



## Assessing law students in a GenAI world to create knowledgeable future lawyers

Amanda Head & Sonya Willis

**To cite this article:** Amanda Head & Sonya Willis (2024) Assessing law students in a GenAI world to create knowledgeable future lawyers, International Journal of the Legal Profession, 31:3, 293-310, DOI: [10.1080/09695958.2024.2379785](https://doi.org/10.1080/09695958.2024.2379785)

**To link to this article:** <https://doi.org/10.1080/09695958.2024.2379785>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 19 Jul 2024.



Submit your article to this journal [↗](#)



Article views: 1044



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

# Assessing law students in a GenAI world to create knowledgeable future lawyers

Amanda Head and Sonya Willis

Law School, Macquarie University, Sydney, Australia

## ABSTRACT

Assessing law students has always been a challenging task, but the introduction of Generative Artificial Intelligence (GenAI), such as ChatGPT, compounds the problems already caused by increased student numbers, contract cheating and budget cuts in universities. As GenAI rapidly develops, legal educators must find ways to accommodate, and even incorporate, GenAI into their curricula and assessments so that law graduates can understand its capabilities and limitations within legal practice. Simultaneously, many jurisdictions, including Australia, have legislative obligations to deliver law graduates who satisfy legal knowledge-based “eligibility” requirements for admission into practice. This article introduces a knowledge framework for managing GenAI in legal education consisting of three pillars: Substantive Legal Knowledge, GenAI Ethics Knowledge, and GenAI System Knowledge. The authors argue this framework can assist legal educators in designing optimal assessments in an AI-disrupted world. The article employs the knowledge framework to examine the experiences and views of Australian law students’ engagement with GenAI outputs in completing a compulsory legal ethics assessment in 2023. This empirical case study demonstrates that effective assessment design incorporating GenAI can enhance law student and graduate outcomes despite the ongoing challenges for legal educators and the profession associated with GenAI.

## ARTICLE HISTORY

Received 21 March 2024

Accepted 10 July 2024

## 1. Introduction

The public release, in November 2022, of ChatGPT, and the many Large Language Models (LLMs) that have followed, have imposed a significant new teaching and assessment burden on legal academics, attracting worldwide attention. This article introduces a knowledge framework for assessing law

**CONTACT** Amanda Head  [amanda.head@mq.edu.au](mailto:amanda.head@mq.edu.au)

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

students in a Generative Artificial Intelligence (GenAI) world, presenting the results of an Australian case study that focuses on the experience and views of first-year law students engaging with output from GenAI in researching and writing a legal ethics essay. Our empirical evidence from the case study informs our analysis of ways to incorporate GenAI meaningfully and ethically in student learning.

Student voice is essential when considering any impact on student learning (Sullivan *et al.* 2023). Student views on using GenAI in higher education have received some attention (Chan 2023; Chan and Hu 2023; Farhi *et al.* 2023; Firat 2023; Huallpa *et al.* 2023; Shoufan 2023; Smolansky *et al.* 2023). There is, however, minimal research analysing the views of law students engaging with GenAI technology in a compulsory university assessment task. Our research also contributes to scholarship on the benefits of reflective practice (Casey 2014), particularly its interaction with the use of GenAI (Ajevski *et al.* 2023).

This article begins with an overview of GenAI and the existing literature engaging with the effects of GenAI on higher education, particularly in the legal education sector. To evaluate GenAI's impact on law graduates entering the legal profession, we have formulated a knowledge framework covering three key areas, which, we argue, are essential to optimising assessment design for graduate capabilities in a GenAI disrupted world. This article employs our knowledge framework to analyse the results of an empirical study undertaken with first-year law students. Our empirical analysis of an assessment incorporating GenAI demonstrates the potential to embed an understanding of GenAI and its inherent risks and benefits into student learning outcomes through considered assessment design. We conclude by identifying ongoing challenges for legal educators.

