INNESSCO LEGAL TECHNOLOGY WHITE PAPER

AI and Legal Practice

What Lawyers Need to Understand About AI Capability, AI Literacy and Professional Responsibility

Developed following the Federal Court of Australia AI Symposium August 2026

Robert Cox & David Loughry
Directors, Innessco

August 2026

Table of Contents

1. EXECUTIVE SUMMARY 4

2. INTRODUCTION 5

3. A STRIKING DEMONSTRATION OF HOW QUICKLY AI LEGAL CAPABILITY IS IMPROVING 6

4. THE LEGAL PRODUCTIVITY OPPORTUNITY GOES WELL BEYOND DRAFTING 7

5. WEATHERALL: THE PRACTICE NOTE STARTS WITH "KNOWLEDGE" 8

6. FROM "CAN MACHINES THINK?" TO CHATGPT AND CLAUDE 8

6.1 1950: Alan Turing asks the fundamental question 8

6.2 1956: artificial intelligence becomes a field 9

6.3 Neural networks: instead of writing the rules, let the machine learn them 10

6.4 Periods of enthusiasm and disappointment 10

6.5 1997: a machine defeats the world chess champion 11

6.6 2012: deep learning accelerates 11

6.7 2017: the Transformer changes natural-language AI 12

6.8 2018–2020: large language models begin to show general capability 12

6.9 2022: models learn to follow human instructions 13

6.10 2023–2026: from chatbot to professional work platform 13

7. WHAT DO WE MEAN WHEN WE SAY "AI"? 14

7.1 Artificial intelligence 14

7.2 Machine learning 14

7.3 Foundation models 15

7.4 Large language models 15

7.5 Generative AI 15

7.6 Multimodal AI 16

7.7 AI assistants and chatbots 16

7.8 Retrieval and grounding 17

7.9 AI agents 17

7.10 Agentic AI 18

7.11 A useful mental model 18

8. WHAT IS A LARGE LANGUAGE MODEL ACTUALLY DOING? 19

8.1 Step 1: text is converted into tokens 19

8.2 Step 2: the model learns patterns in language 19

8.3 Step 3: the Transformer considers context 20

8.4 Step 4: the model generates its answer 21

9. WHY DOES AI HALLUCINATE – AND WHY CAN THE SAME PROMPT PRODUCE DIFFERENT ANSWERS? 21

9.1 Plausibility and truth are different things 22

9.2 The model may know the pattern better than the underlying fact 22
9.3 Training information can be incomplete, inconsistent or outdated 22
9.4 The model responds to the framing of the question 23
9.5 Why can the same prompt produce different answers? 23
9.6 Fluent writing can create an illusion of certainty 24
10. HOW SHOULD A LAWYER USE AI IN PRACTICE? 25
10.1 Define the task and its boundaries 25
10.2 Ask for analysis rather than confirmation 25
10.3 Require the AI to show the basis for important conclusions 25
10.4 Tell the AI what to do when the information is insufficient 26
10.5 Break complex tasks into stages 26
11. THE OPPORTUNITY FOR BOUTIQUE LAW FIRMS 26
12. THE REAL LESSON FROM THE FEDERAL COURT SYMPOSIUM 27

1. Executive Summary

Artificial intelligence is rapidly moving from an experimental technology to a practical tool capable of performing substantial elements of legal work. The Federal Court of Australia's AI Symposium on 19 August 2026 reinforced two important messages for the legal profession: the capability of AI is improving extraordinarily quickly, and lawyers who use it are expected to understand how it works, where it can fail and how to use it responsibly.

Dr Matthew Collins AM KC illustrated the pace of improvement through a series of experiments involving AI completing a master's-level law examination. In late 2025, he used Claude Opus 4.5 to complete the examination. Without attending any of the course lectures, the AI achieved approximately 70 per cent, compared with an average student mark of 74 per cent. When Collins repeated the experiment three times in the weeks before the August 2026 symposium using newer frontier models, he considered each result to be of unquestionable H1 standard, with properly formatted legal citations and no discernible hallucinations. The experiment demonstrates both the legal capability of current AI systems and the speed at which that capability is developing.

Collins also demonstrated practical legal applications, including using AI to analyse a 2,205-page trial transcript in approximately five minutes, identify relevant witnesses, compare their evidence, highlight inconsistencies and provide links back to the source transcript. His central point was not that AI should replace the judgement of counsel, but that it can perform large amounts of searching, organisation and first-pass analysis extremely quickly, allowing lawyers to focus their time on interpretation, strategy and professional judgement.

Professor Kimberlee Weatherall focused on the corresponding obligation of lawyers to develop AI literacy. She emphasised that the Federal Court Practice Note begins with knowledge: practitioners are expected to understand the capabilities and limitations of the technology. In her view, it is not sufficient simply to know that generative AI can hallucinate. Lawyers should understand, at a practical level, why hallucinations occur, how large language models generate their answers and why some limitations remain inherent in the technology.

At a basic level, modern large language models do not operate as conventional databases of verified facts. They are trained on enormous volumes of information to identify patterns and relationships in language and then generate responses by predicting appropriate sequences of words or "tokens" from the available context. This explains both their extraordinary capability with language-intensive work and one of their principal weaknesses: a response can be highly plausible, articulate and professionally presented without necessarily being factually correct. A fabricated authority may, for example, look entirely authentic because the model understands the structure and language of legal citations even where it does not have reliable information about the particular authority.

For lawyers, the practical response is not to reject the technology but to use it in ways that reduce these risks. AI becomes considerably more reliable when it is grounded in authoritative source material — for example, when asked to analyse the actual contract, judgment, transcript, correspondence or matter documents rather than relying solely on information absorbed during the model's original training. Important outputs must nevertheless remain subject to human verification, particularly where legal authorities, evidence, factual propositions or advice are involved.

For boutique law firms, the commercial implications are significant. AI can reduce the time required for document review, comparison, summarisation, chronology preparation, first-pass legal analysis, citation checking and drafting. This potentially enables smaller firms to undertake work that previously required substantially larger teams, improve turnaround

times and fixed-fee profitability, and allow senior lawyers to devote more of their time to client advice, negotiation, strategy and judgement.

The key message is therefore neither that AI will replace lawyers nor that AI is too unreliable to use. Rather, current AI systems are becoming capable enough to materially improve the productivity and quality of legal work, but their effective use depends upon knowledgeable lawyers who understand the technology's limitations, select appropriate tools, provide reliable source information and retain responsibility for the final work.

