

Understand AI before trusting it

Artificial intelligence is moving rapidly from something that once seemed futuristic into an everyday workplace tool. Yet, as adoption grows, so does the need to understand what sits behind the technology – and where its limitations lie. “Before you can understand where the risks lie with your tool, before you can properly make maximum use of the tool in the most appropriate and responsible way, you need to understand what it can do, what are the limits, and what are the no-go areas completely,” said Azhar Aziz-Ismail, AI Training and Development Lead, Practice Innovation at Baker McKenzie, speaking at the 16th Annual FISA Conference sponsored by Hollard Court Bonds.

Aziz-Ismail, a disputes lawyer by background, works in AI and practice innovation at the global law firm. He argues that one of the most useful ways to understand AI is simply to use it – provided it is done responsibly and appropriately.

The black box problem

It’s important to remember that AI doesn’t operate like conventional software, which follows predefined instructions. AI systems learn from data and can generate new content based on patterns and predictions. This distinction becomes particularly important with large language models (LLMs), which underpin many of the generative AI tools now being used in business.

One of the challenges with LLMs is that they can behave in ways that aren’t always predictable. They are probabilistic rather than deterministic, meaning there is no fixed sequence of rules guaranteeing the same type of response every time. This creates several familiar problems, including hallucinations, where AI produces information that sounds convincing but is incorrect.

There is also the problem of bias. Models are trained on data, and the quality and characteristics of that data can influence their outputs. Users therefore need appropriate safeguards and should not assume that a confident-sounding answer is necessarily correct.

Not all AI tools are the same

Another important distinction is between open and closed AI environments. Open tools may draw on information available across the internet, while closed systems operate within a defined body of information. For professional applications, particularly



where accuracy and jurisdiction matter, understanding where an AI system is getting its information becomes critical.

A South African professional asking an AI tool to conduct research, for example, may need to specify that it should focus exclusively on South African law. This is particularly important because AI performance is uneven. Systems may be highly effective at summarisation, translation and basic drafting, while struggling with more complex reasoning or research. Aziz-Ismail referred to this as “jagged intelligence” – the reality that AI can be very good at some tasks and considerably weaker at others.

The risks behind the efficiency

Greater efficiency doesn’t remove professional responsibility. In fact, the faster AI can produce an answer, the easier it may become to overlook mistakes. One of the biggest risks is over-reliance. Professionals can become accustomed to trusting AI output in much the same way they may begin to skim the work of a junior colleague once confidence has been established.

Aziz-Ismail argued that every important output needs to be checked against a reliable source, including citations, names and references. AI can produce an answer that looks authoritative, complete with a footnote or citation, while the underlying information is wrong. A confident presentation doesn’t make an answer accurate. There are also concerns around confidentiality, intellectual property, legal privilege, data privacy and cybersecurity.

Simply paying for an AI tool doesn’t automatically make it secure. Professionals need to understand where data is processed, who can access it, whether inputs are retained or used for training, where servers are located, and what contractual protections apply. For organisations handling sensitive information, enterprise-grade tools may offer greater controls, but users still need to understand the terms and limitations of the technology.

Keep the human in the loop

For Aziz-Ismail, the dividing line is between augmentation and abdication. Augmentation means AI performs part of the work while a human reviews and takes responsibility for the output. Abdication occurs when AI generates an output that reaches the client without appropriate human oversight. “The difference is not immediately visible in the document,” he said. “It’s only visible in the process that sits behind it.”

That distinction is particularly important in financial and professional services, where advice involves judgement rather than simply retrieving information. Aziz-Ismail suggested a traffic-light framework. Green tasks can include drafting, summarising, information retrieval and translating technical language into plain English, provided the output is reviewed. Amber tasks involve analysis and comparison, where verification should be explicitly recorded. Red tasks involve professional discretion, advice and final sign-off – areas where human judgement remains essential.

Four questions can help determine where a task belongs: Is judgement involved? Can the output be verified against a trusted source within a reasonable time? Whose data is being used, and is it lawful to put that information into the tool? And would you be comfortable signing the resulting work in your own name? If the answer to any of these questions is no, greater human oversight is required.

Governance matters

Organisations need governance around AI. That means using approved tools, recording review steps, training staff, supervising AI use and assigning responsibility for each tool. For client-facing businesses, it may also mean considering appropriate AI provisions in engagement terms. As Aziz-Ismail put it, the question is not whether AI should be used, but where it should be used, how it should be used and where humans must remain firmly in control.