## 2. Background

### 2.1. The basics of generative AI, LLMs and ChatGPT

Computer programs designed to simulate human communications have existed since at least the 1960s (Adamopoulou and Moussiades 2020). GenAI is a subset of Artificial Intelligence designed to generate new data or content, typically text, images, or sounds, in response to a prompt. The focus of this article is on large language models (LLMs) which simulate human text. There is an increasing number of GPT-trained (Generative Pre-trained Transformer) foundation models for LLMs which power customer-facing interfaces including ChatGPT by OpenAI (which offers instruction-following and chatbot style interfaces), Claude by Anthropic, Google's Gemini family, and Llama by Meta (which has led to the proliferation of open-source models, e.g. Hugging Face (2024)). The release of chatbot interfaces, particularly ChatGPT, brought this technology to the broader community's attention.

LLMs are powered by several layers of transformers. These transformers refer to neural network architecture that processes sequential data and can generate coherent and human-like outputs (Rothman 2022, pp. 169–171). Although neural network architecture has powered OpenAI's GPT for 6–7 years, what distinguished ChatGPT was the high volume of data it was trained on (Rudolph *et al.* 2023(b)). The technology is statistical or prediction-based: it takes an input, draws on the dataset it has been trained on, and provides what it predicts as the most likely output (Ajevski *et al.* 2023). Due to the exceptionally large dataset, GPT can predict complex language nuances, enabling it to effectively mimic human output. However, ChatGPT and the other LLMs, being statistical models, do not comprehend correctness or bias and are prone to hallucination and confabulation (Ajevski *et al.* 2023). Unlike most humans, LLM output employs an authoritative tone, regardless of its content accuracy.

## **2.2. Scholarly recognition of GenAI's effect on the education sector**

Despite its recency, the launch of ChatGPT has already been identified as a significant disruptor to the education sector generally (Chan 2023; Farazouli *et al.* 2023; Halaweh 2023; Rudolph *et al.* 2022(a)) and legal education specifically (Ajevski *et al.* 2023). Notable academic commentary, particularly from the US, reports the success of ChatGPT in passing some law exams (Choi and Schwarcz 2023; Perlman 2022; see also Hargreaves 2023) and discusses how legal educators might design assessments to combat the threat GenAI poses (Croft 2023; Hargreaves 2023; Katz *et al.* 2023; Ryznar 2023; Sullivan *et al.* 2023). Existing literature also identifies employer expectations for law graduates to be technologically competent, including understanding technology like ChatGPT (Sharma 2023) and its ethical use (Tarves 2023).

Legal educators must, therefore, balance potentially conflicting objectives. They must engage students with the new technology in terms of its capabilities and risks whilst simultaneously designing assessment regimes that probe students' knowledge and abilities beyond what an LLM can produce.

## **3. The knowledge framework**

This article introduces a knowledge framework consisting of three knowledge areas that we argue should govern the assessment requirements for today's law students. First, students require substantive legal knowledge in key areas (substantive legal knowledge). The ubiquity of GenAI has now imposed the second and third knowledge areas for optimal legal education, namely, the knowledge of the legal and ethical risks of GenAI (GenAI ethics knowledge) and the skills to use GenAI effectively (GenAI system knowledge). Historically,

engaging with any technology played a secondary role in legal education (An 2023, p. 280). This is arguably no longer a tenable position. The three knowledge areas comprising our framework are discussed below.

### 3.1. Substantive legal knowledge

Legal education in most jurisdictions, including Australia, is highly regulated (Galloway 2020). All Australian jurisdictions mandate minimum academic knowledge requirements for the practice of law (traditionally called the Priestley 11). In New South Wales, Australia's most populous state (ABS 2023), these requirements are embedded within the *Legal Profession Uniform Law 2014* (NSW) (LPULNSW) and the *Legal Profession Uniform Admission Rules 2015* (NSW) (LPULRNSW). Section 17(1)(a) LPULNSW requires, as a prerequisite for admission, that a person has met the "specified academic qualifications prerequisite" (applicants for admission also require practical legal training: s 17(1)(b), and to be a fit and proper person: s 17(1)(c)). The "specified academic qualification prerequisite", requires completion of an "academic course in Australia" that is the equivalent of "at least 3 years' full-time study of law", "is accredited", and will allow the student to "acquire and demonstrate appropriate understanding and competence in each element of the academic areas of knowledge set out in Schedule 1" (r 5 LPULRNSW). Other Australian states and territories have identical or similar requirements (Dal Pont 2021, p. 40). Therefore, legal education providers are legislatively required to assess students' ability to "demonstrate appropriate understanding and competence" in all 11 knowledge areas.

