

# Professionals Beware: The Opportunities and Risks of Generative AI in Legal Practice

Associate Professor Evana Wright<sup>1</sup>

## Introduction

It has been a little over a year since ChatGPT controversially burst onto the scene in November 2022, promising users a human-like conversational experience that would change the way they live, learn and work. We have witnessed a rapid uptake of generative AI tools, including ChatGPT, Bing CoPilot, Dall-E and Midjourney, for a wide range of use cases: from the innocuous such as writing a poem or creating a recipe, to the more complex, such as drafting documents, reviewing contracts or providing legal advice. These tools have rapidly become more powerful and accurate. For example, the GPT-4 Technical Paper released by OpenAI reported that GPT-4 performed in the top 10 per cent of test takers in a simulated Bar exam compared to the performance of GPT-3.5 which fell into the bottom 10 per cent of test takers.<sup>2</sup> A proliferation of free and enterprise tools that have been released to the market, or incorporated into traditional products, perform a wide range of tasks, with some companies partnering with external service providers to create their own generative AI systems for specific internal use. Generative AI tools have been adopted in the legal profession, performing tasks such as legal research, contract review, and legal drafting.<sup>3</sup>

However, there are risks that come with the opportunities presented by these technologies, risks that legal practitioners and professionals must take into account when deciding whether or how to use them. Generative AI tools have been known to “hallucinate”, generating answers that are incorrect, inaccurate or completely fabricated and outputs may be biased or discriminate against a particular group. The development and use of generative AI tools may infringe copyright in works incorporated into training data sets. As the example of Mr Schwartz demonstrates below, not only is a practitioner’s reputation at risk, but misuse of generative AI may be in breach of a practitioner’s obligations under applicable professional conduct rules. It is critical that legal practitioners keep up with developments in technology and develop a working understanding of how generative AI tools before deploying them. This article will provide an overview of issues that legal practitioners and professionals should keep in mind when making the decision to adopt or use generative AI.

## What is generative AI?

There are many definitions of AI, with definitions shifting and changing as the technology develops. The OECD defines an “AI system” as:

*a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.*<sup>4</sup>

The Australian Government uses a somewhat similar definition and draws on the ISO Standard<sup>5</sup> to define AI as:

*an engineered system that generates predictive outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives or parameters without explicit programming. AI systems are designed to operate with varying levels of automation.*<sup>6</sup>

Generative AI is a type of AI system based on large language models (“LLMs”) or multimodal foundational models (“MFMs”) that use algorithms to make predictions based on inputs.<sup>7</sup> Commonly known examples of generative AI tools include ChatGPT, Bing CoPilot, Bard, Dall-E and Midjourney. These tools generate text or images in response to a text prompt. As set out in the Rapid Response Information Report on Generative AI:

*[w]hat all these models do is recognize patterns in data and produce sophisticated answers based on those patterns. The models are not intelligent or able to necessarily determine fact from fiction in their inputs or training data.*<sup>8</sup>

It is this last sentence that legal practitioners and professionals should pay attention to. Generative AI is not intelligent and it is this fact that contributes to some of the risks discussed in this article.

## **Hallucinations**

Many practitioners are now aware of the risk of generative AI tools “hallucinating”.<sup>9</sup> That is, generating output that may be inaccurate, fabricated or even non-sensical. This risk was made excruciatingly clear in the case of two lawyers representing the plaintiff in a personal injury case in the United States of America.<sup>10</sup> In this case, the lawyers relied on the content generated by ChatGPT to prepare a brief that was filed with the court. The brief contained references to six non-existent cases, and the error was brought to the court’s attention by opposing counsel who could not locate many of the authorities relied upon by the plaintiff. When questioned by the court and asked to produce further information on the cases, including case reports, the lawyers produced more content generated by ChatGPT, including extracts of case reports that had been essentially fabricated by the generative AI tool. When questioned by the Judge, it became apparent that the lawyer who prepared the brief, Mr Schwartz, had relied on ChatGPT as a legal research tool and did not conduct any independent checks as to the veracity or quality of the response generated by the tool. Responding to questioning from the court, Mr Schwartz stated that he did not believe that the tool would generate inaccurate or fabricated cases.<sup>11</sup> In finding that the lawyers acted in bad faith and issuing sanctions, the Judge observed that while there was nothing wrong with relying on databases and other tools to prepare submissions, it is the role of the lawyer to ensure that all filings are accurate.<sup>12</sup> The lawyers were fined US\$5000 and required to contact their client and the judges who were referenced in the falsified case reports and inform them of the sanctions imposed by the court.<sup>13</sup> Of course, in addition to the official sanctions of the court, the lawyers involved must also suffer the humiliation and loss of reputation that must flow from the significant, international media attention paid to this case.

When testing the capacity of generative AI to hallucinate, in mid-2023 the author tested ChatGPT3.5’s response to the prompt “Give me the names of some cases dealing with inventorship and patents in Australia”. ChatGPT generated the following text in response:<sup>14</sup>

*Here are a few notable cases in Australia that have addressed issues related to inventorship and patents:*

*1. Commissioner of Patents v. Microcell Ltd. (1959): This case dealt with the question of who can be considered an inventor and the requirements for validly claiming inventorship in a patent. It established that the true inventor must be a person who*

*contributes to the inventive concept and has a mental element in the formation of the invention.*

2. *Northern Territory v. Collinsville Sawmillers Pty Ltd. (1976)*: This case examined the concept of inventive step and whether an invention was obvious to a person skilled in the art. It clarified the importance of the inventive step requirement in determining the patentability of an invention.

3. *Pacific Enterprises (Aust) Pty Ltd. v. Rowlands (2002)*: This case focused on the ownership of an invention and the assignment of patent rights. It discussed the importance of properly documenting and establishing the chain of title for patents and highlighted the significance of clear contractual agreements in determining ownership.

4. *Meat and Livestock Australia Limited v. Cargill, Inc. (2006)*: This case involved the question of patentable subject matter and whether a method for selecting livestock based on genetic traits could be patented. It examined the criteria for patent eligibility and the boundaries of what can be considered a patentable invention.

