AIFS0085

Written evidence submitted by Moneybox

Introduction

As a leading British digital wealth management platform, Moneybox empowers individuals across the UK to achieve their financial goals and build wealth with confidence through integrated services, including saving, investing, home-buying, and retirement planning. Launched in 2015, Moneybox is a British tech success story and today over 1million customers trust Moneybox to manage more than £14 billion in assets.

Moneybox has ambitions in the coming years to transfer our AI expertise, currently used heavily in the back-office, to the consumer space and we welcome this opportunity to respond to this consultation. As our submission outlines, it is clear that AI will be revolutionary for consumers' finances in a variety of ways. The impact of AI will be felt particularly in the financial guidance and advice space, which should be an area of focus and attention. We welcome the potential of AI to help better understand and support the diverse needs of vulnerable customers, but it is vital to stress that risks and biases in AI should be effectively identified and mitigated against as the industry moves forward.

How is AI currently used in different sectors of financial services and how is this likely to change over the next ten years?

Artificial Intelligence has been a fixture in financial services for decades. Since the early 2000s, the financial services sector has used AI to complete internal processes to improve workforce efficiency. This has primarily involved automated processes such as credit scoring, portfolio management and fraud detection.

As such, it could be argued that the financial services sector is better prepared for the increased adoption of AI than many other industries.

In the near future, SMEs should benefit from 'off the shelf' AI models, which currently only large companies can afford to develop, but will become more widely available as a service for purchase. However, widespread adoption will require a framework to be developed on risk and a method for insurers to underwrite the risk of generative AI if it was to be deployed at scale to customers.

Currently, the biggest barrier to the expansion of AI in the financial services sector is that companies struggle to gauge the risk of using AI within the regulatory framework. If companies are unable to gauge the risk associated with a product, they are unlikely to use it. This is especially true with the use of generative AI (GenAI) and its potential to generate false or misleading outputs, often referred to as 'hallucinations,' in a consumer-facing environment.

Nevertheless, as GenAI improves in the next ten years, we expect the greatest change in the financial services sector to be in the consumer-facing space. Financial services will

become simpler for consumers as more sophisticated virtual assistants and chatbots support potential queries at scale, with scaled oversight of human operators and agents.

To what extent can AI improve productivity in financial services?

Moneybox has always prioritised the customer when considering our product and underlying processes. Looking to the future, we believe that the best use cases for AI will be consumer-orientated:

- **Customer Services**
  - A good use case for AI will be a highly repeatable process that does not materially change from customer to customer.
  - The productivity of the frontline element of the service industry is going to dramatically increase as more capable AI models are deployed at scale.

- **Addressing the Financial Advice Gap**
  - There is currently a significant financial advice gap with only 8% of consumers regularly availing of financial advice. This is primarily driven by the high operational costs associated with financial advisors and the requirement to effectively pass these costs directly to the consumer via adviser charging rules, which limits access for entry-level customers.

    Over the next decade, AI has the potential to transform this model, enabling advisors to serve thousands of clients more efficiently by handling the analytical workload. This may be under the current regulated advice framework, or through personalised guidance, depending on the complexity of customer need. This democratisation of financial advice is already emerging, as evidenced by consumers' increasing use of AI platforms such as ChatGPT for advice.

    However, non-regulated AI platforms lack important controls (and have no licence to give financial financial advice) which is an area of concern. In addition, their safeguards are very generic (as they are generalised models). In contrast, specialised models would be able to deal with guardrails and specify information in a way that is consistent with regulatory requirements.

  - Given this, in the future, we can expect regulated financial services firms to provide AI-assisted advice with appropriate safeguards in place to ensure the accuracy and reliability of recommendations, offering a more secure alternative to unregulated AI tools.

What are the risks to financial stability arising from AI and how can they be mitigated?

Whilst AI will undoubtedly make ‘good actors’ within the financial services sector more efficient, it will simultaneously make the work of ‘bad actors’ just as efficient. It is vital that financial regulators are aware of the issues and are equipped to combat them.

- **Scale of Scamming Operations**
  - With the rise of GenAI, scammers will easily be able to scale up their operations. Whilst currently, scamming operations might require a large workforce, it could be possible in the future to automate a significant proportion. This will make their operations more efficient and will increase the number of threat actors operating in the scamming sector. As the number of threat actors rises, this will lead to additional pressures on consumers and, thereby, increasing the pressure on financial institutions to fight the scamming risk.

- **‘Backdoor’ Security**
  - By leveraging GenAI, engineers can enhance their productivity by utilising it for code development.
  - However, this introduces security considerations. As these AI models are often sourced from external providers, as part of the rising AI as a service for purchase, there is a proven risk that third-party developers could embed hidden instructions within the AI. These concealed directives have the potential to create backdoors in the generated code, introducing a new security vulnerability that organisations will have to carefully mitigate.
  - Several studies have demonstrated the scale of this problem, with evidence demonstrating vulnerability across various parameter sizes. In addition, there is evidence that current safety techniques can ignore security vulnerabilities and can create a false impression of safety.