GenAI's ability to pass many traditional legal assessments potentially undermines this substantive legal knowledge requirement, posing a significant risk to the legal profession and the law academy. Threats to the integrity of substantive legal knowledge assessments are not new. Contract cheating, for example, poses a similar threat. However, the Australian higher education regulator, the Tertiary Education Quality and Standards Agency (TEQSA), can block websites offering cheating services (Hargreaves 2023, p. 87). A similar ban on GenAI is neither possible nor desirable (Hargreaves 2023, pp. 84–88; Cotton *et al.* 2024, p. 8). Indeed, TEQSA has recently acknowledged that a total ban would be an "oversimplified" response given the complexity of the problem and the likely future place of GenAI in society (TEQSA 2023, p. 2). Instead, TEQSA focuses on "ways assessment practices can take advantage of the opportunities, and manage the risks of AI, specifically generative AI" (TEQSA 2023, p. 1). Therefore, legal educators must acknowledge GenAI when designing assessments to evaluate students' "understanding and competence" of substantive legal knowledge.

### 3.2. GenAI ethics knowledge

In addition to substantive legal knowledge, law students and graduates must understand GenAI's ethical implications. Using GenAI to assist with legal work poses several risks at both student and professional levels.

#### 3.2.1. Poor academic practice

Legal educators must educate students on academic integrity (Universities Australia 2017). The academic integrity concerns surrounding using GenAI, particularly ChatGPT, are well documented (Chan 2023; Plata *et al.* 2023; Sullivan *et al.* 2023). The ease of access to this technology and its ability to rapidly produce grammatically correct original human-like content on legal topics, leaves these tools open to misuse and abuse. Studies have demonstrated that GenAI can create a passable answer to some law assessments with minimal student effort in prompting the GenAI (Katz *et al.* 2023; Perlman 2022, p. 3).

Although studies have found that engagement with GenAI has the potential to improve student motivation and engagement (Muñoz *et al.* 2023; Shoufan 2023), there is a risk students may become dependent on these tools (Huallpa *et al.* 2023 p. 111) which may erode their autonomy, proactiveness, initiative and curiosity (Yu 2023) and encourage laziness and apathy in learning (An 2023, p. 280). Additionally, although using GenAI in completing assessments may assist lower-performing students, it may also negatively impact top-performing students (Choi and Schwarcz 2023, p. 17).

#### 3.2.2. Professional responsibility

A lawyer's duty of competence and diligence is fundamental. In Australia, it is a stand-alone duty under professional conduct rules but also part of lawyers' duties more broadly to the court and the administration of justice. However, the conduct rules in Australian jurisdictions do not explicitly refer to technological competence. There has been commentary on whether an Australian lawyer's duty of competence includes some level of technological competence in the context of law firm security of confidential data stored in the cloud (Herbert-Lowe 2021). Similar concerns apply to the use of GenAI technology in legal practice. Lawyers may risk breaching client confidentiality and their duty of competence if, for example, they enter clients' confidential information into LLMs to assist with document drafting without first carefully considering the terms of service and the risks associated with using the technology, including who may have access to the client's entered information, and how to properly review the output.