5. *RPL Central Pty Ltd. v. Commissioner of Patents (2013)*: This case addressed the patentability of a computer-implemented invention related to the assessment of skills and recognition of prior learning. It explored the patentability of business methods and computer-implemented inventions and provided guidance on the patentability of software-based inventions.

*Please note that this is not an exhaustive list, and there may be other cases relevant to inventorship and patents in Australia. It is advisable to consult legal databases and seek guidance from a qualified intellectual property attorney for a comprehensive understanding of the current case law and its implications.*

The first case identified is *Commissioner of Patents v Microcell Ltd (1959)*. Those readers who are familiar with patent law in Australia will immediately recognise this case. *Commissioner of Patents v Microcell Ltd [1959] HCA 71* is a High Court of Australia decision dealing with issues relating to patenting new uses of known things. This case involved a patent application for a “self-propelled-rocket projector” made of reinforced plastic and dealt with the issue of whether a new use for a known material could be patentable. It does not deal with, as ChatGPT describes, issues relating to who may claim to be an inventor.

The second case, *Northern Territory v Collinsville Sawmillers Pty Ltd (1976)*, appears to be entirely fabricated. A search of reputable sources fails to identify such a case. However, for those versed in patent law, the case name does seem somewhat familiar. Could ChatGPT have generated a perversion of the actual case, *Northern Territory v Collins [2008] HCA 49*, a case dealing with contributory infringement? When queried on the full citation for *Northern Territory v. Collinsville Sawmillers Pty Ltd (1976)* ChatGPT responded as follows:<sup>15</sup>

*I apologize for the confusion in my previous response. Upon further investigation, I could not find a specific case titled “Northern Territory v. Collinsville Sawmillers Pty Ltd. (1976)” in the Australian legal database. It’s possible that the case may have been misattributed or the details provided were not accurate.*

*To ensure accuracy, it is recommended to consult legal databases or seek guidance from a qualified legal professional for the correct and complete citation of any specific case*

*you are referring to. They will be able to assist you in finding the accurate case and relevant information associated with it.*

Two things are of interest in this response. First, ChatGPT immediately retracted the reference and advised the user to refer to legal databases or a qualified lawyer. Second, the tool responded in the first person and used the language of an apology further reinforcing the idea that the user is in a dialogue or conversation with an assistant rather than engaging with an artificially intelligent system.

The author was also unable to locate a case by the name *Pacific Enterprises (Aust) Pty Ltd v Rowlands* (2002) although there are a number of cases that involved Pacific Enterprises. Again, when queried on the full citation, ChatGPT apologised for the confusion and stated that it could not find the case.<sup>16</sup> In regard to the fourth case on the list, *Meat and Livestock Australia Limited v Cargill, Inc* (2006), the citation and description of principles are almost correct, however, the year of the decision is incorrect. When questioned about the full citation for *Meat and Livestock Australia Limited v Cargill, Inc* (2006), ChatGPT issued an apology almost identical to the one set out above.<sup>17</sup> The author questioned ChatGPT further, asking “Was Meat and Livestock Australia Limited v Cargill Inc actually a case in 2018?” Again, ChatGPT apologised and insisted that “I couldn’t find a specific case titled ‘Meat and Livestock Australia Limited v Cargill Inc’ in the Australian legal database for the year 2018 or any other year.”<sup>18</sup> This denial is instructive. When questioned on the veracity of information, ChatGPT tends to “apologise” and advise users to consult legal databases or seek advice from a lawyer even when prompted towards the correct answer.

Finally, we turn to the fifth case on our list, *RPL Central Pty Ltd v Commissioner of Patents* (2013). Here, ChatGPT has generated a reference to an existing case and describes the legal principles correctly.<sup>19</sup> However, this achievement must be viewed in the context of the whole response and glaring inaccuracies in the other answers. Four out of five references contained errors or were complete fabrications and none of the cases provided actually dealt with the issue of inventorship in Australia.

It is important to acknowledge that ChatGPT has improved somewhat in the year since it was publicly released. Repeating the task outlined above in early 2024, questioning ChatGPT-4 on notable cases on inventorship and patents in Australia, the tool produced a list of five cases with complete citations (including year and court report).<sup>20</sup> Generally speaking, the descriptions of each of the cases were largely correct with the exception of the description for *SNF (Australia) Pty Ltd v Ciba Specialty Chemicals Water Treatments Limited* [2012] FCAFC 95 that referred to requirements for novelty and inventive step when the case actually dealt with the requirements for novelty and *innovative* step. Interestingly, none of the cases really dealt with the question of who can be an inventor, despite the request for “cases on inventorship”. This perhaps illustrates the importance of using highly specific prompts to elicit a useful response. Further questioning of ChatGPT-4 with specific questions as to who can be an inventor of a patent resulted in one case dealing with inventorship (along with three other hallucinations);<sup>21</sup> however, it took specific questioning as to the outcome in “Thaler and Commissioner of Patents” to get an answer dealing with AI inventorship.<sup>22</sup>

Practitioners are advised to exercise caution when using generative AI tools for legal research purposes. The responses generated by such tools should be verified using other methods such as searching reputable legal databases or drawing on professional experience and practice.

Platform owners such as OpenAI even draw the attention of users to this fact. OpenAI makes it clear that responses generated by ChatGPT can be incorrect with a warning under the chat interface on ChatGPT-4 stating “ChatGPT can make mistakes. Consider checking important information.” Furthermore, specific reference to the ability of ChatGPT to “hallucinate” is made in the service’s terms and conditions. The OpenAI “Terms of use” contain specific clauses dealing with “Accuracy” and advise users that “[g]iven the probabilistic nature of machine learning, use of our Services may, in some situations, result in Output that does not accurately reflect real people, places, or facts.”<sup>23</sup> Under the Terms of use, users agree not to rely on the content generated by the tool and to “evaluate Output for accuracy and appropriateness for your use case, including using human review as appropriate, before using or sharing Output from the Services.”<sup>24</sup>

The issue with hallucinations is not unique or limited to ChatGPT. Other generative AI tools acknowledge the risk that they may generate inaccurate or fabricated results. The “CoPilot AI Experiences Terms” include provisions similar to those in the OpenAI Terms of use, stating:

*The Online Services are for entertainment purposes; the Online Services are not error-free, may not work as expected and may generate incorrect information. You should not rely on the Online Services, and you should not use the Online Services for advice of any kind. Your use of the Online Services is at your own risk.*<sup>25</sup>

To avoid becoming a “teaching moment”, like Mr Schwartz in the case described above, users of generative AI are advised to review any output with a critical eye to validate the content and to ensure that they are not sharing inaccurate content or providing inaccurate or misleading advice to clients or the courts.