As the Federal Court's approach makes clear, the technology may change rapidly, but the underlying professional obligation does not: the lawyer remains responsible for understanding the tools being used, exercising independent judgement and ensuring that the work ultimately relied upon is accurate.

2. Introduction

Artificial intelligence has moved remarkably quickly from being an interesting emerging technology to something that is already capable of performing substantial elements of professional legal work.

For law firms, this creates two competing risks. The first is obvious: using artificial intelligence poorly can result in inaccurate advice, fabricated authorities, breaches of confidentiality and other failures of professional responsibility. The second risk receives less attention: firms can become so focused on what can go wrong with AI that they fail to understand how capable the technology has become and consequently fail to take advantage of potentially significant improvements in legal productivity.

Both issues were central to the Federal Court of Australia's AI Symposium held on 19 August 2026.

The Court's position is not that lawyers should avoid generative AI. Its recently issued Use of Generative Artificial Intelligence Practice Note (GPN-AI) expressly recognises that generative AI has the potential to increase efficiency, reduce legal costs, improve access to justice and enhance the quality of the administration of justice. At the same time, the Court expects anyone using generative AI in connection with Federal Court proceedings to have "a basic understanding of its capabilities, and its limitations and risks."

Two discussions during the symposium are particularly helpful in understanding what this means in practice.

The first was a series of experiments described by Collins showing just how rapidly the legal capabilities of leading AI systems are improving.

The second was Weatherall's observation that the Federal Court Practice Note starts with knowledge. Her point was that it is no longer sufficient for lawyers simply to have heard that AI sometimes "hallucinates". Lawyers who use the technology need some understanding of how it works, why hallucinations occur, where its limitations come from and how those limitations should influence the way AI is used in legal practice.

That is an important development.

A lawyer does not need to become a computer scientist. However, just as a lawyer does not need to be an engineer or doctor to work effectively with expert evidence, they do need enough understanding to ask the right questions, recognise the limits of the expert's methodology and assess whether the conclusions can safely be relied upon. The same principle applies to generative AI: lawyers increasingly need sufficient AI literacy to understand what the system is doing and where its limitations lie.

3. A striking demonstration of how quickly AI legal capability is improving

Collins gave perhaps the clearest demonstration during the symposium of why law firms should be taking developments in AI seriously.

Collins described an experiment he conducted late in 2025 while teaching a master's-level law subject at Melbourne Law School.

He used Claude Opus 4.5, which he described as a frontier AI model at the time, to create a simulated student response to an examination. The AI was supplied with the course materials but, obviously, had not attended any of the approximately 25 hours of lectures delivered during the subject. It therefore lacked some of the contextual advantages enjoyed by the actual students, including knowing which issues and cases had been particularly emphasised during teaching.

Collins placed a fictitious student number on the examination paper and inserted it into the normal marking process without initially telling his teaching colleague that the answer had been produced by AI.

The result was 70 per cent, which Collins characterised as an H2B result. The average student mark was approximately 74 per cent.

Importantly, Collins said he had not changed a word of the AI-generated substantive answer. His intervention was limited to some formatting of the footnotes.

At the time, Collins regarded the result as impressive but also somewhat reassuring. One way of looking at it was that AI might be capable of producing a competent answer, but the challenge for a good student remained demonstrating the additional knowledge, reasoning and judgement necessary to lift an H2B answer into H1 territory.

Then the technology changed.

In the three weeks before the August 2026 symposium, Collins repeated the experiment three times using newer frontier AI models.

He reported that on each occasion the result was, in his assessment, unquestionably H1 standard. He described the resulting documents as well formatted, with properly presented footnotes and citations compliant with the Australian Guide to Legal Citation, and said he could identify no discernible hallucinations.

Collins did not identify the particular models used for those three later tests, so it would be wrong to attribute those results to Claude or any other particular product. The important point was the change in capability.

Collins' broader observation was that the progress in AI had become extraordinarily rapid. He said that, in his view, more progress had occurred during approximately the preceding six months than perhaps during the period since the introduction of ChatGPT 3.5, and that AI was already beginning to reshape legal practice in a manner reminiscent of the arrival of email and the internet.

Independent evidence supports the broader proposition that frontier AI capability is advancing rapidly. Stanford University's 2026 AI Index reports substantial improvements during 2025 across language, reasoning, coding, mathematics and other demanding benchmarks, while also cautioning that benchmarks themselves are increasingly struggling to keep pace with model capability.

For a law firm, this has an important practical consequence:

An assessment of AI made twelve months ago – or even six months ago – may no longer be a reliable assessment of what current systems can do.

A partner who tested an AI tool in 2024 or 2025 and concluded that it produced mediocre drafting, struggled with legal analysis or fabricated too many authorities may have been entirely correct at the time. It does not follow that the same conclusion should be applied to the models available today.

AI evaluation therefore needs to become an ongoing process rather than a one-off technology decision.

4. The legal productivity opportunity goes well beyond drafting

The examination experiment was not the only demonstration of capability during the symposium.

Collins also demonstrated AI analysing the 2,205-page transcript of the Lehrmann v Network Ten proceeding. He asked the system to identify witnesses who had given evidence concerning a particular event, extract the important components of their evidence, identify consistencies and inconsistencies, organise the results into a table and consider what could responsibly be said about witness credit.

Within approximately five minutes, the system had produced an analysis linked back to the underlying transcript passages so that the lawyer could inspect the source material supporting the propositions it had identified.

Collins' point was not that AI should decide whether a witness is credible.

His point was almost the opposite.

A barrister who had lived through the trial and understood the evidence could use AI to undertake the enormous mechanical task of searching, collating and comparing the material, while retaining responsibility for assessing what it meant. He contrasted five minutes of machine-assisted analysis with work that might otherwise occupy a junior lawyer or junior barrister for much of a night at the end of a long trial.

That distinction is important when considering AI within a boutique law firm.

The most valuable applications may not be asking an AI:

"What is the answer to this legal problem?"

They may instead be questions such as:

- Analyse these 800 documents and identify every reference to the termination provision.
- Compare the accounts given by these six witnesses and identify inconsistencies.
- Review this affidavit and identify possible evidentiary objections for the solicitor to consider.
- Summarise the chronology from the correspondence and link each proposition to its source.
- Compare this draft agreement against our standard precedent and identify substantive changes.
- Check every authority cited in these submissions and identify the proposition for which it is relied upon.
- Review the matter file and prepare a first-pass briefing identifying outstanding issues.

