Review into the long-term impact of AI on retail financial services (The Mills Review)

Citizens Advice Response, March 2026

Introduction

In Citizens Advice frontline advice services, we see people every day who need help with their finances. In 2024-25, Citizens Advice supported 2.71 million people, including 426,000 people with debt. Our webpages had over 61 million views, including 2m for financial services and 5m pageviews on debt. Our role as a frontline organisation means that we are already seeing early evidence of how AI is shaping people's experiences and the problems they face.

We're also responding to this consultation as an organisation that provides regulated debt advice, which gives us insights into the benefits and risks that AI poses in this space. We are already taking steps to embrace responsible AI in delivering our work. We are scaling widespread use of an adviser-supporting RAG chat bot which uses trusted sources to support human-led advice. We use AI tools to transcribe and summarise advice in local offices. We're developing technical frameworks to ensure the accuracy and quality of AI-supported advice across our service.

But we are also concerned. Alongside the benefits, AI also poses real risks to real people, especially in the area of financial advice, consumer rights and changes to the retail financial market.

This document summarises our response to the FCA's review into the long-term impact of AI on retail financial services (The Mills Review).

Contents

Theme 1: Future evolution of AI technology

2. Agentic AI: What do you see as the future potential and direction of agentic AI? What are the implications for retail finance over the coming decade (including accountability, assurance, and market structure)?

Agentic AI has many potential benefits to consumers. It could support less biased, better informed financial decision making, help identify fraud or support consumers to research their options with ease. But there are also risks to consumers:

This is due to the following:

- **Users may not always understand the degree of agency they are giving away.** Currently, there are limited quality controls linked to what AI tools claim they can do vs what they can do, due to a lack of benchmarking, guidance and enforcement. This means the claims from companies of what is possible with agentic AI will likely dwarf the actuality. Without upfront transparency, consumers are more likely to delegate away agency without full understanding of the steps taken.
- **There is currently no enforced regulation on the warnings of what agentic AI might do.** Many models remain unpredictable. By definition, they will make decisions without user consent. The possible outcomes are varied.
- **It is currently unclear where the liability of decision making sits from agentic AI.** We are concerned by users agreeing in the fine print to give away agency, trusting a robotic system. We're also concerned by a lack of clarity on whether use of an AI agent would meet threshold conditions for regulation to apply. For example, it is currently unclear at what point a developer's testing and verification processes constitute the 'reasonable care and skill' condition in the Consumer Rights Act. This lack of recognised standards means that it's not clear where many regulations apply.
- **Records of decision making by agentic AI are currently often inaccessible or unreadable without expert understanding**. For consumers adopting agentic AI products themselves, for example to

maximise trading efficiencies or ask them to switch providers, there is currently often a lack of transparency in agentic AI records. This means in many products, users cannot easily understand the steps an agent took, what tools they accessed or decisions made. When errors are made, consumers will need to be able to understand what their agent did, to be able to untangle the problem and fix it. Consumers deserve to understand what decisions were made on their behalf without having to employ professional services. This is also an area where outcomes based regulation could be helpful by focusing on the impact on the consumer, and placing responsibility on firms to support good outcomes. However the challenge under the current regulatory regime here is likely to relate to where activities fall in and out of the FCA perimeter.

- **Agentic AI, especially OS-embedded systems, opens up new data security risks.** Consumers may not understand the degree of risk they are exposed to by using agentic AI. When consumers allow agentic AI to access their 'context' (ie. vast quantities of triangulated data), which can 'undo' in-app encryption through screen shot protocols and store previously separated data in a single, cloud-based dataset, there are new security concerns. Security researchers are arguing that model Context Protocol (MCP) used in agentic AI standardises the exfiltration path for malicious actors. This could mean agentic AI could leave users more vulnerable to attacks. This is especially concerning given how scammers are increasingly using data context around a person to target individuals.

We believe it's possible that as providers respond to AI-driven switching practices, the market offers of financial institutions adjust towards AI priorities. For example, if an agentic AI is looking to switch current accounts on a user's behalf, it may prioritise value for money (fees) and switching bonuses over customer service. Over time, this could lead to banks focusing on fees and switching bonuses over customer service, to enable AI-driven behaviour to switch towards them, and gain more customers. For customers in vulnerable circumstances who value and/or need good customer service to navigate banking, the market may then shift away from their needs.

We recommend the FCA considers:

- **Regulation on labelling of AI products linked to financial decision making**, with a focus on transparency of agentic AI risk. This could be akin to food labelling or financial warnings on capital being at risk. It should cover the risks of delegation and possible data risks.
- **Mandated, clear and upfront warnings about liability about agentic AI decisions**. Consumers need to understand upfront if they will be held accountable for decisions made on their behalf by agentic AI.
- **Regulation to ensure that consumers have access to readable records of decision making made by AI, especially agentic AI.** This is vital to support consumer understanding and the ability to seek redress. We support calls for a statutory duty to provide a right to explanation in AI driven decision making in financial services.