## **Copyright and generative AI**

AI has, and continues to, challenge the boundaries of intellectual property law. This is no different in the case of generative AI tools. The development and use of generative AI tools like ChatGPT and Dall-E raise challenges for copyright law. Does the use of images or text in training data sets infringe copyright? Can the outputs of generative AI tools reproduce a substantial part of a copyright work? Can a work created by, or in collaboration with, generative AI be protected by copyright?

### ***Copyright infringement***

Generative AI uses vast data sets to train its systems. These data sets may comprise millions of images or pages of text that have been scraped from the internet or otherwise obtained without the consent of the author or owner of the work and generative AI uses this data to make predictions and generate new outputs. Numerous court cases challenging the unauthorised use of content in training data sets are a testament to the outrage of creators and rights owners worldwide. For example, Stability AI and Midjourney are being sued by artists for copyright infringement claiming the use of their work in training sets amounts to unauthorised reproduction of copyright work and that outputs from generative AI are unauthorised derivative works.<sup>26</sup> Similarly, Getty Images is suing Stability AI in the US and UK for unauthorised copying of images, along with accompanying captions and metadata with Getty Images claiming that Stability AI has engaged in “brazen infringement of Getty Images’ intellectual property on a staggering scale”.<sup>27</sup> According to their complaint, more than 12 million Getty Images photographs have been copied by Stability AI to train their system.<sup>28</sup> Software

developers are suing GitHub, Microsoft and OpenAI, alleging unauthorised copying of software code in breach of relevant open source licences for use in developing the CoPilot tool.<sup>29</sup> Thomson Reuters is suing Ross Intelligence, alleging unauthorised copying of Westlaw database content for training data.<sup>30</sup>

The most recent, and perhaps the most high-profile, case dealing with copyright infringement and generative AI is that brought by *The New York Times* against Microsoft and OpenAI, alleging that Microsoft and OpenAI have copied millions of copyright protected articles in the course of training their model and operating their products.<sup>31</sup> *The New York Times* filed their complaint on 27 December 2023, following attempts to settle with Microsoft and OpenAI, in an attempt to “ensure it received fair value for the use of its content, facilitate the continuation of a healthy news ecosystem, and help develop GenAI technology in a responsible way that benefits society and supports a well-informed public.”<sup>32</sup> The trial pleading outlines several claims for copyright infringement including: unauthorised reproduction of works during GPT model training; embodiment of unauthorised reproductions and derivatives of works in GPT models; and unauthorised public display of works in GPT product outputs.<sup>33</sup>

Along with the expected claims relating to the incorporation of copyright works in training data and reproduction in derivative works, *The New York Times* claim also raises concerns regarding “memorization” by generative AI tools. This occurs when the tool generates content that is not only similar to, or derivative from, its training data; it is an almost complete reproduction of all or parts of the copyrighted work.<sup>34</sup> In documents filed with the court, *The New York Times* provides examples of outputs from ChatGPT and Bing Chat (now known as CoPilot) that allegedly copies *The New York Times* content, highlighting the similarity between passages as illustration.<sup>35</sup> In response, OpenAI has published an article titled “OpenAI and journalism” setting out how OpenAI believes the use of copyright content to be “fair use” under US copyright law and that the company is working to limit instances of memorization or regurgitation of content.<sup>36</sup> OpenAI also questions the prompts used by *The New York Times* to generate the examples included in the court documents and suggests that *The New York Times* manipulated prompts to generate examples of memorization or regurgitation that would not ordinarily be generated by ChatGPT in response to typical user inquiries.<sup>37</sup>

As identified by OpenAI in their response to *The New York Times* complaint, the application of exceptions to copyright protection, such as the doctrine of fair use in the United States and the text and data mining exceptions in the European Union, will be critical to the outcomes in these cases. Whether the use of copyright material is “fair use” is to be assessed against four fairness factors:

- (1) *the purpose and character of the use, including whether such use is of a commercial nature or is for non-profit educational purposes;*
- (2) *the nature of the copyrighted work;*
- (3) *the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and*
- (4) *the effect of the use upon the potential market for or value of the copyrighted work.*<sup>38</sup>

The application of these factors will depend on the specific facts of the case. Given the time it may take for these cases to progress through the courts, it may be some time before there is

clarification as to whether the activities of generative AI companies and compilers of training data sets could be considered “fair use”. It is possible that the decision in *Thomson Reuters v ROSS Intelligence* may come first, with a tentative trial date set for May 2024.<sup>39</sup> The text and data mining exceptions under European Union law may also provide protections for some of the activities of generative AI companies and those that compile training data sets. The companies that market generative AI tools may not be responsible for compiling the training data set. For example, the data set used by Stability AI was created by the Large-Scale Artificial Intelligence Open Network (“LAION”) based in Germany and it is likely that these activities would fall under the European Union text and data mining exceptions.<sup>40</sup> In the event that similar claims might be made in Australia, it is unlikely that the use of copyright works for these purposes will fall under the more limited fair dealing provisions under the *Copyright Act 1968 (Cth)*.<sup>41</sup>

Regardless of the outcome in these cases, it is significant to note that Microsoft is offering an indemnity for intellectual property infringement to commercial customers of Microsoft’s CoPilot and Azure OpenAI Service, provided the customers have implemented certain guardrails and content filters.<sup>42</sup> While this practice is common in the technology industry in relation to all kinds of technology including software, it is an important signal to customers that Microsoft is willing to take responsibility for intellectual property claims against paying customers and that their technology will operate in a way that minimises the risk of such claims.<sup>43</sup> The existence of such an indemnity may serve as an additional incentive to invest in subscriptions to enterprise versions of generative AI tools rather than rely on free or cheaper subscriptions. Users of generative AI tools are advised to confirm whether their specific tool provides such an indemnity. For example, both ChatGPT and CoPilot specifically disclaim responsibility for the infringement of intellectual property rights.<sup>44</sup> In fact, users may be required to indemnify the generative AI service from and against third-party claims relating to the use of the service or any output.<sup>45</sup>

In addition to issues of copyright infringement, consideration should also be given to the potential violation of an author’s or artist’s moral rights in the copyright work. In the Australian context, authors have the right of attribution of authorship, the right not to have authorship of a work falsely attributed and the right of integrity of authorship of a work.<sup>46</sup> It is unclear how the unauthorised copying of copyright works for inclusion in training data sets can be consistent with an author’s moral rights, in particular the right of attribution of authorship and the right of integrity of authorship. In what way does the training data set record or attribute authorship in relation to an included literary or artistic work? Could inclusion and use of literary or artistic work in a data set constitute derogatory treatment?