- Rewrite this first draft so it conforms to the firm's established style.

These are often labour-intensive tasks involving searching, sorting, comparison, classification and synthesis.

They are precisely the kinds of tasks at which modern AI can provide significant augmentation while leaving the ultimate legal judgement with the lawyer.

5. Weatherall: the Practice Note starts with "knowledge"

Against this backdrop of rapidly increasing capability, Weatherall made an equally important point about professional responsibility.

Discussing the Federal Court's Practice Note, she drew attention to its first substantive expectation: knowledge.

She emphasised that users are expected to understand the limitations of the technology. But she went further. Lawyers should not simply be told:

AI hallucinates. Be careful.

They should understand, at least at a practical level, why it hallucinates.

Weatherall argued that lawyers should learn broadly how the technology works, why its particular limitations arise and why some of those limitations are likely to remain relevant even as the systems become more capable. She described understanding the technology as empowering because it removes some of its apparent "magic".

The Practice Note itself makes the requirement explicit. Anyone using generative AI in Federal Court proceedings is expected to have a basic understanding of its capabilities, limitations and risks. It also warns users that generative AI may produce fictitious cases and citations, incorrect statements of law, factual errors and even confident confirmation that inaccurate information is correct.

What, then, does a lawyer actually need to understand?

To answer that question it helps to understand how we arrived at today's generative AI systems.

6. From "Can machines think?" to ChatGPT and Claude

Artificial intelligence did not begin with ChatGPT.

The intellectual history extends back more than seventy years, and many of the ideas underlying today's systems are decades old.

6.1 1950: Alan Turing asks the fundamental question

In 1950, British mathematician and computer scientist Alan Turing published his famous paper *Computing Machinery and Intelligence*.

Rather than becoming trapped in the philosophical problem of defining what it meant for a machine to "think", Turing proposed what became known as the *Imitation Game*, and later the Turing Test: could a machine converse in such a way that a human evaluator could not reliably distinguish it from another human?

Turing predicted that within about 50 years computers would become sufficiently capable that, after five minutes of questioning, an average interrogator would have no more than about a 70 per cent chance of correctly distinguishing the machine from a human. He also

predicted that attitudes would change to the point where people could speak of machines as "thinking" without expecting to be contradicted. This was a remarkable prediction in 1950, when electronic computing itself was still at a very early stage of development.

Turing also considered how such a machine might be created. Rather than attempting to program every element of an intelligent adult mind, he proposed building what he called a "child-machine" – a comparatively simple system capable of learning, which could then be educated and improved through experience, including forms of reward and punishment. The important idea was that intelligence would not be created by programmers manually writing every rule the machine should follow. Instead, the machine would progressively learn from experience and feedback. Although Turing could not have anticipated modern large language models, this idea – that sophisticated behaviour could emerge through learning rather than explicit programming – became one of the fundamental principles underlying modern artificial intelligence.

6.2 1956: artificial intelligence becomes a field

The term "artificial intelligence" is generally traced to a proposal prepared in 1955 for the Dartmouth Summer Research Project on Artificial Intelligence, which was held in 1956. John McCarthy, Marvin Minsky, Nathaniel Rochester and Claude Shannon were among the central figures involved. The proposal advanced the ambitious idea that aspects of learning and intelligence could, in principle, be described sufficiently precisely for a machine to simulate them.

Early AI research was highly optimistic. Researchers attempted to develop systems capable of solving mathematical problems, proving theorems, playing games and reproducing aspects of human reasoning. Much of this work followed what became known as symbolic AI – an approach in which knowledge and reasoning were represented using symbols, rules and logical relationships.

The basic idea was that intelligent reasoning could be expressed through explicit rules:

If X is true and Y is true, then conclude Z.

This approach is readily understandable to lawyers. A legal problem can often be expressed in similar terms: identify the relevant rules, establish the facts, determine which rules apply and reason towards a conclusion. In a symbolic AI system, programmers attempted to encode large bodies of this type of knowledge and the logical relationships between them.

Symbolic AI achieved significant early successes. One of the best-known examples was Logic Theorist, developed by Allen Newell, Herbert Simon and J.C. Shaw in the mid-1950s, which was able to prove mathematical theorems by applying logical rules and searching through possible solutions. The same rule-based approach later developed into expert systems capable of applying specialist knowledge in fields such as chemistry and medicine. Its limitations, however, became increasingly apparent when researchers attempted to apply the approach to the complexity of the real world. Human knowledge contains ambiguity, exceptions, context and assumptions that are difficult to capture comprehensively through explicitly written rules.

The central challenge was therefore one of scale and complexity: could intelligence really be created by attempting to tell a computer every rule it needed to know? Increasingly, researchers began exploring a different approach – rather than programming all of the rules, could a machine learn patterns for itself from examples and experience?

6.3 Neural networks: instead of writing the rules, let the machine learn them

Running alongside symbolic AI was a very different idea, inspired loosely by the way biological neurons interact.

Rather than programmers explicitly writing every rule, a neural network could learn patterns from examples. Early work in this area included mathematical models of artificial neurons in the 1940s and Frank Rosenblatt's perceptron in the late 1950s.

Early neural networks were generally developed to perform relatively narrow, clearly defined tasks. A network might be trained to recognise a particular type of image, classify a signal, or decide whether an input belonged to one category or another. The problem was usually defined first, and the network was then trained on examples relevant to that specific task.

A neural network is made up of layers of processing units connected by adjustable numerical values known as weights. A weight can be thought of as a measure of influence: it determines how strongly one piece of information affects what happens next inside the network.

For example, a simple system being trained to identify spam email might learn that several suspicious links should strongly increase the likelihood that a message is spam, while the fact that the sender is a trusted contact should reduce that likelihood. The programmer does not usually specify these relationships manually. Instead, the system learns them by examining many examples, making predictions and adjusting its weights when it gets the answer wrong.

In 1986, David Rumelhart, Geoffrey Hinton and Ronald Williams published influential work demonstrating the effectiveness of back-propagation, a method for adjusting the weights in multi-layer neural networks so that future predictions improve.

This created an important distinction between traditional software and machine learning. With traditional software, a programmer might write:

If the amount is greater than $50,000, apply Rule A.

With machine learning, the programmer does not necessarily specify every relevant rule. Instead, the system learns statistical relationships from data.