Theme 2: Future impact of AI on markets and firms

4. Regulatory perimeter: Could AI systems provide services functionally equivalent to regulated activities such as advice or intermediation, while remaining outside the regulatory perimeter? How might this occur in your market, and what proportion of value could migrate to such unregulated services?

More and more consumers are using and trusting LLM's for financial advice. Research from Lloyd's Banking Group found that 56% of adults – around 28.8 million people – say they've used AI in the past 12 months to help manage their money. Amongst people who experienced problem debt, 15% of survey respondents used an AI tool, such as ChatGPT for debt advice over the last 6 months (Citizens Advice polling, Sep-Oct 2025, Sample: 2,000 adults in England and Wales (18+) in problem debt).

Where this advice is accurate and robust this could provide benefits to consumers, helping them to make more informed decisions and access information that is more personalised to their situation. But this also brings substantial risks, particularly where LLMs issue incorrect financial advice. The accuracy of LLM responses vary. The more complex the situation, often the more information is needed and the less accurate LLM responses are likely to be. We're seeing a lack of nuance in the financial advice offered to consumers who come to our offices, and instances of hallucinations.

> "The information is not necessarily correct for that particular client, it depends on their situation. AI can be very generalist." - CitA Adviser

> "[AI] Advice is often incorrect. Fictional cases quoted." - CitA Adviser

At Citizens Advice, we are seeing early evidence of clients experiencing harm as a result of inaccurate or incomplete advice from AI tools, as illustrated in the following case studies:

- A client received advice from an AI tool to file for bankruptcy and followed this advice independently. He then regretted this because he was unaware of other options available to him. He struggled to find ways to cancel or annul his bankruptcy and is now subject to restrictions. He has

incurred further debt post bankruptcy and is now unsure if this was the best option to take in his circumstances.

- A client used an AI tool to seek debt solutions. The technology directed him to a private company, who he got in touch with and looked to enter him into a debt management plan. The client was unaware the company charged a fee until speaking to Citizens Advice, who advised him of the charges.
- A client is in arrears to a utility company, which was transferred to a debt collection agency. After incorrect advice from an AI tool, the client sent a letter to the agency with incorrect information. This has led to further enforcement and an increased amount payable.

Research from Which? Also demonstrates how AI tools can give inaccurate advice. Which? Reviewed 6 AI tools and under lab conditions, asking 40 common questions across four key life areas: money/finance, legal, health/diet and consumer rights/travel. Scoring across accuracy, relevance, clarity, usefulness and ethical responsibility, the tools scored a range of 55-71%, and some gave inaccurate, unclear and risky advice. This included directing users to premium tax-refund companies unnecessarily and giving incorrect advice about ISA allowances, which could cause someone to breach the rules.

We're also concerned that inaccurate advice from LLMs may undermine genuine regulated advice, with our advisers reporting difficulties with clients not believing their advice when it has come into conflict with the advice or information a client has received from an LLM.

Our hypothesis is that this is happening because of a widely studied cognitive bias called the 'continued influence effect', where information that has landed in someone's memory continues to influence judgements even after more recent information has discredited it.

We believe contributing factors to the reports we're receiving include:

- LLMs produce comprehensive financial or debt advice in a confident, professional, believable tone. Evidence shows that users are more receptive to responses delivered in assertive language.
- Some LLMs have agreeability built into their model, meaning the response may align with what the user wants to hear. This can emphasise a confirmation bias in AI use, whereby consumers tend to seek information

that aligns with existing beliefs.
- LLM responses often use source links in replies to reliable sources (including Citizens Advice), even if those source links are misgrounded, ie. lead to pages which do not support the prompted response. Evidence shows that users are more receptive to responses which include concrete evidence or cite external sources.
- Summarising tools such as Gemini distill complex information, which can miss nuance.
- Because these tools are often accurate, consumers are therefore unaware of the rates of inaccuracy, and that it can be right about one thing but wrong about another.
- Algorithmic judgements are sometimes preferred to human ones, with a high dependency on context.

In the legal industry, a rush to adopt AI tools has led to widespread use of generative AI resulting in hallucinations that are incorrect (use a factual error) or misgrounded (the description is correct but cites a source which doesn't support its claims). A study by Stanford shows that tools such as retrieval-augmented generation (RAG) can reduce hallucinations but not eliminate them. They found bespoke legal systems can still hallucinate up to 34% of the time. OpenAI has recently updated its 2025 usage policy to remove legal and medical advice, making it now officially against its usage policy. However it's unclear exactly what changes to the model have been made or if users will continue to seek advice. We believe the pitfalls seen in the legal industry offer a clear warning to the potential problems of unregulated AI-generated financial advice.

There is also a continued desire from consumers to access financial advice from humans. Research from Unbiased shows that 74% of respondents are open to an adviser-led model, compared to just 6% who would choose an AI platform alone. Respondents repeatedly cited trust and personal connections as a reason for preferring human advice, emphasising being able to 'speak to a person' when making complex financial decisions.