### ***Ownership of and rights in outputs***

In addition to the issue of copyright infringement, there are concerns regarding the rights that may attach to content generated by, or in collaboration with, generative AI. This is important as it may have a significant impact on the way the creator of such content may use, or prevent others from using, the work. A number of recent cases in the United States can provide guidance on the attribution of authorship with respect to AI generated works. The United States Copyright Office has provided guidance on authorship stating that, to be eligible for copyright protection, an author must be a human being.<sup>47</sup> According to the Copyright Office, the key question is:

*whether the ‘work’ is basically one of human authorship, with the computer [or other device] merely being an assisting instrument, or whether the traditional elements of authorship in the work ...were actually conceived and executed not by a man but by a machine.*<sup>48</sup>

Decisions relating to copyright registration of AI generated works in the US provide guidance on the level of collaboration between an author and a generative AI tool that is necessary for a work to be eligible for copyright protection. The first example is the comic book *Zarya and the Dawn*, registered with the United States Copyright Office.<sup>49</sup> While the comic book cover named both the author and Midjourney, the application for registration did not specifically address the fact that Midjourney was used to create some of the content.<sup>50</sup> The Copyright Office became aware of the use of generative AI following social media posts by the author discussing the work and reviewed the registration looking at who was responsible for creating specific elements of the work.<sup>51</sup> In this case, the author created the text and arranged elements of the work while images were generated by Midjourney with some modifications made to the images by the author.<sup>52</sup> The Copyright Office cancelled the registration in relation to content generated by Midjourney and a more limited registration was issued covering those elements of the comic book that were created solely by the artist.<sup>53</sup>

A subsequent case involving the Copyright Office and Dr Thaler, the well-known litigant in various patent litigation around the world seeking recognition of an AI system known as DABUS as an inventor,<sup>54</sup> provides further guidance on the level of collaboration that may be necessary for a work to be eligible for copyright protection.<sup>55</sup> Dr Thaler is the creator and owner of the “Creativity Machine”, and sought to register an image generated by the Creativity Machine, “A Recent Entrance to Paradise” with the US Copyright Office. The application was denied, and Dr Thaler challenged the decision arguing that there is nothing in US Copyright Law that requires an author to be a human.<sup>56</sup> Instead, Dr Thaler argued that it was appropriate to view the issue within the context of the doctrine of “work for hire” allowing Dr Thaler to own the copyright as the owner of the Creativity Machine in a way similar to the way employers can own the copyright in works created by an employee.<sup>57</sup> The US District Court for the District of Columbia held that a human author was required and that while advances in technology “will prompt challenging questions regarding how much human input is necessary to qualify the user of an AI system as an ‘author’ of a generated work”, this is not one of the cases.<sup>58</sup>

The position of the US Copyright Office was confirmed in the more recent decision of the US Copyright Review Board in the artistic work “Théâtre D’opéra Spatial” produced by Midjourney and modified by the artist Jason M Allen. The image was the result of more than 624 prompts from the artist to Midjourney and subsequent editing by the artist in Adobe Photoshop.<sup>59</sup> The Copyright Review Board affirmed the decision to refuse to register the artistic work stating that the work contains “more than a de minimis amount of content generated by artificial intelligence” as the Midjourney image “remains in substantial form in the final Work”<sup>60</sup> requiring that this work be disclaimed<sup>60</sup> before the Copyright Office can proceed to register those elements of the work that are the human author’s contribution.<sup>61</sup> How these kinds of issues would be resolved in Australia is yet to be seen. Drawing on existing case law, it would appear that a human author exerting ‘independent intellectual effort’ would be required in order for a work to be eligible for copyright protection under the *Copyright Act 1968 (Cth)*.<sup>62</sup>



In addition to issues surrounding authorship and whether a work may be eligible for copyright protection, users of generative AI should confirm the specific rights granted to them by the owner of the generative AI tool in question and the extent to which those terms grant them the right to use the resulting output. These terms vary and may limit the way that a user may use a particular work. For example, the current CoPilot AI Experiences Terms acknowledge that outputs generated by CoPilot may not be unique<sup>63</sup> and clearly state that “Microsoft does not claim ownership of Prompts, Creations, customizations, instructions, or any other content you provide, post, input, or submit to, or receive from the Online Services”.<sup>64</sup> In return, users grant Microsoft a broad licence to use “Prompts, Creations, customizations (including GPTs), and related content in connection with the operation of its businesses.”<sup>65</sup> Previous versions of the Bing Conversational Experiences and Image Creator Terms limited the right to use output generated by the service for “any legal personal, non-commercial purpose.”<sup>66</sup> The OpenAI Terms of Use also acknowledge that content created by OpenAI services may not be unique and grants to users ownership of any output, explicitly assigning to the user all “right, title, and interest, if any, in and to the Output”.<sup>67</sup> OpenAI retains the right to use both inputs and outputs to “provide, maintain, develop, and improve our Services” however, it does provide an option for users to opt-out of this, noting that if they do, “this may limit the ability of our Services to better address your specific use case.”<sup>68</sup> Users of generative AI tools are advised to check the terms of services to ensure that they have the rights (for example, an assignment or licence) to use any output for their specific, intended purpose.