In modern neural networks, these learned relationships are spread across enormous numbers of weights. An individual weight does not usually correspond neatly to a word or concept. Instead, complex patterns emerge from the interaction of many weights across many layers.

This approach is far more flexible than trying to manually encode every rule, but it is also less transparent. A neural network can learn highly sophisticated relationships without those relationships being easily reducible to a set of human-readable rules.

Symbolic AI attempted to give the machine the rules. Early neural networks instead learned patterns from examples, usually to solve a specific task. Modern AI builds on the same principle at a vastly greater scale.

6.4 Periods of enthusiasm and disappointment

AI development did not proceed in a straight line. The field experienced repeated cycles of optimism followed by disappointment, including periods in the 1970s and again in the late 1980s and early 1990s commonly described as AI winters. During these periods, technical limitations, insufficient computing power and expectations that had run ahead of practical capability caused investment and enthusiasm to decline.

Expert systems became commercially important during the 1980s, but they also exposed the limitations of rule-based AI. Building and maintaining them required large amounts of

specialist knowledge to be manually converted into explicit rules, which became increasingly difficult as problems grew more complex.

Over the following decades, however, three things began to change dramatically:

data, computing power and algorithms.

The growth of the internet produced enormous quantities of digital information. Graphics processing units, originally developed principally for computer graphics, proved extraordinarily effective at performing the parallel mathematical calculations required to train neural networks. At the same time, researchers developed better techniques for training much larger and deeper networks.

By the 2010s, these three developments had begun to reinforce one another, creating the conditions for the rapid advances in deep learning that followed.

6.5 1997: a machine defeats the world chess champion

A highly visible milestone occurred in 1997 when IBM's Deep Blue defeated reigning world chess champion Garry Kasparov in a six-game match.

Deep Blue could evaluate approximately 200 million chess positions per second, combining enormous computational power with specialised algorithms and techniques for evaluating chess positions. It represented an extraordinary achievement, but its intelligence was narrowly focused: Deep Blue had been specifically designed to play chess. Its remarkable ability at chess did not give it the ability to draft a letter, analyse litigation, translate a document or hold a conversation.

This distinction between narrow AI and today's much more general-purpose AI systems is important. For most of AI's history, systems were developed to perform particular tasks, and expertise in one task generally did not transfer to unrelated problems.

Deep Blue could outperform the world's best human chess player at chess, but it could do little else.

6.6 2012: deep learning accelerates

Another major milestone came in 2012 when Alex Krizhevsky, Ilya Sutskever and Geoffrey Hinton entered a deep neural network, later known as AlexNet, in the ImageNet image-recognition competition.

AlexNet combined a large neural network, approximately 1.2 million labelled training images and GPU computing to achieve a major improvement over competing image-classification systems. Its top-five error rate was approximately 15.3 per cent, compared with 26.2 per cent for the second-best entry.

The result became a defining moment in the emergence of the modern deep-learning era. It demonstrated that neural networks could become dramatically more capable when three things came together: large datasets, substantial computing power and improved training techniques.

Similar approaches soon produced major advances in image recognition, speech recognition, translation and other fields. Language, however, presented an additional challenge: understanding a sentence requires a system to recognise relationships between words and ideas that may be separated by considerable amounts of text.

6.7 2017: the Transformer changes natural-language AI

One of the most important developments behind modern generative AI occurred in 2017 when researchers at Google published a paper titled "Attention Is All You Need." The paper introduced the Transformer, a new neural-network architecture built around a mechanism known as attention.

When text is processed by a Transformer, it is divided into pieces known as tokens. Through a process called self-attention, the model can assess which other tokens in the surrounding text are most relevant to interpreting each token.

Consider:

The contract, which had been amended several times during the negotiations, was terminated because the purchaser failed to pay the deposit.

To understand this sentence, the model needs to recognise relationships between words and concepts that are separated by other text – for example, that the failure to pay relates to the purchaser and provides the reason for the termination of the contract. Attention provides a mathematical mechanism for identifying and giving greater importance to these relationships.

The Transformer had another major advantage. Earlier recurrent neural networks generally processed language sequentially, whereas Transformers could perform much more of their training in parallel. This made far more effective use of modern computing hardware and helped make it practical to train increasingly large neural networks on enormous quantities of text.

The Transformer architecture subsequently became the foundation for most modern large language models.

6.8 2018–2020: large language models begin to show general capability

In 2018, OpenAI published work on generative pre-training, showing that a Transformer model could first be trained on large quantities of general text and then adapted to perform a range of different language tasks.

In 2020, GPT-3 dramatically increased the scale of this approach. GPT-3 contained approximately 175 billion parameters – learned numerical values within the neural network that encode patterns and relationships acquired during training.

More importantly, GPT-3 demonstrated that a sufficiently large language model could perform many different tasks simply from instructions and examples provided in ordinary language, often without requiring a separately trained model for each task.

This represented an important change. Instead of building or training separate systems for:

- summarisation;
- translation;
- classification;
- drafting; and
- question answering,

a single general-purpose language model could perform all of these tasks by being told, in natural language, what the user wanted it to do.

The task was increasingly defined by the prompt, rather than by building a different AI system for each problem.

6.9 2022: models learn to follow human instructions

A raw language model is not naturally designed to behave like a useful assistant. Its initial pre-training teaches it to recognise and generate patterns in language, predominantly by learning to predict what text is likely to come next. This gives the model broad language capability and allows it to learn many relationships contained within its training data, but it does not necessarily teach the model how to answer a user's question clearly, follow instructions reliably or behave in a way that people find useful.

Researchers therefore developed additional post-training techniques designed to shape this underlying capability into more useful behaviour.

One important technique was reinforcement learning from human feedback, or RLHF. The model would produce different possible responses to an instruction, and human reviewers would indicate which responses they considered better. For example, if asked to explain a legal concept in plain English, reviewers might prefer a response that was clear, concise and directly answered the question over one that was overly technical or confusing. Those human preferences were then used as training signals so that, over many examples, the model became more likely to produce responses similar to those people preferred.

Importantly, RLHF was not principally about teaching the model new facts. Much of the model's underlying knowledge and language capability had already been acquired during pre-training. RLHF helped teach the model how to use that capability when interacting with people — for example, how to follow an instruction, answer a question, explain something clearly or respond in an appropriate style.

OpenAI's 2022 InstructGPT work demonstrated the significance of this approach. Human evaluators preferred responses from a 1.3-billion-parameter InstructGPT model over responses from the original 175-billion-parameter GPT-3, despite the instruction-following model being more than 100 times smaller.