The US Model Rules of Profession Conduct go further than Australia by providing inferential guidance on lawyers' technological competence. Comment 8 to the Model Rule 1.1 on Competence reads:

To maintain the requisite knowledge and skill, a lawyer should keep abreast of changes in the law and its practice, including the benefits and risks associated with

relevant technology, engage in continuing study and education and comply with all continuing legal education requirements to which the lawyer is subject. (ABA 2023)

Ensuring GenAI output is closely reviewed is, therefore, becoming an essential part of legal practice, much like the supervision of junior lawyer output has always been. However, GenAI does not understand, review or check its output for correctness, so there is a significant risk of inaccuracies (Murray 2020; Yu 2023, p. 222). Additionally, LLMs typically produce output that is confidently expressed despite containing inaccuracies (Leiser *et al.* 2023, p. 87), which, without adequate technological competence, may mislead supervising lawyers. The risk was demonstrated by two US attorneys who unwittingly submitted GenAI-produced “fake” authorities to the court. Notably, the lawyers were reprimanded and fined, not for using GenAI technology, but for their lack of oversight of their submission to court (*Mata v Avianca Inc.* (U.S. Dist., SD NY, No 22-cv-1471, 22 June 2023); De Poloni 2023).

### 3.2.3. GenAI system knowledge

Despite the identifiable risks and ethical concerns surrounding this technology, ignoring GenAI is unlikely to be a tenable long-term approach. GenAI can potentially transform how students and lawyers learn and work. There appears to be growing consensus that GenAI will be integrated into the workplace to assist lawyers rather than replace them, (Ajevski *et al.* 2023, p. 357; Perlman 2022, pp. 22–23) at least in the short to medium term (Alarie *et al.* 2018). Indeed, law firms appear to be moving in this direction. Australian firm, MinterEllison, for example, has developed an in-house advice generator powered by OpenAI’s LLM, which it claims produces work to the standard of a graduate lawyer in a fraction of the time (Yim 2023; see also Allen & Overy 2023). Competition and efficiency factors will drive the uptake of this technology in legal practice, increasing the importance of effective (and appropriate) use of GenAI. Professional reliance on GenAI constitutes a powerful rationale for legal educators to train students on the effective use of this technology (Brescia 2023; Perlman 2022, p. 22).

Moreover, GenAI jeopardises the employment prospects of law graduates, particularly given that GenAI output can be equal or superior to that of paralegals and graduate lawyers (Ajevski *et al.* 2023, pp. 356–357; Iu and Wong 2023, p. 17). In an increasingly competitive graduate market, there is a clear employability risk for students who have not developed technological competence at university.

## 4. GenAI case study

The GenAI case study analysed the reflections of students on employing GenAI outputs in a legal ethics assessment task.



#### **4.1. Overview and methodology**

The case study assessment was conducted in Semester 2 of 2023 in a first-year legal ethics subject titled Law, Lawyers and Society at Macquarie University in Australia (GenAI Assessment). The GenAI Assessment required students to submit a 1500-word essay responding to an essay question on whether the legal profession is best viewed as a profession or a business. Students were required to support their answers with references to the NSW law governing professional conduct including references to at least four good-quality secondary sources. As part of this assessment, the subject coordinator provided students with two GenAI outputs – from ChaptGPT 3.5 and Google Bard – and students were instructed to engage with the outputs as a starting point for their response. The GenAI outputs were produced by the subject coordinator using the prompt: “Is legal practice a business or a profession? Include in your answer references to scholarly sources and to the uniform law as it applies in New South Wales”. ChatGPT produced a 507-word response and cited three fictitious sources. Google Bard produced a 345-word response with four fictitious sources. The students were explicitly notified that the responses included fictitious sources.

Students were also required to submit 500 words reflecting on their experience engaging with the GenAI outputs. Students were provided with some suggestions for matters they may wish to reflect on, including: initial reactions and whether the outputs assisted them, whether parts of the outputs were compelling or requiring improvement, how they critiqued the output and improved upon any aspects of the output they adapted, reasons for not using the output, their critique of the prompting used and suggestions for improvement, the value of the counterarguments in the outputs, and the potential impact of GenAI on the legal profession. A core aim of the GenAI Assessment was to increase students’ GenAI ethics knowledge. Providing GenAI outputs and incorporating reflective practice were critical aspects of the assessment design in achieving this aim.