Because of this, we recommend the FCA considers if and how financial advice offered from generative AI tools should be regulated. Early data shows that people are inclined to believe incorrect advice from LLMs, and that use of LLMs risks eroding trust in professional advice.

Theme 3: Future consumer trends

1. **Benefits and risks: How might consumers benefit from AI-enabled retail finance from 2030 and what do you foresee as the greatest risks for consumers?**

Where deployed well there are a number of advantages that AI could bring to consumers in retail finance, for example:

- Improved fraud prevention where technology is used to detect potential anomalies before payments are taken.
- Increasing the reach of supervision and monitoring e.g. AI could potentially help regulators to identify and take down higher numbers of misleading financial promotions.
- Automated budgeting and savings tools that support consumers to identify where they can put money aside.
- Improved analytic tools to help consumers manage their budget.
- Managing simple queries to free up customer service resources for handling more complex cases or cases where additional help is needed.

But AI could also bring new risks or extend risks in some areas, including:

- **Poorer customer support** - for example if FS providers deploy AI agents in place of customer service representatives in areas of high risk, situations of vulnerability or acute need. Where these tools are deployed it is also important that customers have access to escalation channels in case an AI agent is unable to resolve their issue, or fails to understand their problem.
- **Poor quality advice** - this may include advice gathered from general purpose AI tools (e.g. ChatGPT) as well as through FS providers own tools. Information provided by AI tools can be inaccurate, but consumers may rely on this information. This could result in poor product decisions, increase the risk of mis-selling or at the sharp end result in consumers facing legal issues or taking out products that are not suitable for their circumstances. It will therefore be critical for the FCA to consider what guardrails need to be in place to protect consumers, and what access to redress they will have where they have received poor or inaccurate information from AI tools.

- **Credit access and approvals** - use of AI in decision making around credit applications, mortgages or insurance decisions could result in changes in patterns in acceptance/decline rates. This will need to be monitored closely to ensure that AI doesn't have intolerable error rates, and introduce biases against specific groups of consumers.
- **An increase in the scale and sophistication of AI-driven fraud**, and the disproportionate impact of AI-driven fraud on customers in vulnerable circumstances.
- **Greater information asymmetries and reduced right to redress**: There is an opacity between companies and consumers when it comes to understanding when automated decision making is being used. Without transparency, customers a) are at increased risk of discrimination and b) will find it harder to seek redress when a poor choice is made.
- **Greater data and privacy risks**, including where companies are using AI to bring together multiple data sets or involving third party organisations in processing individual data: Consumers need transparency to make informed choices about how their data is used, and manage the associated privacy and security risks. The current regulatory framework under the UK GDPR offers only limited protection against harms arising from AI systems. For example, Articles 5(1)(a), 13, 14 and 22C primarily require controllers to provide general information to data subjects, rather than information tailored to the circumstances of a particular individual. In addition, Articles 13 and 14 contain only limited requirements regarding the explanation of algorithmic decisions, while guidance issued by the Information Commissioner's Office (ICO) on explaining AI-driven decisions does not have legal force. The ICO itself has constrained enforcement powers and a broad regulatory remit, and a finding by the ICO does not in itself provide individual vindication or remedy for an affected data subject. Moreover, the UK GDPR largely relies on private rights of redress, which can create significant practical barriers for individuals seeking to pursue claims where wrongdoing has occurred. For these reasons, the protections available under the UK GDPR and oversight by the ICO should be regarded as offering only limited practical recourse for consumers seeking redress for harms caused by AI systems.
- **Increased digital exclusion and a widening digital divide:** in the UK there are 7.9m people who lack basic digital skills. This means they can't do all 8 tasks of the Essential Digital Skills Framework, including turning on

a device, using a mouse and keyboard, setting up a Wi-Fi connection and opening an internet browser. In the UK, there are 1.6m adults who don't have a phone, tablet or laptop and 1.9m households who struggle to afford their mobile contract. The opportunities of AI will not be available to these people. The acceleration of AI adoption will likely widen the digital divide, as the opportunities and advantages are not possible for these groups.

2. **Inclusion versus exclusion: Which consumer segments might 'win' or 'lose' in this new world of AI-enabled retail finance?**

There is a risk that AI-enabled finance could further entrench existing financial inclusion and capability issues. There are groups we would expect to face exclusion, particularly those who already experience digital exclusion or who have low digital confidence. But AI in retail finance could also result in additional groups being harmed or being unable to access the potential benefits. This may include:

- **People who are overconfident when it comes to AI** - for example where they may trust the information it provides without seeking to verify the information independently. This could put consumers who may be considered to have reasonably good levels of financial capability at risk - at the sharp end this could include people relying on financial guidance or information that could conflict with laws. With further use of agentic AI this could also risk consumers being less engaged with their own finances which may result in delays before they notice an issue has occurred.
- **People who are not AI confident** - this could be a very widespread group of consumers, and those included in this group are likely to be quite different from those facing low digital confidence or digital exclusion, as this may include those working in industries where AI adoption has been slower, or who have not had to engage with AI within employment or education. This could result in big gulfs in understanding among different consumer groups, resulting in varying levels of trust and confidence in AI.
- **People with concerns around data sharing and privacy** - many people may have understandable concerns about how their data is being used, how much is being accessed and by what tools which may put customers

off accessing tools. This may mirror issues that have impacted consumer engagement with open banking.
- **Consumers from minoritised groups** - there is evidence of AI systems resulting in biased outcomes.