On 30 November 2022, OpenAI released ChatGPT, using a conversational model related to InstructGPT that could respond to follow-up questions and interact with users through an accessible dialogue interface.

This marked the point at which generative AI moved from being principally a research and specialist technology into mainstream public use and awareness.

6.10 2023–2026: from chatbot to professional work platform

Since the release of ChatGPT, development has accelerated rapidly.

Leading AI systems can now work with large documents and document collections, images, spreadsheets, computer code and other forms of information. They can also search external sources, retrieve information from connected business systems, use software tools and undertake multi-step tasks.

Legal-specific platforms can combine powerful general-purpose AI models with authoritative legal sources, firm knowledge, precedents and matter documents. This means that the AI system being used by a lawyer may no longer be simply a standalone language model.

It can instead be thought of as several layers working together:

the underlying AI model

plus

the firm's documents, precedents and knowledge

plus

authoritative legal and external information

plus

search and retrieval mechanisms

plus

tools and connections to other software systems

plus

security, permissions and governance controls

plus

workflow instructions and processes.

This distinction is important. The underlying language model provides the general capability to understand and generate language, but the systems surrounding it can provide the model with relevant and authoritative information, control what information it can access and determine what actions it is permitted to perform.

These additional layers can substantially improve the usefulness and reliability of AI in professional legal work and reduce some of the limitations of a standalone chatbot.

They do not, however, eliminate those limitations.

7. What do we mean when we say "AI"?

Before considering how modern AI systems work, it is useful to distinguish several terms that are often used interchangeably but describe different things.

7.1 Artificial intelligence

Artificial intelligence, or AI, is the broadest term.

It refers generally to computer systems designed to perform tasks that would ordinarily require aspects of human intelligence, such as recognising patterns, interpreting language, making predictions, solving problems or selecting between possible actions.

AI therefore includes many technologies that pre-date ChatGPT. Spam filters, image-recognition systems, recommendation engines, speech recognition and IBM's Deep Blue chess system can all properly be described as forms of artificial intelligence.

AI is therefore best understood as the umbrella term under which the other technologies discussed in this paper sit.

7.2 Machine learning

Machine learning is one approach to creating AI.

As discussed earlier, traditional software generally relies on rules explicitly written by programmers. Machine learning instead allows a system to learn patterns from data.

Neural networks are an important form of machine learning. During training, their internal numerical values are repeatedly adjusted until the network becomes better at recognising useful patterns and producing the desired result.

Historically, many machine-learning systems were developed to perform relatively narrow tasks, such as recognising an object in an image or determining whether an email was likely to be spam.

Modern AI builds on the same principle, but at a vastly greater scale.

7.3 Foundation models

A major development in recent years has been the emergence of foundation models.

Rather than being trained to perform only one narrowly defined task, a foundation model is trained on very large and diverse datasets so that it develops broad capabilities that can later be applied to many different tasks.

This represented an important change in the development of AI.

Instead of building one model to translate text, another to summarise a document and another to answer questions, a sufficiently capable foundation model can potentially perform all of those tasks depending on the instructions it is given.

7.4 Large language models

A large language model, or LLM, is a type of AI model trained principally to work with language. Models in the GPT family and the models underlying Claude are examples.

The word large refers to the scale of these systems: they are trained on very large quantities of information and contain enormous numbers of learned numerical parameters representing patterns and relationships acquired during training.

An LLM is therefore not synonymous with AI. It is one particular – although currently very important – form of artificial intelligence.

LLMs are capable of performing many language-based tasks, including:

- summarising documents;
- drafting correspondence;
- explaining concepts;
- comparing information;
- extracting facts;
- translating text;
- analysing documents; and
- answering questions.

This is particularly significant for the legal profession because so much legal work is fundamentally language-based.

7.5 Generative AI

Generative AI describes AI systems designed to create new content. That content might include:

- text;

- images;
- audio;
- video;
- computer code; or
- other digital material.

Large language models are therefore an important form of generative AI because they generate text and other language-based outputs. However, generative AI is broader than large language models. An image-generation system, for example, is also generative AI even though it may use a different underlying model architecture.

The distinction can be summarised simply:

- AI is the broad field.
- Generative AI is AI that creates new content.
- An LLM is a type of AI model particularly focused on understanding and generating language.

7.6 Multimodal AI

Modern AI models are increasingly **multimodal**.

A multimodal model can work with more than one type of information. Rather than being limited to text, the same model may be capable of analysing combinations of:

- text;
- images;
- audio;
- video; and
- documents containing different forms of information.

For example, a lawyer might upload a scanned contract containing both text and diagrams and ask the system questions about the document as a whole.

This is one reason the term large language model is becoming an imperfect description of some of the most capable modern models. Language remains central to how users interact with them, but their capabilities increasingly extend beyond language alone.

7.7 AI assistants and chatbots

It is also important to distinguish the underlying AI model from the product through which a user interacts with it.

A model such as GPT is the underlying AI engine.

A product such as ChatGPT is an application built around one or more models.

That application may add capabilities such as:

- conversation history;
- document uploads;
- web search;
- access to company information;

- security and permissions;
- connections to other systems; and
- tools that allow the AI to perform additional tasks.

The distinction is important because two products using similarly capable underlying models can behave very differently depending on the information, security controls, tools and workflows surrounding the model.

The same applies to legal AI platforms. A legal AI product may combine a general-purpose model with authoritative legal sources, firm precedents, matter documents and additional legal workflows.

The model provides much of the underlying language and reasoning capability, but the complete AI system determines what information the model can access and what it is able to do.

7.8 Retrieval and grounding

A modern AI system does not necessarily have to rely only on information learned during its original training. It can also be provided with relevant information at the time a user asks a question.

For example, rather than asking an AI model to answer a question about a contract from its general knowledge, the system can retrieve the actual contract and provide the relevant provisions to the model before it generates an answer.

This approach is commonly referred to as retrieval-augmented generation, or RAG.

In a legal environment, the information supplied to the model might include:

- matter documents;
- correspondence;
- precedents;
- firm knowledge;
- legislation;
- judgments; or
- material from an authoritative legal research service.

This process is often described more generally as grounding the AI in reliable source material.

Grounding is particularly important in professional legal work because it reduces the extent to which the model must rely solely on patterns learned during its original training.

It does not eliminate the possibility of error, but it can significantly improve the relevance and reliability of the output.