The data for the case study comprises quantitative metrics from the GenAI Assessment and a similar assessment in a prior year (non-GenAI Assessment) and qualitative data consisting of a set of volunteered student reflections from the GenAI Assessment. The authors obtained ethics approval from Macquarie University under the Australian National Statement on Ethical Conduct in Human Research (2007) updated 2018, and invited students to share a copy of their reflections anonymously. Twenty percent of the cohort (23 out of 114 students) participated in the research (the participants) and shared their reflections (labelled P1 to P23), which the authors coded in NVivo.

The following two sections detail the main findings from the case study. Part 5 will review these findings against the knowledge framework.



## 4.2. Subject coordinator observations

The results from the 2023 GenAI Assessment were compared to those of a non-GenAI Assessment from 2022. The 2022 and 2023 cohorts had similar enrolment numbers and identical markers and marking standards. Consistent with other studies in this area (Choi and Schwarcz 2023, p. 25), the subject coordinator observed a negligible change in the median and mean average marks between the GenAI Assessment in 2023 and the non-GenAI Assessment in 2022. However, fail grades fell from around 4% in 2022 to 0% in 2023 (controlling for late penalties), and the AI assisted cohort also recorded significantly fewer late submissions (8.6% in 2023 compared with 19.1% in 2022). These results are consistent with findings from other studies that GenAI assistance tends to most benefit lower-performing students (Choi and Schwarcz 2023) or students at risk (Sullivan *et al.* 2023, p. 37).

## 4.3. Findings from the qualitative analysis

In their reflections, participants expressed a wide range of views. Many participants in the case study found engaging with the GenAI outputs beneficial in completing their assessment, though numerous participants also expressed concern about GenAI's limitations. In particular, participants commonly reported that the GenAI output assisted in the initial stages of their research despite observing the unreliability of the GenAI output.

### 4.3.1. Positive experiences with GenAI

Most participants reported that the GenAI outputs assisted them in completing the assessment. Of the 23 participants, 78% noted that the GenAI assisted them early in the assessment, particularly with commencing the writing process. Typical responses were that GenAI was helpful as a “starting point” for both formulating the paper's main ideas and the structure of their papers. For example:

Overall, I found that the use of [GenAI] proved to be a valuable resource during the early stages of my assessment ... [T]he [outputs] both offered foundational knowledge and arguments on the topic. (P14)

The [outputs] were a helpful starting point for this assessment. (P15)

It also ensured that I began my research essay promptly rather than staring at a blank document. (P20)

The GenAI outputs particularly assisted participants with understanding the set essay question. P1 noted that it was “beneficial in understanding the question in simple terms” and “gave me an idea of the structure that I could follow, and it also did provide good points and definitions to help me gain an idea of what the

question was asking”. P18 observed that the GenAI outputs were in “plain English wording and did not excessively produce academic vocabulary, which made them accessible for anyone interested in the topic”. Similarly, P7 noted the language used by GenAI was “not too complex or unreadable”, and P11 noted that GenAI assisted in “being able to get an initial understanding of the question”. Numerous students noted that GenAI outputs gave ideas on structuring their essays (P1, P3, P4, P10, P11, P14, P18, P22) and presenting their arguments (P19, P23). GenAI also assisted participants by providing ideas for their main arguments and engaging with various perspectives. For example:

[T]he opposing viewpoints presented in both [outputs] were essential in establishing my argument (P2)

The counterarguments provided by [GenAI] allowed me to strengthen my viewpoints and expand my perspective on the classification of the legal profession as it provided two opposing views. (P3)

Many aspects of the [outputs] were compelling and thought provoking and the resources did assist me in completing my essay, as I was able to use them as a guiding tool to assist me in my writing. (P7)

[T]he [outputs] introduced me to different perspectives on the question that I would likely not have considered on my own. (P11)

These participant reflections were consistent with the subject coordinator’s observations that students asked fewer content and structure-based questions compared with prior offerings of the subject.