## **Bias**

Generative AI tools like ChatGPT generate content by drawing on the LLM or MFMs underpinning the tool. The training data incorporated into this large language model informs the quality and accuracy of the output. Where the training data has gaps or reflects certain biases, the output may be flawed or biased. We shouldn’t be surprised by the capacity of generative AI to produce biased or discriminatory content. This is a known problem with AI generally. For example, in 2016, reporting by ProPublica uncovered significant bias in the COMPAS tool used in the criminal justice system in the United States to conduct risk assessments of defendants.<sup>69</sup> The tool was shown to be slightly more accurate than a coin toss in predicting future reoffending and there were “significant racial disparities” in outcomes.<sup>70</sup> Black defendants were more than twice as likely to be wrongly flagged as reoffenders than white defendants.<sup>71</sup> The Australian Human Rights Commission report on Human Rights and Technology, published in 2021, outlines the risk of “algorithmic bias” and recommends “greater guidance for government and non-governmental bodies in complying with anti-discrimination law in the context of AI-informed decision making.”<sup>72</sup> The issue of bias and discrimination is a key area of concern in numerous AI ethics frameworks including Australia’s AI Ethics Principles.<sup>73</sup>

Testing the capabilities of Dall-E in mid-2023, images generated in response to prompts regarding lawyers demonstrated some bias in representation. Images generated in response to the prompts “impressionist painting of lawyers” and “painting of lawyers digital art” showed little gender diversity with most of the images depicting men. A request asking for “lawyers practising in court” resulted in pictures showing greater diversity in terms of gender representation, whereas a prompt asking for pictures of “lawyers negotiating” resulted in four images depicting only men. Finally, the images generated in response to the prompt “lawyers stressed about new technology” depicted mostly women. To what extent do these images reflect

existing bias? Is the technology responding to historical data regarding the gender make-up of the profession? Or are there more impressionist paintings of men? Are women more likely to be stressed by new technology? These observations relate to a very limited data set; however, research looking at over 100 images generated by AI confirms the existence of bias across a wide range of characteristics, including sex, age, race, class and geographical location.<sup>74</sup>

Asking DALL-E 2 to generate images based on these prompts (in early 2024) resulted in a different set of images. There was still little gender diversity in the images generated in response to the prompts “impressionist painting of lawyers” and “painting of lawyers digital art” with the images depicting mostly men. However, when asking for images of “lawyers stressed about new technology” and “lawyers negotiating”, the generated images showed much greater gender diversity when compared to the images produced in 2023. Again, it is not possible to reach firm conclusions as to bias on the basis of such a limited data set. Given the nature of generative AI, the response to a prompt will likely be different every time. All users of generative AI should continue to be conscious of the potential bias that may be depicted or included in content produced by generative AI and assess the suitability of content for their intended purpose. Particular care should be taken where the output of the generative AI may have an impact on a person’s life such as where generative AI may be used to assist in legal decision making.

### **Privacy and confidential information**

As noted above, generative AI services may retain the right to keep and use any content you upload as prompts to the service. In addition to ensuring that they have the right to upload and share content from an intellectual property perspective, users should also ensure that any content shared with generative AI tools is consistent with privacy regulations and obligations of confidence. In particular, practitioners should ensure that they are not exposing personal or confidential information in breach of privacy legislation, terms of engagement with a client or professional obligations including legal professional privilege. This applies to the use of free and enterprise versions of generative AI, especially where firms may be managing conflicts.

The Italian privacy regulator blocked access to ChatGPT in early 2023, raising concerns about OpenAI’s data collection practices and the extent to which these practices were consistent with requirements under the European *General Data Protection Regulation* (“GDPR”).<sup>75</sup> In response, OpenAI implemented changes in practice, including processes allowing users to request that their information be removed from the service.<sup>76</sup> While privacy legislation in Australia is weaker in comparison to the GDPR, users should still exercise caution and ensure they comply with the *Privacy Act* 1988 (Cth) and the Australian Privacy Principles, and, where necessary, make sure that any terms of service or privacy policies allow for use of personal information including disclosure of information to generative AI tools.

Generative AI tools may also generate or disclose information about a person and this information may not always be accurate. This is acknowledged by OpenAI in their Privacy Policy, stating that where someone wishes to correct an inaccuracy they may submit a correction request however “[g]iven the technical complexity of how our models work, we may not be able to correct the inaccuracy in every instance. In that case, you may request that we remove your personal information from ChatGPT’s output”.<sup>77</sup>

In addition to compliance with relevant privacy legislation, users should ensure that confidential information is not shared with generative AI services. This may include privileged or client information, customer lists, software code, or other trade secrets. In 2023, a number of large companies, such as Apple and Samsung, banned or limited the use of generative AI tools to ensure the protection of confidential information, including trade secrets.<sup>78</sup> The risk of disclosing confidential information may be mitigated by using versions of generative AI tools that do not collect or share data, including for training purposes. This may include bespoke tools developed for a specific company or tools supplied under enterprise licences. Practitioners should take a cautious approach to sharing information until they have confirmed how the data shared with a generative AI service is collected and used.

### **Professional responsibility**

Legal practitioners have duties to the courts, to clients and to the public. These are set out in various conduct rules.<sup>79</sup> Other professionals, such as IP attorneys, have similar obligations under relevant codes of conduct.<sup>80</sup> Some professional bodies have issued guidance to members on the use of generative AI in practice.<sup>81</sup>

In addition to duties to the court and the administration of justice,<sup>82</sup> legal practitioners have duties to clients including acting in a client's best interest, as well as obligations to be honest and courteous; to act competently, diligently and promptly; and to avoid compromise to integrity and professional independence.<sup>83</sup> These duties should be kept in mind when using generative AI in legal practice, in particular, the duties to deliver services competently, diligently and promptly. There are two ways to look at these obligations with respect to the use of technology more broadly and generative AI specifically. First, keeping in mind the case of Mr Schwartz discussed above, would it be competent or diligent to rely on material generated by tools such as ChatGPT without conducting due diligence to confirm its accuracy? Second, would it be unprofessional to not use technology, such as generative AI or other tools, where it may result in a more cost-effective and timely provision of services? The Victorian Supreme Court has already ordered the use of technology-assisted review in a case involving, at a minimum, approximately 1.4 million documents on the grounds that "traditional manual discovery ... is not likely to be either cost effective or proportionate."<sup>84</sup> Practitioners are advised to consider how these tools may be used in the course of their practice, including by other legal practitioners and employees engaged in the provision of legal services for which they are responsible.

