

‘Hallucinations’ in AI refer to instances where large language models (LLMs) infer patterns in input data that do not exist, generating confident sounding but incorrect outputs. Because LLMs are designed to produce the most statistically plausible response, not necessarily a factual one, they can replicate errors in their training data or produce entirely fabricated information.

This presents a significant challenge for the pensions sector as AI use expands. Risks are particularly acute when savers interact directly with AI systems, such as chatbots offering personalised financial guidance. Many schemes now deploy AI that uses member and scheme data to generate personalised guidance in conversational chatbot style formats. Good scheme governance should ensure that any third-party AI integrated into scheme operations is enterprise level – i.e. designed for organisational needs and equipped with stronger safety guardrails to reduce inaccurate outputs.

But the risk is never zero. All current generative AI models can produce inaccuracies, and incorrect information provided to scheme members could lead to uninformed decisions with financial consequences.

The risks are heightened for vulnerable consumers. These savers may be less able to recognise mistakes in AI outputs or understand that AI can be unreliable. Inaccuracies also create financial and reputational risks for schemes, which may face liability for consumer detriment if AI driven guidance leads to harm.

However, there is a counterargument: AI tools can improve access to much needed financial guidance to people who might otherwise receive none. For individuals already making important decisions without advice, low-cost, scalable AI chatbots can offer quick, accessible support. In some cases, the small risk of AI error may be outweighed by the broader access and preventative

Confident, but inaccurate

▶ **Olivia Sizeland explains how pension schemes can manage the risk of AI hallucinations**



guidance these tools deliver – for example, helping members avoid harm when making decumulation decisions.

Demand for guidance continues to grow as savers take more responsibility for managing their retirement outcomes. The FCA’s Advice Guidance Boundary Review sought to address these gaps but remains ongoing. In the meantime, there is evidence that consumers are already turning to mass market AI tools such as ChatGPT or Gemini to help with financial decision-making.

This trend will accelerate with pensions dashboards. As savers access more of their own data more easily, LLMs will too. Consumers are already inputting personal financial information into general purpose AI tools, which then generate ‘personalised’ recommendations about contribution levels, retirement income solutions, or investment decisions. This is risky: These models are not designed for regulated financial guidance and are far more likely to produce incorrect outputs. Their training data often includes large volumes of internet scraped content, such as a Wikipedia page, some of which contain inaccuracies. If flawed information enters the model, flawed recommendations follow.

Enterprise level LLMs, by contrast, use carefully curated datasets – such

as pensions legislation and regulated guidance – and therefore tend to be significantly more accurate. But consumers rarely distinguish between the two.

A further challenge arises from exploitative practices. Some nefarious companies are publishing public webpages that are hard for human users to find. They contain exaggerated claims about products, such as asserting that a particular decumulation solution is the best on the market. LLMs ingest this content and may repeat it as fact. As a result, AI models may endorse products without considering an individual’s full financial circumstances, unlike a regulated adviser.

Regulation compounds this disparity. Consumer-facing tools like ChatGPT are bound only by general UK law, such as GDPR. There is no statutory AI regulation specific to financial advice contexts, leaving consumers unprotected if they rely on AI for retirement decisions.

Pension schemes using enterprise AI, however, must meet standards set by regulators such as TPR and the FCA. The General Code requires due diligence on service providers, and FCA rules apply to firms using AI in regulated contexts.

When trained on high-quality data and used within regulated markets, AI can significantly improve saver outcomes. But without accessible, reliable tools, consumers will continue using unregulated mass market AI to support major financial decisions. The sector must consider how to mitigate this growing source of consumer risk.

▶ **Written by Pensions UK senior policy adviser, Olivia Sizeland**