7.9 AI agents

The next development is the move from AI systems that principally answer questions to systems that can undertake tasks. An AI agent combines an AI model with tools and information that allow it to perform a sequence of actions in pursuit of a particular objective.

For example, a lawyer might ask a conventional AI assistant:

Summarise this matter.

If all of the relevant information has already been provided, the model can analyse that material and generate a summary. An AI agent could potentially go further. Given appropriate permissions and connections to the firm's systems, it might:

- identify the relevant matter;
- retrieve matter details from the practice-management system;
- retrieve relevant emails and documents;
- analyse those materials;
- identify outstanding tasks;
- prepare a matter summary;
- generate a draft document; and
- save that draft against the appropriate matter for lawyer review.

The underlying language model provides much of the reasoning and language capability. The tools surrounding the model allow it to retrieve information and take actions.

This distinction is important. A language model principally generates an output. An agent can potentially use the model repeatedly as part of a broader workflow.

7.10 Agentic AI

The term agentic AI is increasingly used to describe AI systems that can pursue an objective with a degree of autonomy rather than merely responding once to a prompt.

For example, a generative AI system might be asked:

Draft a letter advising the client that settlement has occurred.

An agentic AI system might instead be instructed:

Review today's completed settlements, identify the clients who need to be notified, retrieve the appropriate firm template and matter information, prepare the letters for lawyer approval and save each draft against the correct matter.

The second task requires the system to determine and execute a series of steps rather than merely generate a single response.

For law firms, this distinction is significant because the productivity potential of AI increasingly lies not only in producing text more quickly, but in helping automate complete workflows.

7.11 A useful mental model

The terminology can therefore be understood as a series of related layers:

- *Artificial intelligence* is the broad field.
- *Machine learning* is one approach to building AI by allowing systems to learn patterns from data.
- *Neural networks* are an important machine-learning technique.
- *Foundation models* are large, general-purpose models trained on broad datasets.
- *Large language models* are foundation models principally designed to work with language.
- *Generative AI* uses models to create new content.

- *AI assistants* package models into products people can interact with.
- *Retrieval and grounding* provide models with relevant external information.
- *AI agents* combine models with tools so they can undertake multi-step tasks.
- *Agentic AI* describes the broader move towards systems capable of pursuing goals and completing work with increasing levels of autonomy.

This distinction becomes important when considering both the opportunities and risks of modern AI.

The AI system used by a lawyer today may therefore be considerably more than a chatbot. It may combine a powerful general-purpose model with matter documents, legal information, firm knowledge, search capabilities, software tools and security controls.

At the centre of many of these systems, however, remains the large language model.

Understanding what that model is actually doing helps explain both the extraordinary capabilities demonstrated by Collins and the limitations highlighted by Weatherall.

Why can a system analyse a 2,205-page transcript and produce H1-standard law examination answers, yet occasionally invent a case that does not exist?

The answer begins with understanding how a large language model processes and generates language.

8. What is a large language model actually doing?

The large language model, or LLM, sits at the centre of many of the AI systems now being used by lawyers. Understanding, at a practical level, what an LLM actually does helps explain both its extraordinary capabilities and its limitations.

8.1 Step 1: text is converted into tokens

A language model does not initially see a document in the same way a person does.

Instead, text is broken into small units known as tokens. A token might be a whole word, part of a word, punctuation, a number or another commonly occurring piece of text. For example, a sentence containing 20 words might be represented by somewhat more than 20 tokens because some words may be divided into smaller components.

Each token is then converted into a numerical representation that the neural network can process.

The important point is that an LLM ultimately works with numbers representing language, rather than words in the way a human sees them.

8.2 Step 2: the model learns patterns in language

During training, an LLM is exposed to enormous quantities of text. One of its fundamental training tasks is surprisingly simple: predict what token is likely to come next.

For example:

The High Court of Australia is the highest court in ...

The model might initially assign different probabilities to a range of possible next tokens. If the expected answer is *Australia*, the model's internal numerical parameters are adjusted slightly so that it becomes more likely to make the correct prediction in similar circumstances in future.

This process is repeated across enormous quantities of language.

Through that training, the model learns much more than individual word sequences. It develops statistical representations of patterns involving:

- grammar and sentence structure;
- writing styles;
- facts and concepts;
- legal terminology;
- relationships between ideas;
- common forms of reasoning; and
- the structures of different kinds of documents.

No programmer manually writes all of these relationships into the model. They are learned through the repeated adjustment of the neural network's parameters during training. This is why a model trained broadly on language can later perform tasks that were never individually programmed into it.

8.3 Step 3: the Transformer considers context

Predicting language requires more than simply knowing which words commonly appear next to each other. The meaning of a word or sentence often depends on information elsewhere in the document.

This is where the Transformer architecture discussed earlier becomes important.

Through its attention mechanisms, the model can assess relationships between the tokens in the available context and determine which parts of that context are particularly relevant to what it is currently processing.

For example, if a lawyer provides a contract and asks:

What obligations survive termination?

the model can identify relationships between provisions dealing with concepts such as:

- termination;
- confidentiality;
- indemnities;
- accrued rights; and
- survival of obligations.

It can then use those relationships when preparing its response.

This capability is particularly powerful in legal work because law is heavily language-based. Contracts, judgments, affidavits, correspondence, legislation and submissions all contain structured relationships between words, concepts, facts and legal principles.

Large language models are exceptionally capable at identifying and working with those patterns.

8.4 Step 4: the model generates its answer

When an LLM produces an answer, it does not normally retrieve a completed response from a database. It generates the response progressively. At each point, the model calculates which token is most appropriate to come next having regard to information such as:

- the user's question or instructions;
- the preceding conversation;
- documents supplied to the model;
- information retrieved from other sources;
- its learned parameters; and
- any additional system instructions or tool results.

It generates one token and then repeats the process using the newly generated token as part of the context for deciding what should come next. This occurs again and again until the response is complete.

The apparent simplicity of this process can be misleading. Because the model has learned extraordinarily complex relationships from vast quantities of language, predicting what should come next can result in sophisticated explanations, summaries, arguments and analysis.

However, it also reveals an important limitation.

The model's underlying task is not inherently:

"Determine what is legally true and only state verified facts."

Its underlying task is closer to:

"Given everything available to me, generate the most appropriate continuation."

Most of the time, those objectives may point towards the same answer. But they do not always. A response can therefore be linguistically convincing, logically structured and entirely plausible while still containing information that is incorrect.

That distinction is central to understanding why large language models can hallucinate.