### **Conclusion**

This article outlines some of the risks involved in using generative AI tools, both generally and specifically in legal practice. It is now more important than ever for lawyers to take opportunities to understand new technologies and how to use them. The rapid adoption of generative AI raises the question of the kind of skills legal practitioners and other professionals need to effectively use these tools in compliance with their obligations under professional conduct rules. Understanding the risks associated with generative AI is the first step. However, legal practitioners and professionals must continue to engage in professional development to keep up with advances in technology and develop new skills in areas such as prompt engineering. Generative AI represents an opportunity for the profession, allowing practitioners and professionals to focus on higher-value tasks that leverage the higher-order skills that legal training provides such as critical thinking, advocacy and negotiation skills. However, it must

be deployed effectively and responsibly to uphold obligations to the courts, the administration of justice, and to clients.

---

<sup>1</sup> Associate Professor, Faculty of Law, University of Technology Sydney. This article is adapted from the presentation “Using ChatGPT and similar tools” given at an IPSANZ Professional Development Event on 20 July 2023. The author is grateful to Dr Genevieve Wilkinson and Fiona Rotstein for thoughtful feedback on earlier drafts.

<sup>2</sup> Open AI, *GPT-4 Technical Report*, (Report, 19 December 2023) 1 <<https://arxiv.org/pdf/2303.08774.pdf>>.

<sup>3</sup> See, e.g. Microsoft, *Generative AI for lawyers* (October 2023)

<<https://www.microsoft.com/cms/api/am/binary/RW1dRAC>>; ‘LexisNexis Launches Lexis+AI, a Generative AI Solution with Linked Hallucination-Free Legal Citations’ *LexisNexis* (Web Page, 25 October 2023)

<<https://www.lexisnexis.com/community/pressroom/b/news/posts/lexisnexis-launches-lexis-ai-a-generative-ai-solution-with-hallucination-free-linked-legal-citations>>; Jane Wang and Sharyn Ch’ang, ‘Are you using AI yet? How lawyers can use it to their advantage’ *PwC* (Web Page) <<https://www.pwc.com/gx/en/about/pwc-asia-pacific/are-you-using-ai-yet-how-lawyers-can-use-it-to-their-advantage.html>>.

<sup>4</sup> OECD, *Recommendation of the Council on Artificial Intelligence* (OECD/LEGAL/0449, 8 November 2023) I.

<sup>5</sup> ISO Standard: Information technology – Artificial intelligence – Artificial intelligence concepts and terminology (ISO/IEC 22989:2022) 3.1.4 ‘artificial intelligence system’.

<sup>6</sup> Australian Government, *Safe and responsible AI in Australia: Discussion Paper* (June 2023) 5.

<sup>7</sup> Genevieve Bell, Jean Burgess, Julian Thomas and Shazia Sadiq, *Rapid Response Information Report: Generative AI – language models (LLMs) and multimodal foundation models (MFMs)* (Australian Council of Learned Academies, 24 March 2023) 2.

<sup>8</sup> Genevieve Bell, Jean Burgess, Julian Thomas and Shazia Sadiq, *Rapid Response Information Report: Generative AI – language models (LLMs) and multimodal foundation models (MFMs)* (Australian Council of Learned Academies, 24 March 2023) 2.

<sup>9</sup> The Cambridge Dictionary Word of the Year 2023 was ‘hallucinate’. See Cambridge Dictionary, *The Cambridge Dictionary Word of the Year 2023 is ...* (Web Page)

<<https://dictionary.cambridge.org/editorial/woty>>. The Dictionary defines “hallucinate” with respect to artificial information as producing “false information”. See *The Cambridge Dictionary* (online at 8 February 2024) “hallucinate” (definition 2).

<sup>10</sup> *Roberto Mata v Avianca, Inc.* (US S.D.N.Y, 1:22-cv-01461-PKC, 22 June 2023) Opinion and Order on Sanctions (“*Mata v Avianca*”).

<sup>11</sup> *Mata v Avianca*, [12].

<sup>12</sup> *Mata v Avianca*. The sanctions related not only to the reliance on fabricated case law generated by ChatGPT but also other false statements made by the attorneys during the course of proceedings, including relating to the availability of counsel.

<sup>13</sup> *Mata v Avianca* (Judge Castel).

<sup>14</sup> Output from ChatGPT, OpenAI to Evana Wright, 2023. Section 3 ‘Content’ of the OpenAI terms and conditions of use provides that users “receive all right title and interest in and to the Output’ and confirm that users can “use Content for any purpose, including commercial purposes such as sale and publication.”

<sup>15</sup> Output from ChatGPT, OpenAI to Evana Wright, 2023.

<sup>16</sup> Output from ChatGPT, OpenAI to Evana Wright, 2023.

<sup>17</sup> Output from ChatGPT, OpenAI to Evana Wright, 2023.

<sup>18</sup> Output from ChatGPT, OpenAI to Evana Wright, 2023.

<sup>19</sup> Output from ChatGPT, OpenAI to Evana Wright, 2023.

<sup>20</sup> Output from ChatGPT, OpenAI to Evana Wright, 2024. The cases generated were *D’Arcy v Myriad Genetics Inc* [2015] HCA 35; *Commissioner of Patents v RPL Central Pty Ltd* [2015] FCAFC 177; *AstraZeneca AB v Apotex Pty Ltd* [2015] HCA 30; *SNF (Australia) Pty Ltd v Ciba Specialty Chemicals Water Treatments Limited* [2012] FCAFC 95; *Research Affiliates LLC v Commissioner of Patents* [2014] FCAFC 150.

<sup>21</sup> *University of Western Australia v Gray* [2009] FCAFC 116. The other cases cited (hallucinations) were *Collins v Northern Territory of Australia* [2008] FCA 1366; *Griffith Hack v Nine Network Australia Pty Ltd* [2009] FCA 1066, and *Emmett’s Pty Ltd v Sharpe’s (North West Bend) Pty Ltd* [2019] FCA 1526. A search of Austlii and the Federal Court of Australia Judgments Digital Library was unable to locate any of the last three cases.

<sup>22</sup> ChatGPT was unable to respond to queries regarding the Full Federal Court of Australia decision, *Commissioner of Patents v Thaler* [2022] FCAFC 62 claiming that it was after its last update of April 2023.