The assessment also prompted some participants to reflect on the perceived efficiencies of using GenAI in the legal profession. P15 noted that “large workloads and the culture of demanding efficiency may pressure lawyers towards using it.” Of particular note was the link made by some participants between these efficiencies and the requirement for ethical use of GenAI. For example (emphasis added):

AI has the potential to make a lawyers work more efficient *but only if its writing is sufficiently reviewed by a lawyer.* (P9)

I recognise that AI has the potential to save time and money *if used appropriately.* (P15)

*If AI is used correctly,* it can be a very useful tool in the profession, where work can be made more efficient. (P21)

#### 4.3.2. Concerns with GenAI

The concerns participants expressed with GenAI demonstrated the success of embedding the GenAI ethics knowledge requirement in the assessment task. In their reflections, many participants noted that GenAI should be used with

caution and the need for close scrutiny of GenAI's output. For example, P11 stated, "points and statements made by [GenAI] should be investigated further and cannot be relied upon as a trusted and accurate source". P22 noted the risk of being deceived by GenAI due to its ability to produce "elements of truth, for example, correct author names but incorrect journal or article titles". P8 noted the "importance of using AI as a supplementary tool rather than a replacement for human judgement". Several responses reported a lack of trust in the GenAI outputs and difficulties with verifying the GenAI outputs' assertions. For example:

While both responses answered the question they were equally lacking in detail and did not use enough primary or secondary sources to support their arguments which weakened the responses. (P7)

This is a concerning possibility where someone takes an idea from the AI and cannot find the necessary evidence to support it. (P9)

I noticed that there was no depth to the responses, in that there were no factual evidence or references to any sources mentioned. (P11)

I quickly realised that attempting to find evidence to support the [GenAI]'s assertions was somewhat difficult. (P12)

I could not locate it within the supplied source. As a result, I became certain I would not utilise any of the cited sources. (P14)

Due to the nature of generative AI and its propensity to make facts up, I have a distrust of AI content. (P15)

However, both [GenAI outputs] lacked credibility in their counterarguments lacking analysis using examples from primary and secondary sources, like cases and peer-reviewed journal articles. (P16)

Several participants noted the potential for GenAI to encourage poor academic practice:

I believe students should be strengthening their research and writing skills in order to prepare them for their professional careers. However, the use of Artificial Intelligence limits the effectiveness of this as students are able to acquire a full-length response to any given question. This may cause many students to lack the development of skills during their degree which can hinder their performance once they have completed their qualifications. (P3)

[T]he move away from reading actual literature and doing your own research, to access the more convenient yet completely falsified information a robot provides [and] to use AI would be to show a lack of competence or even laziness, and therefore not be competent, as the AI is doing the work for you. (P6)

This experience helped me realise that although AI may initially assist brainstorming, it can also obstruct further research due to its choices in selecting and presenting ideas. (P12)

[I]t discourages me from thinking on my own and coming up with my own ideas. (P13)

If lawyers or even students wanting to practice law, get their hands on AI, they would tend to waste their skills, knowledge, and creativity and use baseless arguments in their work. (P17)

However, there is also a risk that some rely too much on AI and thus lose their own skills in terms of imagination, analysis, and writing. (P21)

Most participants expressed concern as to the superficiality of the GenAI outputs, observing that they were vague, contained basic arguments, were shallow or had little or no depth (P1, P2, P3, P5, P6, P11, P14, P21, P23). Participants noted the lack of precision with the GenAI outputs' broad-brushed reference to the governing legislative regime:

It is worth noting that the [output] didn't provide any relevant sections part of the Uniform Law, it only cited it. (P13)

[T]here were no specific statements from different sections of the law to support the points within the response. (P16)

Another problem is that very few sources are used. There were no specific references to legal paragraphs. (P21)

Although acknowledging the value of GenAI as a starting point, some participants reported that their final paper did not rely on the GenAI outputs due to concerns with the integrity of the GenAI output:

