling and recommendations as well as protection for customers in vulnerable circumstances. Firms must act to deliver good outcomes for retail customers and are required to act in good faith and avoid causing foreseeable harm.

- **Accountability and governance:** firms should ensure appropriate oversight and senior management responsibility, including documentation of AI decision-making processes, with clear accountability for outcomes.

- **Contestability and redress:** customers should be able to challenge AI-driven decisions and seek remediation if a firm's use of AI results in a breach of FCA rules.

Promoting innovation

While emphasising that existing laws and regulations apply to the responsible use of AI, the FCA has adopted a 'tech-positive' stance to encourage safe and responsible innovation. Initiatives like the Supercharged Sandbox and AI Live Testing are designed to help firms develop and deploy new technologies in a controlled environment, with regulatory oversight to ensure consumer protection.

AI in practice – two use cases for identifying vulnerability

Roundtable speakers presented two approaches where AI is being applied to identify and support customers who might be in vulnerable circumstances:

- **Proactive identification – flagging early warning signs:** AI-powered predictive analytics use customer transactional data to help firms identify vulnerabilities, spot early indicators of financial hardship and plan early interventions. This enables firms to act before problems escalate, for example, reaching out to policyholders showing early signs of payment stress to discuss payment holidays or policy adjustments before they lapse their coverage.

- **Reactive identification - augmenting human interactions:** AI tools that transcribe and analyse client interactions with firms can examine: conversational cues that human advisers might miss in real-time, including tone, pace, and sentiment; emotional stress signals such as raised pitch, pauses, or hesitations; and language indicating hardship like job loss, illness, or bereavement.

AI alone is not a solution

Participants agreed that AI's greatest value lies in augmenting human judgment rather than replacing it, acting as a safety net that provides an additional layer of support.

It is important to keep in mind that for AI solutions to be effective, they require appropriate data infrastructure, systems, and processes to enable adequate vulnerability data management. Without these foundations in place, AI tools cannot deliver their intended benefits.

From a compliance point of view, AI-generated assumptions may fail GDPR data accuracy requirements unless verified through direct customer contact, meaning firms should not act on such inferences as if they were factual vulnerability data.

These applications raise ethical questions and practical challenges that formed the heart of our subsequent roundtable debate.

The core debate: ethical tensions

The discussion revealed four areas of tension where the promise of AI clashes with the potential for harm.

Identifying vulnerability: promise versus risk

Using AI to identify vulnerability is a double-edged sword. While it offers the potential for earlier and more consistent support, it also introduces new risks.

Promise includes:

- **Spots missed signals:** umans can miss subtle cues during pressured and complex interactions.
- **Reduces unconscious bias:** When properly designed, AI can flag vulnerability indicators more objectively than human judgment alone.
- **Scalability:** AI can monitor thousands of customer interactions simultaneously, extending vulnerability identification capacity.
- **Reduces reliance on disclosure:** Identifies customers who need support but may not realise help is available, or don't know how to ask for it.
- **Pattern recognition across time:** Detects gradual changes in customer behaviour or circumstances that might escape notice in isolated interactions (for example, incremental withdrawal of funds or changing contact patterns).
- **Consistent recording:** enables an auditable process for identifying and logging potential vulnerability so long as the firm has the right data infrastructure to begin with, and inferences are confirmed.

Risks include:

- **Single-channel blind spots:** Solutions focused only on one channel, like telephony, will miss vulnerabilities expressed elsewhere.
- **False positives:** might over-flag, leading to incorrect assumptions and breaching UK GDPR data accuracy requirements if not verified through direct customer contact.
- **Data avalanche:** firms can become overwhelmed with data, leading to evidence of vulnerability without a clear plan to act.
- **Context blindness:** AI may flag temporary stress (a bad day) as evidence of a vulnerability or miss cultural nuances that impact how people express themselves (if training data only reflects dominant cultures).
- **Privacy erosion:** Analysing tone, emotion, and behavioural patterns represents intensive surveillance that customers may not expect or consent to.

Reactive detection versus systematic identification

Many AI solutions focus on augmenting reactive human interactions, such as analysing calls, reading emotional cues, interpreting conversations. However, this approach assumes vulnerability identification depends primarily on frontline staff judgment during customer contact. Firms should develop objective assessment frameworks that identify vulnerability systematically, not just tools that make reactive interactions slightly better at catching what could have been flagged by an objective assessment earlier.

Transparency and consent

A debate emerged around what, when, and how firms should disclose and seek consent about their use of AI to detect characteristics of vulnerability:

- ***What*** should be disclosed? Is there a difference between using AI for reactive note-taking versus proactive predictive monitoring of a customer's data for signs of future vulnerability?

- ***When*** should disclosure happen? Should it be a one-off notice at the point of sale, or should it be reinforced at every interaction?