- **Hallucinations**
  - ‘Hallucinations’ refer to a response generated by AI that contains false or misleading information presented as fact. Today, there's less risk simply as consumers do not trust AI. In a future where a trust relationship has been formed with providers of AI, the risk of hallucination escalates.

- **Financial Stability**
  - The number of private traders using AI for advice will proliferate in the coming years. Therefore, there is growing potential for an increased number of unusual shifts in market dynamics as when previously the only component to market dynamics had been human behaviour, now you have AI prompting human behaviour. This could in turn lead to herding behaviour, which has the risk of destabilising financial institutions.

What are the benefits and risks to consumers arising from AI, particularly for vulnerable consumers?

Much of the financial services sector has already optimised its internal algorithms using AI. The main future potential for AI is its ability to enhance financial services for individual consumers by simplifying complex financial language and tailoring communication. Through the use of AI-powered tools, firms can explain products and financial concepts more effectively at scale. This accessibility ensures that consumers, particularly vulnerable customers, can gain a clearer understanding of their financial decisions.

AI could be revolutionary in helping businesses identify and support vulnerable consumers. Historically, addressing the full complexity of vulnerable consumers (VC) has been challenging due to the limitations of supporting VCs at scale within traditional systems and because characteristics of vulnerability are often nuanced and interrelated. In the future, AI's capability to analyse large datasets could offer companies a more nuanced understanding of various dimensions of vulnerability through its use of improved language.

By recognising multiple types and levels of vulnerability, financial firms can implement more personalised and effective support strategies for vulnerable consumers. Harnessing the potential of AI in this area will support early detection of financially vulnerable consumers and help predict levels of risk. This will allow companies to provide personalised and fast support to a range of vulnerable individuals, significantly improving experience and outcomes for consumers and customers.

Nevertheless, just as AI will help companies identify and support vulnerable consumers, there are several risks that the financial services sector will need to be wary of. One primary concern is the risk of 'hallucinations.' Consumers currently using GenAI tools for financial advice, such as ChatGPT, may receive inaccurate or misleading financial information. In the future, regulated firms employing AI for advice will bear responsibility for the accuracy of AI-generated content. Regulatory frameworks should ensure firms implement strict oversight, including multiple 'lines of defence,' before AI-driven advice reaches consumers.

One cause of hallucinations is the increasing embedded bias in AI models, as these models are inherently influenced by the data on which they are trained, reflecting existing societal biases. Unlike humans, whose biases can be directly addressed through training and oversight, mitigating bias in AI necessitates complex interventions. If left unchecked, AI bias may result in unfair or discriminatory financial decisions, disproportionately affecting vulnerable consumers. The risk of bias would be particularly acute with firms using smaller data sets. Encouraging industry-wide collaboration to create shared datasets could democratise AI development and promote a more innovative landscape to the benefit of the consumer.

How can Government and financial regulators strike the right balance between seizing the opportunities of AI but at the same time protecting consumers and mitigating against any threats to financial stability?

As AI develops and becomes more sophisticated, Government and financial regulators must maintain oversight over the training and validation of AI models. It is essential that further resources are provided to create a centralised independent AI regulatory sandbox. This would have a number of benefits including:

- **Independent Data Sets for Validation:** AI systems should undergo rigorous validation using data that is not from the internet and, therefore, has not been included in their training. This necessitates the creation of independent data sets to establish objective benchmarks.

- **Centralised Validation Systems:** A centrally managed evaluation system would enable regulators to oversee AI performance consistently. Without such a system, firms may validate their models in isolation, leading to inconsistent standards and a lack of comparability across the industry.

- **Bias Detection and Safeguards:** By implementing common evaluation metrics and benchmarks, regulators can more effectively identify and mitigate biases in AI systems, enhancing consumer protection.

Nevertheless, there is a fine line to walk between the necessary validation checks for AI models and over-regulating. The regulation of AI should not be focused on the AI itself, but rather on the consumer impact, similar to existing regulations for other financial products and it should be noted this is the approach taken by the FCA in their 2024 AI paper.

Finally, if we want the implementation of AI in financial services to be beneficial for consumers, algorithms must move away from the traditionally used “tractable” models. A tractable model is one where the exact process can be traced from start to finish in a linear fashion. These tractable models hinder the outcome for consumers as they limit the scale to which they can be deployed.

As an industry, financial services should be changing their perception towards AI algorithms. Instead of exclusively using tractable models and being solely concerned about the linear decision-making process, companies should be allowed to evaluate how much risk they can manage with AI models.

Without this, GenAI will never be used in a consumer setting which would be to the detriment of consumers. Therefore, the FCA should instead be looking towards regulating based on positive customer outcomes rather than exclusively tractability. This would have the potential to significantly improve the financial outcomes for millions of consumers.

April 2025