AI enabled finance could also rapidly increase the risk of scams, which may result in increased demand for non digital services. We already see firsthand how consumers value person-delivered support in areas like fraud and scams, where people may prefer the reassurance of speaking to a person or presenting at an in-person service. This is particularly important as experiencing a fraud or scam issue can make people more concerned about how they can verify the identity of who they are engaging with.

We also expect that AI finance could result in wider gaps in behaviour between different cohorts of consumers, which could impact how the benefits of AI enabled finance are accrued across society. For example there may be some groups of consumers who have more exposure to AI in their daily life which gives them greater levels of confidence using AI tools and understanding of how they work. In contrast, there will be some consumers with limited exposure to AI, which is likely to include people over retirement age, as well as people with limited access to digital tools at home or in the context of their work or education.

We also know that consumers in vulnerable circumstances are often less able to reap the benefits of competitive consumer markets.

For example, Citizens Advice research on the loyalty penalty shows that older people, those on lower incomes and people with mental health problems find it harder to shop around or switch providers to get a better deal.

Our research on hidden deals also shows that people with mental health problems and those experiencing financial difficulty are more likely to find negotiation processes difficult and more likely to experience negative consequences as a result.

It is possible we'll see these patterns replicated as AI is adopted.

It is possible that as providers respond to AI-driven switching practices, the market offers of financial institutions adjust towards AI priorities.

For example, if an agentic AI is looking to switch current accounts on a user's behalf, it may prioritise value for money (fees) and switching bonuses over customer service. Over time, this could lead to banks focusing on fees and switching bonuses over customer service, to enable AI-driven behaviour to switch towards them, and gain more customers. For customers in vulnerable circumstances who value and/or need good customer service to navigate banking, the market may then shift away from their needs.

3. **Changes to products and services: How might AI drive changes and personalisation in products and services, and what impact will evolving consumer expectations have? This could be to do with evolving price, value, fraud, security, mis-selling, advice, or other topics pertinent to you.**

We've grouped our responses around key themes below.

Price and value

We expect that use of AI in financial services could result in the market offering increased levels of personalised pricing and tailored decision making using AI, much like other consumer markets. This could result in costs being increased for certain consumer segments and may result in substantial price differentials for different consumer groups. There is a risk that this could entrench systematic inequalities.

Fraud and scams

AI could increase the risk of fraud and scams, as these become even more sophisticated, particularly as tools like voice cloning software become more readily available. Against this backdrop it will be important for the FCA to consider whether the security measures firms have in place are adequate to protect consumers. This will also need to consider where firms may be deploying off the shelf AI products, which could present new data security risks.

Advice

It is likely we will see more and more consumers using AI for advice. Where the advice given is not appropriately tailored to an individual's circumstances or

includes inaccurate information, this could increase the risk of consumers experiencing poor or sub-optimal outcomes. Some areas of advice will necessarily carry higher risks for consumers. We would be particularly concerned about this in the context of debt solutions, where the current suite of debt solutions carry different risks and downsides for consumers. In this space we are also concerned that there may be a high risk of mis-selling, because of existing sharp practices which exist in these markets, which could impact upon the information AI tools are drawing on. For example in the IVA market there have been widespread issues around poor take-on practices, which include poor advice. If AI tools are drawing on this information when responding to consumer queries this risks baking in known issues within the market.

Customer support

AI is changing how customer support is delivered. For example firms may seek to triage customer contacts using AI tools or use AI tools to make decisions or respond to consumer queries. Where tools like this are deployed it is important that this does not compromise good customer care. We are seeing increasing examples of consumers being sent into 'automated loops' and finding themselves unable to speak to a human when they need to. We are concerned that companies are moving too fast to replace human agents and we hear increasing reports of people 'giving up' on having their query resolved after spending too long chasing automated systems. For some cases, this can result in real financial harm. This is especially true for digitally excluded people who struggle to engage with chatbot processes.

It's important that any evolution here is mindful of consumer expectations and understanding of where consumers are likely to retain a preference for speaking to a person - particularly in the context of consumers in more vulnerable circumstances of dealing with a situation that is likely to cause considerable stress (e.g. fraud, scams or problems with arrears), where they may really value the reassurance of support from a person. For individuals with accessibility needs or in high-complexity cases, speaking to a human may not be a preference so much as a need.

Wider issues - transparency, privacy and redress

We're also concerned about the lack of transparency available to consumers, both:

- when companies use AI to make decisions; and
- where companies use third party AI tools within their service.