- ***How*** should firms disclose? Is a line in a privacy policy sufficient, or is a more active, verbal confirmation required?

There were two broad positions identified by our roundtable participants, in favour of and against seeking consent before engaging with a service:

- **In favour of explicit consent:** one perspective holds that disclosure must happen *before* a customer engages with an AI enabled service. This is seen as a prerequisite for trust, giving customers the agency to choose a provider whose methods they are comfortable with, especially for those with a history of trauma who are sensitive to decisions being made about them without their knowledge.

- **Against explicit consent:** a counterargument highlighted real-world complexities of seeking consent. Forcing a disclosure on certain customers, such as an individual with low digital capability, could cause unnecessary stress and confusion. There's also the real challenge of 'consent fatigue' that many consumers experience, often clicking 'agree' on complex terms without true understanding, making the ideal of informed consent difficult to achieve in practice.

Responsible AI kitemark

Participants suggested that an independent body could provide a kitemark or certification of trust to organisations using vulnerability management AI tools responsibly. This could help mitigate consent fatigue by offering customers a simplified signal of trustworthy practice. A kitemark would also create accountability through external standards rather than relying solely on individual firm transparency efforts.

The ‘human in the loop’ - but any human?

Human oversight is non-negotiable, but not all human oversight is equal. Participants noted that simply having a person in the loop does not guarantee good outcomes. Humans bring their own cognitive biases and inconsistent decision-making and may lack the capability to effectively challenge AI outputs. The question is not whether there’s a human involved, but whether that human is equipped to provide meaningful oversight.

Supporting effective human oversight

Firms must provide ongoing AI literacy training programmes that are role and context specific and ensure that an *AI fluent and ethical human* is in the loop, supported by accurate, unbiased tools that are used for the right reasons.

From theory to practice: procurement, governance and implementation

The discussion turned to how firms can responsibly procure and implement AI solutions for vulnerable customers, and suggested the following guiding principles:

1. **Start with a clear framework:** Firms must first define vulnerability using an established, credible framework, such as the framework used by the FCA in FG21/1 Vulnerability Guidance. This ensures the AI is being directed toward a clear and consistent objective.

2. **Prioritise consumer outcomes over efficiency gains:** Firms must ensure their primary goal is aligning with the Consumer Duty and achieving good outcomes, not simply optimising for profit.

3. **Leverage existing standards:** Firms should use established resources such as the AI Procurement Lab and existing ISO standards that provide robust, pre-vetted guidance on AI governance. See appendix for links.

4. **Redesigning services:** The true opportunity of AI is not merely to enhance existing processes but to fundamentally redesign them around the customer. One participant articulated the risk of firms merely "bolting tech onto a failing process." The goal should be to use technology to create customer-centric journeys from the ground up, ensuring the right support is embedded in the service, not just added as an afterthought.

These guiding principles should be considered across all phases of AI implementation. Firms should consider the following across each phase:

Phase	Considerations
Due Diligence	**What should firms demand from AI vendors?** Before selecting an AI vendor, firms must scrutinise the system's foundations and ask: What data sources does the model use, and how representative are they? What methods ensure the model treats different customer groups fairly? Can the system explain its decisions in ways that support compliance and customer redress enquiries? Does the vendor's approach align with FCA principles and Consumer Duty requirements?
Internal Preparedness	**How should firms prepare for implementation?** Successful implementation requires more than technical deployment. Teams must be trained to interpret AI outputs critically, challenge recommendations when they appear inconsistent with individual circumstances, and respond to customers with empathy. Before full rollout, firms should conduct pilots and stress tests using real-world scenarios involving vulnerable customer groups to identify potential failure points and refine processes.
Post-implementation governance	**What oversight should be in place?** Firms should establish clear accountability by designating specific owners for monitoring system performance and outcomes. This includes creating escalation pathways when the system produces concerning results, implementing regular ethical reviews to ensure the AI continues achieving its intended goals without causing unintended harm, and maintaining the capability to conduct root cause analysis when adverse outcomes occur.

Measuring success: monitoring for genuinely better outcomes

Outcome monitoring is a non-negotiable component of the Consumer Duty, and this principle applies when measuring the success of AI solutions to support vulnerability management. It is not enough to *believe* an AI-enabled service is helping; firms must be able to prove whether it is improving or worsening outcomes for vulnerable customers.

Firms will already have metrics for tracking outcomes. However, in the context of AI use, the critical challenge lies in the ability to investigate and address adverse events. When things go wrong, firms must be able to conduct root cause analysis, and this is only possible if systems are transparent in their decision-making processes. To enable effective investigation, AI systems must:

- Log the decision pathways taken for specific customer outcomes
- Flag anomalies or unusual data points that decisively influenced a decision
- Identify the features or data points that were most salient in reaching a conclusion

Without this level of transparency, firms cannot learn from failures, correct systemic issues, or provide satisfactory, evidenced explanations to customers and regulators when adverse outcomes occur.

