Issues Arising from the Use of AI Language Models (including ChatGPT) in Legal Practice

12 July 2023

NSW Bar Association – Issues arising from the use of ChatGPT and other AI Language Models in Legal Practice

Approved by Bar Council on 22 June 2023

This document is intended to assist NSW barristers to understand their duties under the *Legal Profession Uniform Conduct (Barristers) Rules 2015* (NSW) (**Barristers Rules**) and the importance of their role in the administration of justice in the context of the consideration of whether to use artificial intelligence (**AI**) language models, such as ChatGPT, in their legal practice. This document should not be read as encouraging or approving the use of such models. It is a guide in relation to professional obligations that must be considered before, and not breached in the event of, such use.

About this document

This document provides guidance for barristers in relation to the consideration of use of AI language models, including ChatGPT, in their practice.

The use of AI-based models is now widespread. New models and techniques are emerging that may engage different ethical and practical considerations. Barristers should always be mindful of their professional obligations when considering whether to use new technology as an aide to their practice.

What GPTs do

AI language models that operate as generative pre-trained transformers (**GPTs**) are a group of language models 'pre-trained' on a large data set to generate human-like text responses. One currently prominent GPT is 'ChatGPT', which operates as a chatbot that users can ask questions of, including requests to generate documents.

A feature of ChatGPT is that it is possible to use 'prompts' which direct the kind and format of the answers received in response to a request to ChatGPT. At a high level, the writing of prompts is similar to using programming languages. The 'art' and 'science' of using ChatGPT emerges through the sophisticated use of prompts.

There have been reports of the use of ChatGPT in the drafting of legal documents, analysing arguments, and conducting legal research. For example, it is possible to ask ChatGPT to 'write' a legal letter, and ChatGPT will use the information contained within the query to generate a response in letter format, using information sourced from the web-based data set from which it has learned. These reports include those of research producing cases that do not exist and the production of false information, otherwise known as 'artificial reality' or 'hallucination'.

Under the Uniform Law provisions, barristers are bound by professional conduct rules and ethical obligations, which include providing competent and diligent representation, maintaining independence and integrity, and maintaining the confidentiality of client information.

When considering whether to use ChatGPT or any other tool which co-creates content, barristers should ensure that they are complying with those rules and obligations. Issues can also arise in relation to the intellectual property in any work product generated from AI tools, just as they can arise from any work product generated from other sources.

Barristers Rules: Independence, competence and diligence

The consideration of use of ChatGPT warrants careful consideration of the Barristers Rules. The relevant foundational rules are those contained in Rules 11, 13, 23 and 24, as informed by the Principles in Sub-rules 4(c) and (e) which may easily be breached by the use of GPTs, namely:

Rule 4

(c) **barristers as specialist advocates** in the administration of justice, must act honestly, fairly, skilfully, bravely and with competence and diligence.

(e) **barristers should exercise their forensic judgments and give their advice independently** and for the proper administration of justice, notwithstanding any contrary desires of their clients.

Rules 11 and 13 provide for the nature of barristers' work.

Rule 23 provides:

> A barrister has an overriding duty to the court to

act with independence in the interests of the administration of justice.

Rule 24 provides:

> A barrister must not deceive or knowingly or recklessly mislead the court.

Rule 35 provides:

> A barrister must promote and protect fearlessly and by all proper and lawful means the client's best interests to the best of the barrister's skill and diligence, and do so without regard to his or her own interest or to any consequences to the barrister or to any other person.

Rule 42 provides:

> A barrister must not act as the mere mouthpiece of the client or of the instructing solicitor and must exercise the forensic judgments called for during the case independently, after the appropriate consideration of the client's and the instructing solicitor's wishes where practicable.

These rules reflect the expectations of the community of barristers as specialist advocates. They require barristers to apply their own skill and exercise their own independent judgment when performing work for a client.

Additionally, barristers owe a duty to exercise reasonable care and skill in providing professional legal services to their clients, requiring them to demonstrate the competence and skill reasonably expected of barristers practising their profession and taking proper care.