The use of undeclared algorithmic decision-making, AI or automated decisions could increase the risk of consumers facing discriminatory outcomes, where biases are built into the tools being used. At an individual level it can be hard for consumers to identify that this is happening - for example if a certain group of consumers faced higher claim rejection rates an individual would not know. This makes it critical that monitoring of AI decision making is being undertaken at a system level, to ensure biases are not being entrenched within decision making processes. Both firms and the FCA need to be carefully monitoring the outcomes consumers are receiving as new AI systems are being adopted in order to quickly identify and address potential issues. It is also critical that decisions made by AI systems are explainable - where issues are identified firms must be able to interrogate and explain what has happened, and cannot rely on the fact that the decision was made by an AI agent.

We already see algorithmic decision-making can result in bias and discrimination against minority groups. There is a growing evidence base that shows predictive algorithms often produce biased results, such as this example of software reported by Propublica used to predict criminal behaviour in the US which was "likely to falsely flag black defendants as future criminals, wrongly labeling them this way at almost twice the rate as white defendants." Systemic bias has also been detected in DWP decision-making systems in the UK, through FOI requests.

We are concerned that due to both training data and design, predictive models will embed systemic discrimination in decision making. When financial decisions are made - about insurance, pensions, credit cards - people deserve to understand why they got the outcome they did. If these are contained in a "black box", which is not understood, this could make it much harder for wronged consumers to seek redress.

Under Consumer Duty, organisations will need to be able to explain the outcomes of decisions. There is already growing evidence of financial organisations such as banks being unable to explain to consumers why a loan was denied, because the decision was made by AI.

We are concerned that there are currently limited avenues by which consumers can seek right to redress due to algorithmic decision-making, mainly because there is a lack of transparency of its use but also because the recent amendments to the UK's GDPR laws weakens the 'right to explanation' for automated decision making.

We welcome the FCA and ICO's announcement in June 2025 of the creation of a joint statutory code of practice for firms.

We urge the FCA to consider:

- How AI use can be made more transparent to consumers,
- How companies will put necessary safeguards in place to protect consumers in high risk markets such as financial services , and
- How to balance the needs of intellectual property and a competitive market with the rights of the consumer.

We support calls for a statutory duty to provide a right to explanation in AI driven decision making in financial services. This is supported by public opinion and bank workers, where 66% of polled respondents thought it was very important that customers had a legal right to an explanation, and 98% of participants said they thought customers are entitled to an explanation when refused a bank loan.

We understand the need to support market competition and the intellectual property involved in AI powered decision-making; these cannot come at the expense of the rights of the consumer.

The FCA should also take steps to increase transparency to consumers when companies are using third-party AI tools.

When a company uses AI tools without consumer knowledge, it reduces a consumer's ability to make informed choices, consent to its use and manage the associated privacy and security. Transparency is essential for building trust with consumers, and supporting rights to redress.

Genuine consent can only be achieved with information, and currently providers must offer very little information or explanation of how data is used, often using lengthy privacy policies. Although technically compliant with UK GDPR, consumers rarely interact with these policies meaningfully nor understand them. We believe consumers are consenting in an informational void, with little to no understanding of the risks involved.

We propose the FCA should explore introducing further measures which increase the obligations of financial firms to a) disclose where AI is in use and b) explain the implications of AI use on consumer data.

In the same way that financial institutions must warn investors that their capital is at risk, or food labels disclose what's inside, using AI within financial services should be clearly disclosed. The industry has obligations to help consumers understand that their privacy and data security is taking on a different risk profile through use of AI. See our response on Agency and Understanding for further detail.

4. **Agency and understanding: With the balance shifting between consumer agency and delegation to AI, how might this affect consumer understanding, financial literacy and vulnerability?**

Agentic AI has many potential benefits to consumers. It could support less biased, better informed financial decision making, help identify fraud or support consumers to research their options with ease.

But, we believe that the risk of harm may be greater to people in vulnerable circumstances or who have lower AI competency. This is due to the following:

- **Users may not always understand the degree of agency they are giving away.** Currently, there are limited quality controls linked to what AI tools claim they can do vs what they can do, due to a lack of benchmarking, guidance and enforcement. This means the high claims from companies of what is possible with agentic AI will likely dwarf the actuality. Without upfront transparency, consumers are more likely to delegate away agency without full understanding of the steps taken.
- **There is currently no enforced regulation on the warnings of what agentic AI might do.** Many models remain unpredictable. By definition,

they will make decisions without user consent. The possible outcomes are varied.