9. Why does AI hallucinate – and why can the same prompt produce different answers?

The technology industry increasingly uses the term confabulation, although hallucination remains the more familiar expression. The US National Institute of Standards and Technology describes confabulation as a situation in which a generative AI system produces confidently stated but erroneous or false content.

Importantly, hallucination is not simply a software defect that can necessarily be eliminated by fixing a bug. It arises in part from the way generative AI works.

As discussed in the previous section, a large language model is trained to recognise patterns in enormous quantities of information and then generate an appropriate continuation from the context available to it. Its underlying objective is not inherently to retrieve and state a verified fact.

This distinction explains both why hallucination can occur and why two apparently identical requests can sometimes produce different answers.

9.1 Plausibility and truth are different things

A language model is extraordinarily good at recognising what a legal citation looks like. It has encountered large numbers of examples containing structures such as:

Smith v Jones [2024] FCA 123

It therefore understands the linguistic and structural pattern of an Australian legal citation.

If asked to provide an authority for a proposition for which it does not have reliable information, the model may nevertheless generate something that fits that pattern. The result can look exactly like a genuine case citation but that does not make the case real.

This is one reason fabricated authorities became an early and highly visible example of AI failure in legal practice. The model may be capable of generating a plausible legal answer where the lawyer requires a verified legal answer. Those are different standards.

9.2 The model may know the pattern better than the underlying fact

Consider the hypothetical question:

What did Justice Smith say at paragraph 127 of Brown v White?

A language model may recognise:

- the structure of the case name;
- the name of the judge;
- the kinds of legal issues that commonly arise in similar cases;
- the style in which Australian judgments are written; and
- the type of answer suggested by the question.

If the model does not actually have reliable access to paragraph 127, however, it may still be able to generate an entirely convincing answer based on those patterns.

This illustrates an important distinction between information that the model can generate from patterns learned during training and information that has been supplied to it from an identifiable source.

If the actual judgment is available to the system, the model can analyse that material directly. If it is not, the model may instead be relying on learned patterns and whatever information about the case is represented in its training.

The resulting answers may look equally polished, even though the evidentiary basis for them is very different.

9.3 Training information can be incomplete, inconsistent or outdated

An LLM learns from information created by humans. That information inevitably contains:

- errors;
- disagreement;

- outdated information;
- commentary of varying quality;
- inconsistent terminology; and
- jurisdictional differences.

The model can therefore reproduce or combine those inconsistencies.

For Australian lawyers this can be particularly significant. A general-purpose model may have been exposed to substantially greater volumes of United States legal material than Australian material. Concepts such as privilege, discovery, employment law or constitutional law may have superficially similar terminology across jurisdictions while operating under materially different legal rules.

A model generating an answer from its general training may therefore produce something that sounds legally sophisticated but reflects the wrong jurisdiction, an outdated legal position or a combination of different sources.

This is one reason access to authoritative and current source material can materially change the reliability of a legal AI task.

9.4 The model responds to the framing of the question

Modern AI assistants are post-trained to follow instructions and provide useful responses. That is one of the reasons they are so effective.

It can also influence the way they respond to the assumptions contained in a user's question. Consider:

Can you confirm that this case supports my argument?

The question itself suggests both a desired conclusion and an assumption that the case does support the argument. A system designed to provide a helpful answer may be more inclined to work within that framing than a lawyer would want from an independent reviewer.

This does not mean the model will always agree with the user. Modern systems are trained to correct false assumptions in many circumstances. However, the wording and assumptions contained in a prompt can influence what the model focuses on and the direction in which its analysis develops.

This is particularly important in legal work, where the user may themselves have formed a preliminary view and unintentionally frame the question in a way that encourages confirmation of that view.

9.5 Why can the same prompt produce different answers?

A conventional computer program will usually produce exactly the same result when given exactly the same inputs. Generative AI does not necessarily behave that way.

When an LLM generates a response, there will often be several possible next tokens that would all make reasonable sense.

For example, after:

The Court considered the application and...

possible continuations might include:

- dismissed;
- allowed;

- determined;
- concluded; or
- held.

The model assigns probabilities to possible next tokens. The generation process may select from among those possibilities rather than always selecting exactly the same one. Once a different token is selected, even relatively early in a response, everything that follows is generated using that new text as part of the context.

A small difference at the beginning can therefore send the response down a different path.

This means that two identical prompts can sometimes produce:

- different wording;
- different examples;
- different emphasis;
- different reasoning paths; or
- in some circumstances, materially different conclusions.

This variability was demonstrated during the Federal Court symposium itself. Collins noted that an AI system produced a less satisfactory result during the live demonstration than it had produced when essentially the same task was tested earlier that day. He observed that the technology generates a new response each time rather than simply retrieving the earlier answer.

There is another complication.

What appears to the user to be the same prompt may not always be the same complete input to the AI system. The answer may also be affected by:

- earlier messages in the conversation;
- documents available to the system;
- information retrieved from search or connected systems;
- system instructions applied by the AI platform;
- tools used during the task; and
- changes to the underlying model or platform.

The user sees the prompt they typed. The AI system may be working with considerably more context. This is another reason an AI system should not be treated like a deterministic legal database.

9.6 Fluent writing can create an illusion of certainty

Perhaps the most dangerous feature of hallucination is not simply that an answer can be wrong. It is that the wrong answer can be extremely well written.

A junior lawyer who is uncertain might say:

"I have not been able to locate an authority directly on point, but there appears to be…"

An AI system may instead produce a polished paragraph containing a legal proposition, case name, quotation and pinpoint citation. The quality of the writing can make the information psychologically persuasive.

Lawyers therefore need to separate "quality of presentation" from "quality of evidence supporting the proposition". An articulate answer is not necessarily an accurate answer.

The practical question is therefore not simply whether AI can hallucinate or produce different answers. It is how a lawyer should use that knowledge when instructing the system and evaluating its work.

10. How should a lawyer use AI in practice?

Understanding how generative AI works should influence how a lawyer instructs and works with it. Good prompting is not principally about discovering clever phrases or secret commands. It is about clearly defining the task, giving the system the information it needs, reducing unnecessary assumptions and making its work capable of being checked.

Several practical techniques follow.

10.1 Define the task and its boundaries

Do not make the AI infer matters that can easily be specified. Where relevant, tell it:

- what task it is performing;
- the jurisdiction;
- the relevant date;
- the documents or information it should use;
- any sources it should not use;
- the particular question it should answer; and
- the form in which the answer should be returned.