[GenAIs] are highly unreliable and should not be associated with most academic endeavours. (P12)

I chose not to use the [GenAI] as [they] do not have any reputability in their writing as opposed to peer review journal articles, cases, textbooks, or legislation. (P16)

[GenAIs] talk around the topic rather than presenting a singular strong argument. (P19)

[W]here in-depth legal analysis is required, [GenAI] may be futile. (P23)

Similarly, participants noted how the use of GenAI could negatively impact the legal profession, particularly with the lack of support for the claims it made:

[T]he technology has blind spots so those who employ it, particularly in the court room, should thoroughly review the content generated to ensure its legitimacy. (P9)

[I]ncorrect use of this technology may result in lawyers citing false information in their cases. (P12)

For a lawyer, sourcing material from CHAT GPT could make things difficult, since the law is all about referencing from genuine sources. (P17)

[T]he generation of fake citations is particularly dangerous in law ... From this experience, I can see the allure of using AI in legal practice. It can quickly generate responses

and is vague enough to be somewhat close to the truth in many aspects. However, especially in a field that requires as much precision as law, I would be wary of using it for anything other than a quick idea generator if I am stuck on how to approach an assignment. (P15)

The use of [GenAI] in legal writing prompted reflection upon the future potential effects of AI in the legal profession, such as the ability to give free legal advice which is more cost-effective and therefore accessible to the general public. However, a lack of specialist advice in a [GenAI output] as well as the ability of the technology to fabricate information and sources, this advice could misinform the public and create legal issues. (P18)

Several participants demonstrated GenAI systems knowledge by noting the limitation of the prompts used to produce the GenAI output and reflecting on how direct engagement with the technology may have assisted their assessment. For example:

Perhaps what could make the generated responses even more interesting to work with would be if their responses were about the legal profession being best understood as only profession or only business, with no counterarguments. (P10)

Using a different prompt within the [GenAI] related to how profit motivations can compromise the core values and responsibilities of professional conduct could have been beneficial. This approach would have aided in brainstorming the fundamental arguments in favour of this perspective. (P14)

A prompt shift that potentially would have yielded a more nuanced response would have been to ask the AI if and where it deemed morality to be involved in the classification of the profession. This would make the [GenAI] give a judgement that was not so binary as: a Profession or Business? (P5)

## **5. Discussion of case study findings against the knowledge framework**

The case study findings support the importance of the proposed knowledge framework, particularly the GenAI Assessment's success in engaging students with GenAI ethics knowledge, which was a central aim. Our case study also supported our hypothesis, consistent with existing literature, that GenAI outputs might assist students in structuring their writing (part of GenAI system knowledge) but provide limited assistance with substantive legal knowledge due to the outputs' unsatisfactory legal references. Additionally, as will be seen, the participants' responses provided valuable insights into how the knowledge framework can assist in designing and managing law assessments in a post-GenAI world.

### **5.1. Substantive legal knowledge**

The substantive knowledge area for Law, Lawyers and Society is the law governing professional conduct, a core knowledge area for admission to legal practice.

The GenAI Assessment was designed to encourage students to engage with issues connected with professional conduct, considering lawyers' codes of conduct and how the practice of law as a profession may be in tension with the commerciality of legal practice.

The GenAI outputs did not materially assist students with their substantive legal knowledge. When devising the assessment, the free versions of ChatGPT and Google Bard engaged minimally with the substantive law. The outputs provided to students referred only in general terms to the "Legal Profession Uniform Law" or "the law in NSW" without specific references to legislative provisions, and several participants reported this in their reflections. As such, the use of GenAI in the assessment did not significantly compromise the GenAI Assessment's ability to evaluate the students' substantive legal knowledge.

Given the rapid development of GenAI technology, this specific limitation may be short-lived, particularly with the increased accessibility of individually customised GenAI, which can be enhanced by pre-existing data, such as the relevant law (Savelka *et al.* 2023). Future GenAI may, therefore, interfere more significantly with assessing substantive legal knowledge. The GenAI outputs in the GenAI Assessment potentially assisted students indirectly with substantive legal knowledge by assisting some students to engage with the substantive law more confidently through their independent research. The observations of the subject coordinator with respect to lower-performing students, consistent with the literature discussed above, may also support this observation. The potential benefits and risks of GenAI assistance for lower-performing students require further research.