- **It is currently unclear where the liability of decision making sits from agentic AI.** We are concerned by users agreeing in the fine print to give away agency, trusting a robotic system. We're also concerned by a lack of clarity on whether use of an AI agent would meet threshold conditions for regulation to apply. For example, it is currently unclear at what point a developer's testing and verification processes constitute the 'reasonable care and skill' condition in the Consumer Rights Act. This lack of recognised standards means that it's not clear where many regulations apply.
- **Records of decision making by agentic AI are currently often inaccessible or unreadable without expert understanding**. For consumers adopting agentic AI products themselves, for example to maximise trading efficiencies or ask them to switch providers, there is currently often a lack of transparency in agentic AI records. This means in many products, users cannot easily understand the steps an agent took, what tools they accessed or decisions made. When errors are made, consumers will need to be able to understand what their agent did, to be able to untangle the problem and fix it. Consumers deserve to understand what decisions were made on their behalf without having to employ professional services.
- **Agentic AI, especially OS-embedded systems, opens up new data security risks.** Consumers may not understand the degree of risk they are exposed to by using agentic AI. When consumers allow agentic AI to access their 'context' ie. vast quantities of triangulated data, which can 'undo' in-app encryption through screen shot protocols and store previously separated data in a single, cloud-based dataset, there are new security concerns. Security researchers are arguing that model Context Protocol (MCP) used in agentic AI standardises the exfiltration path for malicious actors. This could mean agentic AI could leave users more vulnerable to attacks. This is especially concerning given how scammers are increasingly using data context around a person to target individuals.

We recommend the FCA considers:

- **Regulation on labelling of AI products linked to financial decision making**, with a focus on transparency of agentic AI risk. This could be akin to food labelling or financial warnings on capital being at risk. It should cover the risks of delegation and possible data risks.
- **Mandated, clear and upfront warnings about liability about agentic AI decisions**. Consumers need to understand upfront if they will be held accountable for decisions made on their behalf by agentic AI.
- **Regulation to ensure that consumers have access to readable records of decision making made by AI, especially agentic AI.** This is vital to support consumer understanding and the ability to seek redress. It is also completely possible, and should be a bedrock of AI regulation to support the ethical principles of explainability. We support calls for a statutory duty to provide a right to explanation in AI driven decision making in financial services.

5. **Fraud: How could AI-driven fraud evolve as consumers increasingly delegate decisions to AI, and what would this mean for consumer agency, harm, and protection in retail financial services?**

We expect fraud and its detriment to increase in scale and size, due to AI-powered technologies. This is due increasingly sophisticated tools, such as deepfakes and voice cloning, how widely available they are to scammers, and how easy they have become to use.

Bank and credit card fraud is growing. ONS data shows that although the overall number of fraud incidents did not change much from 2024 to 2025 (YE September), there was a 19% increase in bank and credit card fraud. Out of 4.2m incidents of fraud, 3.1m involved a loss.

Natwest reported in November 2024 that 42% of British adults have been targeted by scammers in the past 12 months. Their data shows that AI Voice cloning scams was one of the fastest growing scam areas (increase of 30%) and deep-fake celebrity endorsement scams (increase of 22%).

The reach and ingenuity of fraud has escalated, and we expect this trend to continue in the financial services sector.

Real people present in our services, having been victims of AI-generated harm. A recent example in our data includes a client who was scammed for £3,000 after seeing a deepfake advert on YouTube featuring prominent public figures. The client has since taken out a loan that they cannot afford to pay back and borrowed money from a friend, resulting in relationship breakdown.

We are concerned by the increased cybersecurity risk to consumers posed by agentic AI. Agentic AI is offered with advantages of convenience and money saving, which can be beneficial to consumers. But allowing an agent to pool data in the cloud which gives vast context to an individual poses a risk to the consumer. Where agents interact with sensitive data, the risk is increased. Some AI tools have received criticism for its data risks and security researchers are arguing that model Context Protocol (MCP) used in agentic AI standardises the exfiltration path for malicious actors, which could leave users more vulnerable to attacks. The breadth and depth of data pooled by an agentic AI around an individual could be powerful in the hands of a scammer.

We are concerned that consumers are engaging with agentic AI without understanding these data risks. When consumers take on the benefits of a convenient service, they deserve to understand the risk they are taking. When investing capital, we warn customers that it is at risk. AI usage needs to have similar warnings to be used responsibly, considering the power malicious actors could gain from accessing the levels of context required to run agentic AI.

We are also concerned that consumers assume a level of protection from AI that is not aligned with the level of protection that currently exists. For example we are aware of consumers starting to use AI-driven switching services for utilities that offer to find them better deals on household bills and financial products. But already we can see evidence of potential harm, for example where consumers a) did not understand that the agent would act on their behalf b) claims that it acted on their behalf without permission and c) where the AI took decisions customers did not want, that ended up costing them significantly more money. Consumers deserve clarity on liability when decisions are made on their behalf, before they engage an AI service. We're concerned that this is currently a gap in consumer protection.

We are also concerned by consumers who are disproportionately at risk of AI-related harm. All consumers have some risk of AI harm. But, we think there

are two key consumer groups who are more likely to be vulnerable to AI fraud and therefore face a higher risk of harm:

1) **Consumers in vulnerable circumstances**
2) **Users who are less AI-competent and/or less aware of risks**

We believe there is significant overlap between these two groups, with possible compounding effects. Users may also be confident and highly aware of risks. Users may be confident and complacent about risks, placing an over-reliance on AI decision-making for example.