This segment ended with a call to action: how will firms evidence whether vulnerable customers served by AI enabled services are getting better outcomes or just different ones?

Conclusion

AI solutions genuinely offer the potential to identify and support individuals in vulnerable circumstances, but simply adopting AI to keep pace, without adequate vulnerability management data infrastructure, governance frameworks, and a supportive culture is a recipe for failure. The CII's managing vulnerability guidance provides practical direction on building these foundational capabilities.

Once these foundations are in place, deployment must be guided by clear commitments to fairness, transparency, accountability, and inclusion, principles that align with the UK government's framework for responsible AI.

Beyond these principles, existing laws and regulations apply equally to AI systems: data protection requirements, consumer duty obligations, and vulnerability management standards remain in full force regardless of whether processes involve algorithmic or human judgment.

A significant insight from the roundtable was the recognition of the unresolved tensions that are shaping the debate in this space, including: informed consent versus pragmatic approaches, or prescription versus flexibility in regulatory approaches, to name a few.

These tensions have no simple answers, and to help address them the CII remains committed to:

- **Facilitate dialogue where stakeholders can co-create practical resources**, for example procurement checklists that incorporate vulnerability, adapted from established standards.

- **Developing thought leadership** that examines the ethical tensions inherent in AI-enabled vulnerability management.

Our independence as a Chartered body enables us to convene the honest conversations needed to develop sector-wide solutions that strengthen professional standards and deliver better outcomes for customers in vulnerable circumstances.

Appendix: resources

Vulnerability management:

- Managing customer vulnerability in insurance and personal finance: a practical implementation guide: https://media.umbraco.io/ciigroup-dxp/eaedofqi/managing-customer-vulnerability-in-insurance-and-personal-finance-a-practical-implementation-guide.pdf

Standards:

- ISO AI data standards: https://www.iso.org/sectors/it-technologies/ai

Procurement guidance:

- AI Procurement Lab: https://www.aiprocurementlab.org/

Governance guidance:

- Centre for democracy and technology: https://cdt.org/cdt-ai-governance-lab/
- International Association of Privacy Professionals (IAPP): https://iapp.org/
- Centre for the Governance of AI: https://www.governance.ai/

Mapping risks:

- The MIT AI risk repository: https://airisk.mit.edu/
- The IBM AI risk atlas: https://www.ibm.com/docs/en/watsonx/saas?topic=ai-risk-atlas

AI ethics training certifications:

- IAPP Artificial Intelligence Governance Professional: https://iapp.org/certify/aigp/
- CISI: Certificate in Ethical Artificial Intelligence: https://www.cisi.org/cisiweb2/cisi-website/why-choose-a-CISI-qualification/professional-assessments/Certificate-in-Ethical-Artificial-Intelligence

Participants

- **Chris Adlard**, Elephants Don't Forget, Director of Customer Experience & Compliance
- **Tatiana Caldas-Löttiger,** International WoMenX in Business for Ethical AI, Founder and CEO
- **Prof. Christopher Cowton**, University of Huddersfield, Emeritus Professor
- **Christopher Digby**, Howden Group, Executive Director
- **James Edmonds**, Protect Association, Managing Partner. MorganAsh, Insurance Sector Consultant
- **Kate Gannon**, Themis Wealth Management, Director. NED Personal Finance Society
- **Annabel Gillard**, Conversations on AI, Co-founder
- **Edward Grant**, Cabinet Office Disability and Access Ambassador, NED Personal Finance Society (PFS) and European Financial Planning Association (EFPA), Chair of the Finance in Society Research Institute (FISRI)
- **Martin Grimwood**, FWD Research, Director
- **Andy Harrison**, Fidelity, Senior Manager, Vulnerable Customers Lead
- **Kath Harvey**, Melo, Head of Progress
- **Dan Holloway**, University of Oxford, CEO Rogue Interrobang
- **Vicki Jordan**, FCA, Policy Manager – AI, Remuneration, and Policy Profession
- **Sarah Langley**, Consumer Code for New Homes, Managing Director
- **Claire Massey**, Claim Guardians, Founder and CEO
- **Diane Maxwell**, Association for Insurance and Risk, CEO Designate
- **Ian Roberts**, AXA, Control and Contract Manager
- **Alex Williams-Jones**, FCA, Senior Associate, Consumer Policy and Outcomes

- **Matthew Connell**, CII, Policy and Public Affairs Director
- **Adam Harper**, CII, Executive Director, Strategy, Advocacy and Professional Standards
- **Vanessa Riboloni**, CII, Head of Research and Insight
- **Ian Simons**, CII, Content and Capabilities Director

Chartered Insurance Institute

Artificial Intelligence and vulnerable customers

Roundtable summary report