- 
- <sup>23</sup> *OpenAI Terms of use* (Web Page, updated 14 November 2023, effective 31 January 2024) ‘Accuracy’ <<https://openai.com/policies/terms-of-use>>. (“*OpenAI Terms of use*”).
- <sup>24</sup> *OpenAI Terms of use*.
- <sup>25</sup> *CoPilot AI Experiences Terms* (Web Page, 15 January 2024), s 6 <[www.bing.com/new/termsfuse](http://www.bing.com/new/termsfuse)> (“*CoPilot AI Experiences Terms*”).
- <sup>26</sup> *Andersen et al v Stability AI Ltd et al*, (US N.D.Cal, 3:23-cv-00201, 13 January 2023) Complaint.
- <sup>27</sup> *Getty Images (US) Inc v Stability AI Inc*, (US D.Del, 1:23-cv-00135, 3 February 2023) Complaint [1] (“*Getty Images (US) v Stability AI*”) See also *Getty Images v Stability AI* (UK High Court filing 12 May 2023) seeking an injunction preventing the sale of the product in Britain.
- <sup>28</sup> *Getty Images (US) v Stability AI*, [1].
- <sup>29</sup> *J. Doe 1 and J. Doe 2 v GitHub Inc, Microsoft Corp., OpenAI and Others*, (US N.D.Cal, 3:22-cv-06823, 3 November 2022).
- <sup>30</sup> *Thomson Reuters Enterprise Centre GmbH v ROSS Intelligence Inc*, (US D.Del, 1:20-cv-00613-UNA, 6 May 2020).
- <sup>31</sup> *The New York Times Company v Microsoft Corporation, OpenAI Inc and Others* 2023 WL 8933610 (US S.D.N.Y, 1:23-cv-11195, 27 December 2023) Jury Trial Demand (“*The New York Times v Microsoft*”).
- <sup>32</sup> *The New York Times v Microsoft*, [7].
- <sup>33</sup> *The New York Times v Microsoft*.
- <sup>34</sup> *The New York Times v Microsoft*, [98].
- <sup>35</sup> See e.g. *The New York Times v Microsoft*, 99-[100] and Exhibit J.
- <sup>36</sup> Open AI, *OpenAI and journalism* (Blog, 8 January 2024) <<https://openai.com/blog/openai-and-journalism>>.
- <sup>37</sup> Open AI, *OpenAI and journalism* (Blog, 8 January 2024) <<https://openai.com/blog/openai-and-journalism>>.
- <sup>38</sup> *Copyrights*, 17 U.S.C. § 107 (2018).
- <sup>39</sup> Isaiah Poritz, ‘Thomson Reuters Will Head to Trial in AI Model Copyright Battle’ *Bloomberg Law* (online, 27 September 2023) <<https://news.bloomberglaw.com/ip-law/thomson-reuters-will-head-to-trial-in-ai-model-copyright-battle>>.
- <sup>40</sup> Pamela Samuelson, ‘Generative AI meets copyright’ (2023) 381 (6654) *Science* 158, 159.
- <sup>41</sup> *Copyright Act 1968* (Cth) ss 40-42 covering fair dealing for the purposes of: research or study; criticism or review; parody or satire; and reporting news. See Rita Matulionyte, ‘Generative AI and Copyright: Exception, Compensation or Both?’ (2023) 134 *Intellectual Property Forum* 33, 34.
- <sup>42</sup> Brad Smith, ‘Microsoft announces new Copilot Copyright Commitment for customers’ *Microsoft* (Blog Post, 7 September 2023) <<https://blogs.microsoft.com/on-the-issues/2023/09/07/copilot-copyright-commitment-ai-legal-concerns/>>; Jessica Hawk, ‘Microsoft Azure AI, data, and application innovations help turn your AI ambitions into reality’ *Microsoft* (Blog Post, 15 November 2023) <<https://azure.microsoft.com/en-us/blog/microsoft-azure-ai-data-and-application-innovations-help-turn-your-ai-ambitions-into-reality/>>.
- <sup>43</sup> Brad Smith, ‘Microsoft announces new Copilot Copyright Commitment for customers’ *Microsoft* (Blog Post, 7 September 2023) <<https://blogs.microsoft.com/on-the-issues/2023/09/07/copilot-copyright-commitment-ai-legal-concerns/>>.
- <sup>44</sup> *OpenAI Terms of use*, ‘Disclaimer of Warranties’, *CoPilot AI Experiences Terms* s 6.
- <sup>45</sup> *OpenAI Terms of use*, ‘Indemnity’, *CoPilot AI Experiences Terms* s 6.
- <sup>46</sup> *Copyright Act 1968* (Cth) Part IX – Moral Rights of Performers and of Authors of Literary, Dramatic, Musical or Artistic Works and Cinematographic Films.
- <sup>47</sup> US Copyright Office, Compendium of U.S. Copyright Office Practices § 101 (3d ed. 2021) §306.
- <sup>48</sup> US Copyright Office, Compendium of U.S. Copyright Office Practices § 101 (3d ed. 2021) §313.2.
- <sup>49</sup> ‘Zarya of the Dawn’ United States Copyright Office Registration # Vau001480196.
- <sup>50</sup> Letter from Robert J. Kasunic, United States Copyright Office to Van Lindberg, 21 February 2023, 2.
- <sup>51</sup> Letter from Robert J. Kasunic, United States Copyright Office to Van Lindberg, 21 February 2023, 2.
- <sup>52</sup> Letter from Robert J. Kasunic, United States Copyright Office to Van Lindberg, 21 February 2023, 8-11.
- <sup>53</sup> Letter from Robert J. Kasunic, United States Copyright Office to Van Lindberg, 21 February 2023, 12.
- <sup>54</sup> See e.g. *Commissioner of Patents v Thaler* [2022] FCAFC 62; *Stephen Thaler v Comptroller of Patents Trade Marks and Designs* [2021] EWCA Civ 1374.
- <sup>55</sup> *Stephen Thaler v Shira Perlmutter, Register of Copyrights and Director of the United States Copyright Office, et al*, (US D.D.C, 1:22-cv01564-BAH, 18 August 2023) (“*Thaler v Perlmutter*”).
- <sup>56</sup> Letter from Shira Perlmutter, Suzanne Wilson and Kimberley Isbell to Ryan Abbott, Esq, 14 February 2022, 2.
- <sup>57</sup> *Thaler v Perlmutter*, 3.
- <sup>58</sup> *Thaler v Perlmutter*, 13.