The graphic below outlines where we see groups at highest risk of fraud.

Consumer confidence and risk of susceptibility to scams: a working hypothesis

AI-competent Not risk aware MEDIUM RISK	AI unskilled Not risk aware HIGH RISK
AI-competent Risk aware LOW RISK	AI unskilled Risk aware MEDIUM RISK

Emerging evidence shows that users who are less confident and competent with AI are less likely to experience benefits and more likely to experience harm.

A Dutch study explores how users with these less 'AI-competent' traits are at heightened risk of being influenced, persuaded or manipulated by automated recommendations. Compared to the average user, the most vulnerable groups with the lowest levels of AI knowledge and skills tended to be older, with lower levels of education and privacy protection skills.

Ofcom data shows that only 1 in 10 people are confident in their ability to spot deepfakes (July 2024) and research from Alan Turing showing 90% of survey respondents are concerned about the rise and potential harms of deepfakes. Visa's latest European research reveals that people who mistake AI-generated content for real are almost five times more likely to fall victim to a scam (62% vs. 13%).

Put together, we predict that people with less AI competency, who have less ability to reap the benefits from consumer markets are disproportionately likely to be at risk from AI-generated harm. This harm could be from fraud, but could also be from an inability to reap the benefits of AI, resulting in a widening digital divide.

The increased scale and harm of fraud due to AI will likely encourage digitally excluded people to stay off the internet.

In the UK there are 7.9m people who lack basic digital skills. This means they can't do all 8 tasks of the Essential Digital Skills Framework, including turning on a device, using a mouse and keyboard, setting up a Wi-Fi connection and opening an internet browser. In the UK, there are 1.6m adults who don't have a phone, tablet or laptop and 1.9m households who struggle to afford their mobile contract.

Research shows that fear of being scammed keeps people offline. Almost 3 in 10 (29%) Brits tell loved ones to stay offline to avoid being tricked, which results in increased use of proxy support, with a risk both to the individual and the proxy supporter. We are concerned that as AI driven fraud increases, the digital divide will widen.

We recommend the FCA considers:

- **Regulation on labelling of AI products, with a focus on transparency of agentic AI risk.** This could be akin to food labelling or financial warnings on capital being at risk. It should cover the risks of delegation and possible data risks.
- **Mandated, clear and upfront warnings about liability about agentic AI decisions**. Consumers need to understand upfront if they will be held accountable for decisions made on their behalf by agentic AI.
- **Regulation to ensure that consumers have access to readable records of decision making made by AI, especially agentic AI.** This is vital to support consumer understanding and the ability to seek redress. We support calls for a statutory duty to provide a right to explanation in AI driven decision making in financial services.

Theme 4: Future regulatory approach

1. **Outcomes-based regulation: What are the opportunities and challenges for the FCA in ensuring an outcomes-based approach to retail regulation in an AI-enabled FS industry?**

Outcomes based regulation can be advantageous, as it allows regulators to set out high level principles that they require firms to deliver on. This can help the regulator to stay ahead of potential developments in the market, as it places a responsibility on businesses to demonstrate that their practices and products align with those principles. As more FS providers adopt AI within their delivery this has some benefits, as it ensures that there are expectations around the outcomes businesses are expected to deliver, and enables the FCA to potentially respond more quickly than it may be able to develop new rules.

But outcomes-based regulation is not a binary alternative to prescription. There are areas where a balance of prescriptive rules and outcomes-based rules are necessary to complement each other to drive the right outcomes, avoid customer harm and set baseline expectations and standards. It's important that in its approach to regulation of AI the FCA sets out clear rules that establish:

- **Where the use of AI by financial services firms would not be acceptable and would lead to an unacceptable level of risk or consumer harm-** for example the EU AI Act sets out that this would apply where an AI system uses subliminal techniques to influence financial decision making.
- **Where particular standards or processes must to be in place in order to safely deploy AI systems or use AI within processes** - for example, the FCA should apply very high standards when it comes to:
    - The delivery of advice to ensure that consumers receive advice that is high-quality, reliable and accurate
    - Decisions around credit-worthiness, affordability or whether to offer services to an individual - all of these could result in considerable harm to consumers where inappropriate decisions are made

- Protecting consumers from the risk of fraud, scams and identity theft.

The FCA should also set clear parameters around its expectations of what certain outcomes held in the Consumer Duty require in the context of AI deployment in financial services. One example where we can see a real need for this is in relation to customer support. For example, as more firms deploy AI agents to field queries it's critical that there remain alternative options for consumers to access support. This is essential to ensure that customers are not excluded from accessing support, for example where they don't have access to digital devices, have low digital confidence or have concerns about privacy and trust in AI agents.

This is also essential in guarding against consumer harm, e.g. where digital agents fail to correctly understand a customer's needs, or where consumer circumstances or the particular issue they are facing requires human intervention.