Guidance: Clients are entitled to expect that any work that is done on a brief is a barrister's own work and reflects the application of the barrister's own skill, knowledge, experience and judgment. Barristers are expected to use any technology responsibly. The use of AI tools is no exception. If AI tools are to be used the barrister must ensure that the use supplements, and does not substitute, the barrister's own legal skills and knowledge.

Barristers should only release work under their name to a client or a decision-making body in relation to which they: (i) have exercised independent forensic judgment; (ii) have advised independently; (iii) have reviewed and edited where necessary; (iv) believe is correct and reliable based on their training, experience, and research; and (v) can therefore independently explain and support.

A barrister's duty of independence requires them to ensure that the work they deliver is the product of their own professional judgment. A corollary of that requirement is that they must be in a position to defend the professional judgment embodied in that work. The high level of independence barristers should apply to their work is easily eroded by reliance on or use of GPTs.

Although the design intent underlying language models is that they provide accurate information, for the reasons addressed in the following section that will not always be the case. Further, because language models are continuing to evolve (and the pace of evolution is expected to increase over time), the responses barristers receive from one time to the next in relation to the same prompt might be substantially different.

GPTs are not reliable in providing accurate legal information. Barristers need to understand the strengths and weaknesses of the GPT tools they use. As a minimum first step, a barrister should always verify the accuracy, reliability and currency of AI-generated information to ensure it is consistent with their own legal knowledge and research before relying on it for any purpose. It will generally be prudent for a barrister to choose to keep a record of the prompts they have used (their search history), the choices they have made, and the results generated by the AI tool. A barrister should also ensure that any use of an AI tool is consistent with its terms of service. A barrister should record these terms and the explicit decision they have made to use a particular tool. This is because the terms of service may frequently change and the terms they accepted may be different.

Limitations of GPTs (including ChatGPT)

GPT-based AI systems generate text based on patterns 'learned' from varieties of information on which the systems were trained, largely comprising undisclosed criteria and sources of varying quality. This presents several known limitations on the reliance of machine-generated answers to research queries, including:

(a) *Overestimation and incompleteness*: The body of information on which GPTs are trained is finite. If the data set on which it was trained does not

contain sufficient resources on the queried area of law in a particular jurisdiction, the GPT may produce a seemingly robust answer which is an incomplete or jurisdictionally inaccurate understanding of the law.

(b) *Currency*: Current GPTs have only been trained on data up to a certain date (for ChatGPT v3.5, September 2021) and there is a risk that the answers generated may not contain the current state of legislation or case law.

(c) *'Hallucination'*: GPTs are vulnerable to fabricating facts or sources when they do not have access to sufficient data to produce an answer. ChatGPT, for example, will not indicate where this occurs and the fabricated answer is given with the same level of confidence as one that is factually correct. ChatGPT's hallucination rate has been estimated to be as high as between 15% and 20%, and has been known to affect citation of sources when ChatGPT is asked to produce them.

(d) *Bias*: GPTs are prone to reproduce any bias within the data set on which they are based. Because older and more frequently replicated information within the data set may be favoured by the algorithm deployed by the GPT, there is a risk that outdated interpretations of legal principles may be produced in any answer without qualification. It also may favour information from more common or popular sources, such as media reports or popular websites, over ones that carry legal authority.

(e) *Unexplainability and lack of verifiability*: Added to the foregoing considerations, and a feature which amplifies all risks, is that the tools are proprietary and any hard constraints on their ability to answer have not been disclosed. As models change (or 'evolve'), the responses received from the use of particular prompts may change.

An example of the shortcomings of ChatGPT can be found in a recent case before the United States District Court for the Southern District of New York, in which an attorney prepared and filed a submission based on research using ChatGPT, which contained case references that were entirely false. This was revealed when the opponent and the judge attempted to access the cases cited and found they did not exist. In imposing sanctions on the practitioners involved, the presiding judge observed (*Mata v. Avianca, Inc.*, No. 1:2022cv01461 - Document 54 (S.D.N.Y. 2023)):

> Many harms flow from the submission of fake opinions. The opposing party wastes time and money in exposing the deception. The Court's time is taken from other important endeavors. The client may be deprived of arguments based on authentic judicial precedents. There is potential harm to the reputation of judges and courts whose names are falsely invoked as authors of the bogus opinions and to the reputation of a party attributed with fictional conduct. It promotes cynicism about the legal profession and the American judicial system. And a future litigant may be tempted to defy a judicial ruling by disingenuously claiming doubt about its authenticity.