- 
- <sup>59</sup> Letter from Suzanne V. Wilson, Maria Strong and Jordana Rubel, US Copyright Review Board to Tamara Pester, Esq, 5 September 2023, 2 <<https://www.copyright.gov/rulings-filings/review-board/docs/Theatre-Dopera-Spatial.pdf>>.
- <sup>60</sup> Letter from Suzanne V. Wilson, Maria Strong and Jordana Rubel, US Copyright Review Board to Tamara Pester, Esq, 5 September 2023, 5 <<https://www.copyright.gov/rulings-filings/review-board/docs/Theatre-Dopera-Spatial.pdf>>.
- <sup>61</sup> Letter from Suzanne V. Wilson, Maria Strong and Jordana Rubel, US Copyright Review Board to Tamara Pester, Esq, 5 September 2023, 5 <<https://www.copyright.gov/rulings-filings/review-board/docs/Theatre-Dopera-Spatial.pdf>>.
- <sup>62</sup> *IceTV Pty Limited v Nine Network Australia Pty Ltd* [2009] HCA 14; *Telstra Corporation Limited v Phone Directories Company Pty Ltd* [2010] FCAFC 149.
- <sup>63</sup> *CoPilot AI Experiences Terms* s 2.
- <sup>64</sup> *CoPilot AI Experiences Terms* s 5.
- <sup>65</sup> *CoPilot AI Experiences Terms* s 5.
- <sup>66</sup> Bing Conversational Experiences and Image Creator Terms, s 7 <<https://www.bing.com/new/termsfuse/archives?features=bingnewpage>>.
- <sup>67</sup> *OpenAI Terms of use*, ‘Ownership of Content’.
- <sup>68</sup> *OpenAI Terms of use*, ‘Our Use of Content’, ‘Opt Out’.
- <sup>69</sup> Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ‘Machine Bias’, *ProPublica* (online, 23 May 2016) <<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>>.
- <sup>70</sup> Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ‘Machine Bias’, *ProPublica* (online, 23 May 2016) <<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>>.
- <sup>71</sup> Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ‘Machine Bias’, *ProPublica* (online, 23 May 2016) <<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>>.
- <sup>72</sup> Australian Human Rights Commission, *Human Rights and Technology: Final Report* (2021), 105.
- <sup>73</sup> Australian Government, Department of Industry, Science and Resources, *Australia’s AI Ethics Principles* <<https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-framework/australias-ai-ethics-principles>> ‘Fairness’.
- <sup>74</sup> T.J. Thomson and Ryan J. Thomas, *Ageism, sexism, classism and more: 7 examples of bias in AI-generated images*, *The Conversation* (Web Page, 10 July 2023) <<https://theconversation.com/ageism-sexism-classism-and-more-7-examples-of-bias-in-ai-generated-images-208748#:~:text=There%20were%20also%20notable%20differences,of%20more%20fluid%20gender%20expression.>>>.
- <sup>75</sup> Adam Satariano, ‘ChatGPT Is Banned in Italy Over Privacy Concerns’, *The New York Times* (online, 31 March 2023) <<https://www.nytimes.com/2023/03/31/technology/chatgpt-italy-ban.html#:~:text=Italy%20is%20the%20first%20government,not%20to%20make%20it%20accessible.>>>.
- <sup>76</sup> Supantha Mukherjee and Griselda Vagnoni, ‘Italy restores ChatGPT after OpenAI responds to regulator’ *Reuters* (online, 29 April 2023) <<https://www.reuters.com/technology/chatgpt-is-available-again-users-italy-spokesperson-says-2023-04-28/>>. Under the *OpenAI Terms of Use* ss 2 and 4, users of ChatGPT may opt out of model training and may request that OpenAI remove their personal data from the model. See, OpenAI Personal Data Removal Request <[https://share.hsforms.com/1UPy6xqxZSEqTrGDh4ywo\\_g4sk30](https://share.hsforms.com/1UPy6xqxZSEqTrGDh4ywo_g4sk30)>.
- <sup>77</sup> *Open AI Privacy Policy*, s 4 <<https://openai.com/policies/privacy-policy>>.
- <sup>78</sup> Johan Moreno, ‘Generative AI In The Office: To Ban or Not Ban’ *Forbes* (online, 28 June 2023) <<https://www.forbes.com/sites/johanmoreno/2023/06/28/generative-ai-in-the-office-to-ban-or-not-ban/?sh=292ae306494d>>.
- <sup>79</sup> See e.g. *Legal Profession Uniform Law Australian Solicitors’ Conduct Rules 2015* (NSW).
- <sup>80</sup> See e.g. *Code of Conduct for Trans-Tasman Patent and Trade Marks Attorneys 2018*.
- <sup>81</sup> New South Wales Bar Association, *Issues Arising from the Use of AI Language Models (including ChatGPT) in Legal Practice* (12 July 2023) <<https://inbrief.nswbar.asn.au/posts/9e292ee2fc90581f795ff1df0105692d/attachment/NSW%20Bar%20Association%20GPT%20AI%20Language%20Models%20Guidelines.pdf>>; Law Society of New South Wales, Professional Support Unit, ‘A solicitor’s guide to responsible use of artificial intelligence’, *Law Society Journal* (online, 14 November 2023) <<https://lsj.com.au/articles/a-solicitors-guide-to-responsible-use-of-artificial-intelligence/>>.
- <sup>82</sup> See e.g. *Legal Profession Uniform Law Australian Solicitors’ Conduct Rules 2015* (NSW) s 3.
- <sup>83</sup> See e.g. *Legal Profession Uniform Law Australian Solicitors’ Conduct Rules 2015* (NSW) s 4.
- <sup>84</sup> *McConnell Dowell Constructors v Santam* [2016] VSC 734, [12]. See also Supreme Court of Victoria, Practice Note SC Gen 5 ‘Technology in Civil Litigation (2018) [8.7]-[8.9].’