The FCA should set clear expectations, including ensuring alternative channels are available to consumers, that high quality support is available to consumers in vulnerable circumstances and ensuring that there are clear mechanisms for consumers to escalate their queries or complaints to a person where needed.

Low levels of trust with AI are at least in part driven by consumers seeing that sufficient safeguards are not in place. A clear regulatory approach may therefore encourage consumer adoption and support innovation.

This should also be a focus of ongoing monitoring from the FCA to ensure that consumers are getting high quality support from financial services providers.

Wider challenges

The adoption of AI in the financial sector may also pose challenges in identifying where harm has occurred. For example, if AI is used in the context of advising consumers it is unlikely to be clear to consumers if the advice they have received is inaccurate or sub-optimal in their context, so consumers may not report issues. It's therefore important that the FCA goes beyond relying on complaints

data to identify whether harms are occurring in the market, and is therefore taking steps to proactively monitor emerging use of AI in financial services.

2. **Regulatory levers: Are the key FS regulatory levers (Consumer Duty, Operational Resilience, SM&CR, Critical Third Party regime etc) suitable to manage future risks and to enable firms to fully take advantage of AI?**

Given that the Consumer Duty is still a relatively new part of the FCA's regulatory toolkit, it is hard to make a thorough assessment of whether it is suitable in the context of AI deployment in financial services.

Whilst AI deployment in financial services may provide benefits to consumers, there is also a substantial risk of harm to consumers. It is important that the FCA is proactively considering potential risks and considering how these would or would not be mitigated by its existing tools, to identify where it may require new levers.

We expect there may be some areas where stronger intervention may be needed from the regulator, particularly given the high risk of scams enabled by AI which may mean that substantial evolution is required in privacy and security technology to keep consumers safe. This is an area where we would expect the regulator to be working proactively with firms to ensure that the safety mechanisms used to protect UK consumers remain robust.

The FCA may also need to consider whether additional principles need to be captured by the Consumer Duty in the context of AI. For example, we encourage the FCA to boost expectations around transparency to ensure that consumers understand where they are interacting with human decision makers vs automated systems.

It's important that the SM&CR regime is used to ensure that firms remain accountable for AI failures where they did not provide proper oversight. As AI is adopted in new processes the FCA may need to provide more detailed guidance to set clear expectations around what level of oversight it would expect. This could include reference to:

- Where it would always expect human-in-the-loop protocols to be followed, particularly in the context of AI systems or decisions that may carry higher risks.
- Expectations for firms around monitoring of potential biases within AI decisions or processes.
- Expectations around ongoing monitoring and spotchecking. This will be particularly important in the context of identifying potential errors in AI systems or hallucinations. Within this the FCA may need to consider where it would consider an error rate to be unacceptable.
- Requirements for firms to report issues identified in systems - this is likely to be particularly important where firms may be making use of "off the shelf" systems, where early detection and reporting of problems could help to avoid market wide issues.

AI also brings new risks to consumers in financial services where the regulatory levers may not solely be situated within the FCA's current regulatory remit. For example, the widespread use of LLMs by consumers may extend the risks of consumers receiving poor quality or inaccurate advice or information that could affect their financial choices. This may require greater attention around the information/advice boundary.

3. **Supervisory and enforcement approach: Do you have views on the way the FCA should improve or develop its approach to supervision and/or enforcement to respond to increased AI use in the future, including using AI itself?**

The FCA should consider where it can take advantage of AI to support supervision and enforcement - in particular there could be opportunities to improve monitoring by deploying AI tools to improve the speed with which the regulator can detect bad actors in the market. A potential example would be using AI tools to detect where organisations are providing financial services or advice without appropriate authorisation, or monitoring misleading financial promotions.

It's also important that the FCA is proactively monitoring changes in the FS market, so that it can identify where new adoption of AI processes and systems may be transforming consumer outcomes for better or worse. Examples that the regulator should monitor include things like:

- Changes in acceptance and rejection rates for financial services products
- Changes in credit scoring, premia or risk based pricing across the market
- Customer satisfaction and complaints
- For all of the above, changes in how these are distributed across different consumer groups.

4. **Growth: In what ways can the FCA continue to support growth and competitiveness in an AI-driven financial services industry in the future?**

It's important that when considering how to drive growth through AI, this is appropriately balanced against protection for consumers, and ensuring that they are not exposed to unacceptable risks.

5. **Frameworks for inspiration: Are there other regulatory frameworks (UK or international, other non-FS sectors) which the FCA might consider or emulate to respond to increased AI use in retail financial services?**

The FCA should consider the approaches that have been adopted as part of the EU AI Act. In particular we would encourage the FCA to adopt a similar approach setting out situations where it would not consider use of AI to be acceptable, in order to guard against foreseeable risks to consumers.

The FCA should also engage with other UK regulators, particularly with regards to considering what common standards and approaches might be required in relation to cross-cutting issues like consumer support and redress, data security and privacy, and scam prevention.