For example:

Using only the attached Federal Court judgments, identify the authorities relevant to the interpretation of clause 7. Do not introduce authorities from your general knowledge.

The more clearly the task is defined, the fewer assumptions the model needs to make.

10.2 Ask for analysis rather than confirmation

The framing of a question can influence the answer.

Instead of asking:

Confirm that this clause allows us to terminate immediately.

ask:

Analyse whether this clause permits immediate termination. Identify arguments for and against that interpretation and cite the contractual wording supporting each conclusion.

The first question suggests the conclusion the user wants. The second asks the AI to examine the issue. For legal work, the objective should generally be to use AI as an analytical tool rather than an agreement machine.

10.3 Require the AI to show the basis for important conclusions

Where accuracy matters, ask the system to identify the material supporting its conclusions. For example:

For each conclusion, identify the document, page or paragraph on which you rely.

This does not guarantee that the answer is correct. It does, however, make checking the answer substantially easier. Instead of deciding whether a polished response merely sounds convincing, the lawyer can compare the proposition with the identified source.

10.4 Tell the AI what to do when the information is insufficient

A model should not be encouraged to fill gaps unnecessarily. A useful instruction is:

If the supplied material does not contain sufficient information to answer the question reliably, identify what information is missing rather than making an assumption.

This gives the model an acceptable alternative to generating a plausible answer where the available information does not support one. It can also make the AI more useful by identifying what additional documents, facts or instructions are required before the task can be completed.

10.5 Break complex tasks into stages

For significant or complex work, it may be preferable not to ask AI to move immediately from a large volume of material to a final conclusion. The work can instead be divided into stages. For example:

1. identify the relevant documents;
2. extract the relevant passages;
3. organise the competing evidence;
4. identify the factual or legal issues;
5. analyse those issues; and
6. prepare a draft conclusion.

This structure makes the AI's work easier to inspect and reduces the risk that an error made early in the process will become embedded in the final answer.

A useful way to think about the process is:

Find → Extract → Organise → Analyse → Draft → Verify

The lawyer can then decide where review and verification should occur. These techniques are practical consequences of understanding how generative AI operates.

A useful principle is:

Do not ask AI to guess when you can give it the information, and do not accept an assertion when you can require it to show the evidence.

11. The opportunity for boutique law firms

For boutique firms, AI should not be viewed solely as a technology issue. It has the potential to change the economics and competitive structure of legal practice.

Historically, scale has given large law firms important advantages. A large firm can deploy teams of graduates, junior solicitors and paralegals to review documents, conduct research, prepare chronologies, compare evidence and undertake other labour-intensive work. Boutique firms generally cannot deploy people at the same scale.

Technology has already reduced some of that advantage.

Cloud computing, for example, means that a small law firm can now access many of the same enterprise-grade technology platforms, security capabilities and computing infrastructure as a much larger firm without having to make the same investment in its own infrastructure and IT resources.

AI has the potential to extend that process into the work of the law firm itself.

A smaller firm can increasingly gain access to sophisticated legal knowledge, analyse very large volumes of information and automate or accelerate routine intellectual work that previously required substantial junior resources. Tasks such as document review, chronology preparation, comparison of evidence, first-pass research, summarisation and drafting can increasingly be undertaken with AI assistance.

This does not remove the value of lawyers or legal expertise. It changes where that value is applied.

A boutique firm may be able to use AI to perform more of the information-processing work that once favoured firms with large teams, while applying its lawyers' time to the areas where human expertise remains most valuable: judgement, strategy, negotiation, advocacy and the client relationship.

The potential result is a reduction in the traditional advantage of scale.

For boutique firms, that may mean:

- competing for larger and more complex matters;
- dealing economically with larger volumes of documents and evidence;
- reducing turnaround times;
- improving the economics of fixed-fee work;
- making smaller matters economically viable; and
- allowing experienced lawyers to spend a greater proportion of their time on high-value legal work.

The opportunity should therefore not necessarily be viewed as using AI to employ fewer lawyers. Weatherall noted that some firms were reporting a different effect: greater productivity allowed them to do more work and reach more clients rather than simply reduce the number of people required.

That point applies more broadly across the legal profession.

AI represents a potentially substantial productivity opportunity. Firms that can safely incorporate it into everyday legal work may be able to deliver work faster, analyse more information, provide services more economically and increase the amount of work that each lawyer can undertake.

As with earlier technological changes, the competitive advantage is unlikely to come simply from purchasing the technology. It will come from learning how to use it effectively and embedding it into the way legal work is performed.

The firms that adapt most successfully are therefore likely to be better positioned to compete and grow as the economics of legal practice change.

12. The real lesson from the Federal Court symposium

The juxtaposition of Collins' demonstrations and Weatherall's comments captures the central issue particularly well.

On one hand, AI capability is improving extraordinarily quickly.

An AI model that produced a 70 per cent result in a master's-level law examination late in 2025 was followed within months by frontier models that Collins regarded as producing unquestionably H1-standard work. At the same symposium, AI was demonstrated analysing thousands of pages of trial transcript in minutes, comparing witness evidence, identifying potential evidentiary objections, checking citations and improving legal drafting.

The productivity potential is difficult to ignore.

On the other hand, greater capability does not remove the need for professional understanding.

It increases it.

As AI becomes more capable, its outputs also become more persuasive. A sophisticated, well-written and apparently authoritative answer can make a subtle error considerably harder to recognise.

That is why Weatherall's emphasis on knowledge is so important. Lawyers do not need to become AI engineers, but they should understand enough about how the technology operates to appreciate why it can produce both remarkable results and convincing errors. That understanding should influence the tools they select, the information they provide to those tools, the way they instruct them and the level of verification they apply to the result.

This suggests that a practical level of AI literacy is becoming part of what competent modern legal practice requires. The technology is now too capable to dismiss. It is also too imperfect to use blindly.

The firms likely to gain the greatest advantage will be those that operate between those two extremes: adopting AI deliberately, understanding its limitations, connecting it to reliable legal and firm information, and using experienced lawyers to apply judgement to the result.

That approach is consistent with the position taken by the Federal Court itself.

The Court is not resisting AI. It recognises its potential to increase efficiency, reduce costs and improve access to justice. But responsibility remains where it has always rested – with the people practising before it.

The technology may change very quickly.

The obligation of the lawyer to understand the tools being used, exercise judgement and take responsibility for the result does not.