Guidance: Each of the above limitations requires judgement as to whether the use of AI tools is appropriate to the task the barrister intends to undertake. The guidance in relation to independence, competence and diligence applies.

Where AI tools are used the answers generated should be carefully checked and interrogated before the barrister relies upon them in any way to assist in producing work for clients or courts. If that is not possible, then the barrister should avoid any reliance on those answers as a matter of prudence, professionalism and in the proper discharge of their responsibilities under the Barristers Rules.

Maintain transparency and accountability in relation to use, including transparency as to any impacts of AI use on costs charged to a client

Guidance: A barrister should be transparent with clients about their use of AI tools and their use to assist in legal representation. This should include disclosing to the client the nature of the AI tool the barrister proposes using and acknowledging the known limitations of the use of AI in legal practice. A barrister shall always remain entirely accountable for any legal work based on AI-generated information.

To the extent that AI is able to be utilised as a labour-saving tool to reduce the amount of time required for a particular task, (if using a time-based retainer model)

the amount of time charged to the client must, obviously, reflect the amount of time *actually spent* on the task.

Barristers Rules: Confidentiality and legal professional privilege

It is to be expected that developers of AI language models will record in their systems, and make use of, information entered into their models by users. The terms of use for ChatGPT grant its developer rights over inputs into the chatbot. ChatGPT saves a user's chat history and uses the conversations to train its models further unless the user manually disables the function. Where material entered into ChatGPT is incorporated into the data set it may form part of responses generated by other users' queries.

Barristers are therefore advised not to include any sensitive or confidential information when conversing with the chatbot. Several private corporations are banning the use of AI tools in the workplace following incidents involving leaks of confidential information entered into ChatGPT.

As with electronic records more generally, the inputs entered into an AI tool and the outputs generated as a result may be obtained through compulsory processes such as subpoenas or notices from regulatory authorities.

Guidance: When using an AI language model, barristers should be aware of pitfalls created by AI's terms of use and ensure that they do not disclose or use, in any way, sensitive or confidential information obtained by the barrister in the course of practice. Further, barristers ought to bear in mind that they may only disclose privileged communications if clearly instructed to do so by their client. A waiver of the client's privilege including where the AI's terms of use result in any such waiver, may mean that the barrister is in breach of the Barristers Rules.

Relevantly, **Rule 114** of the Barristers Rules provides that:

> A barrister must not disclose (except as compelled by law) or use in any way confidential information obtained by the barrister in the course of practice concerning any person to whom the barrister owes some duty or obligation to keep the information confidential unless or until:

> (a) the information is later obtained by the barrister from another person who is not bound by the confidentiality owed by the barrister to the first person and who does not give the information confidentially to the barrister, or

> (b) the person has consented to the barrister disclosing or using the information generally or on specific terms.

The requirement that barristers protect the confidentiality of such persons' data/information continues to apply in full when using AI tools.

A barrister should exclude at all times confidential or sensitive data in questions posed to GPTs.

Barristers Rules: Avoid bias and discrimination

As GPTs are trained on a vast source of internet-based data, AI-generated information produced by them can reproduce language contained within the data set that is biased, discriminatory and/or offensive. Such biases could include gendered and racist views.

When engaging with AI, barristers should be conscious of the guiding principle that barristers have a paramount duty to the administration of justice.

Relevantly, **Rule 8** of the Barristers Rules provides:

> A barrister must not engage in conduct which is:

> (a) dishonest or otherwise discreditable to a barrister;

> (b) prejudicial to the administration of justice; or

> (c) likely to diminish public confidence in the legal profession or the administration of justice or otherwise bring the legal profession into disrepute.

Rule 123 of the Barristers Rules also relevantly provides:

> (1) A barrister must not in the course of, or in connection with, legal practice or their profession, engage in conduct which constitutes:

> (a) discrimination,

> (b) sexual harassment, or

> (c) bullying.

> (2) For the purposes of subrule (1), conduct in connection with a barrister's profession includes,

but is not limited to:

(a) conduct at social functions connected with the bar or the legal profession, and

(b) interactions with a person with whom the barrister has, or has had, a professional relationship.

Guidance: In using material generated by an AI language model, barristers should note that the AI's response may include language which may put them in a position of contravening Rule 8 and/or Rule 123. Any output produced from AI tools should be carefully reviewed to ensure that such language is not repeated or endorsed in any way in work produced by, or communications emanating from, the barrister. If biased, discriminatory and/or offensive language is deployed in any communication with the court, a client, an instructing solicitor or the barrister's opponent, a barrister may thereby have contravened the above bar rules and exposed themselves to disciplinary action.

Intellectual property considerations

Barristers should ensure that any use of AI to generate written documents does not infringe the copyright of a third party in that party's work (such as a literary or artistic work) or other copyright subject matter (such as a video) by reproducing or communicating the whole, or a substantial part, of that work or subject matter. Given that the ultimate sources of any text generated by AI in response to a barrister's query or prompt are generally unknown, great care should be taken to avoid replicating that text in advice or other work done under a barrister's name.

As at the date of these Guidelines, litigation has been instituted against the operator of ChatGPT claiming that the data set on which the tool has been trained incorporates copyrighted material without consent of the copyright owner. Such material may be reproduced in ChatGPT's output. Barristers should be aware that if their use of an AI tool generates a document that reproduces a substantial part of copyrighted material without consent of the copyright owner, use of that document by the barrister may involve liability.

A defence to an allegation of copyright infringement is available under section 43 and section 104 of the *Copyright Act 1968* (Cth) provided that the relevant act

is for the purpose of a judicial proceeding; or the provision of professional advice given by a legal practitioner, but only if the use of the copyright work is a 'fair dealing'. This is likely to be informed by various factors including those set out in *Copyright Act* in relation to other exceptions.

Barristers should also bear in mind that if substantial parts of any work provided by the barrister, such as a written advice or submission, is generated directly by AI (albeit based on 'prompts') rather than drafted by the barrister from information provided by the AI tool, it is possible that the barrister will not own copyright in that literary work. In order for copyright to subsist in a literary work, there is a requirement that the 'author' of the work is human: see *Telstra Corp Ltd v Phone Directories Co Pty Ltd* [2010] FCAFC 149; 194 FCR 142.

Barristers should also consider appropriate academic guidelines governing plagiarism and proper attribution of sources but which may affect their (and the profession's) reputation when articles or papers are submitted for publication in journals and elsewhere.

Privacy considerations

As at the date of these Guidelines, litigation has been instituted against Google and the operator of ChatGPT claiming that the data sets on which their AI tools were trained incorporated personal data of individuals without their consent. Obligations governing use and disclosure of personal information vary widely by jurisdiction, and in Australia include the *Privacy Act* 1988 (Cth). Barristers should be mindful of these obligations if an AI tool incorporates such information in its output.

Further guidance

The critical point in relation to AI technology is that it cannot be used as a substitute for the proper exercise of a barrister's professional judgement in matters of law or in ignorance of their professional and ethical obligations. It is not a substitute for a barrister's own work.

In addition to the rules set out in these Guidelines, other rules of relevance include: **Rules 23, 24, 29, 58, 60, 61, and 64.**

These Guidelines may be considered by the Bar

Council in relation to any complaint received regarding a barrister's use of AI language tools for legal research, advice, and analysis.

Engaging in conduct against the Barristers Rules may give rise to a complaint and/or disciplinary action against a barrister.

Finally, the AI field is changing very quickly, and new developments may occur that are not referred to in these Guidelines. Barristers retain the responsibility of ensuring that any tool they use to aid them in practice is appropriate for such use and for ensuring at all times adherence to the Barristers Rules.

If any further guidance is required, please contact Ms Shar Doudman (Lawyer, Professional Conduct) at pcd@nswbar.asn.au or Ms Lucy Kelley (Policy Lawyer) on dppa@nswbar.asn.au or by phone (02) 9232 4055.