## **5.2. GenAI ethics knowledge**

As already indicated, an essential aspect of the assessment design was to assist students in attaining knowledge of the ethical use of GenAI technology and its associated risks. Legal ethics and professional responsibility are particularly suitable subject areas for GenAI ethics knowledge as potential (mis)use of technology is central to legal professional competence, including duty of care, and raises clear concerns about protecting clients' confidential and privileged information. As the reported US cases discussed above demonstrate, it is a breach of a lawyer's duties to their clients and the court to engage with GenAI technology without competent oversight.

To encourage students' consideration of the ethical use of GenAI as a research tool, the assessment explicitly advised students that the GenAI outputs were not authoritative and provided false citations. The findings from the case study demonstrate that the participants engaged well with this aspect of the assessment. Many responses showed that the participants were beginning to understand the ethical aspects of engaging with GenAI. In

particular, they noted the lack of support for GenAI's propositions and the crucial need to oversee GenAI's output. Several participants noted how these risks could impact legal practice. Additionally, participants observed GenAI's propensity to produce vague or superficial output. The potential of GenAI to promote poor academic practice was also noted by participants, particularly the risk that overreliance on the technology could impede skills development and dampen their creativity.

### **5.3. GenAI system knowledge**

In the GenAI Assessment, students' engagement with GenAI was controlled as they were only permitted to engage with the tools' output and not the tools directly. Additionally, the prompt used to generate the output in both GenAI tools was a "basic prompt" (Choi and Schwarcz 2023, pp. 15–16). Notwithstanding these limitations, the case study demonstrated that the GenAI Assessment improved many participants' GenAI system knowledge. In particular, many reflected that they had learned more about GenAI and its capabilities from undertaking the task. Several participants also considered how they might change the prompt or engage further with the technology.

Research indicates that knowledge of prompt engineering will be an essential aspect of GenAI system knowledge as the quality of the prompts can have a significant effect on the quality of the output produced by the GenAI technology (Chan 2023; Choi and Schwarcz 2023; Hargreaves 2023, pp. 79–80). This is not dissimilar to understanding the effective use of Boolean operators when interrogating research databases, where the quality of the search terms influences the quality of the search results (Lowe *et al.* 2020, p. 3). A comprehensive AI literacy programme as part of legal education (including prompt engineering and evaluation of outputs) will help law students better understand and responsibly utilise GenAI technologies in their academic and professional lives.

## **6. Conclusion**

This article has demonstrated three important knowledge-based reasons why legal academics cannot ignore GenAI when educating the next generation of lawyers. GenAI has the potential to enable law students to provide mostly correct answers to basic legal questions and is capable of passing many traditional law assessments without those students knowing the law. The need, both legislative and professional, for lawyers to personally comprehend legal principles means that law assessments, particularly in core law subjects, must address the capabilities of GenAI and be designed in a way that tests the actual knowledge of the students through effectively excluding or, where possible, curating the use of GenAI.

In addition to legal knowledge, the very existence of GenAI obliges legal academics to teach students the risks associated with its use. The academic



integrity and professional conduct risks associated with GenAI are significant. Yet, many students will be unaware of these until they are taught. Banning the use of GenAI in law assessments is an insufficient acknowledgment of the risks. Instead, legal academics must ensure students know the legal and ethical risks of GenAI use in student and professional contexts. As the case study demonstrates, in Australia, professional conduct subjects are an ideal knowledge area to provide GenAI ethical risk education.

Law students and legal employers rely on legal academics to prepare students for the legal profession. As law firms increasingly interact with GenAI through the documents they create and those created by others, the need for lawyers to optimally engage with GenAI will increase (see also Law Society of NSW 2023). Within a few years, competent law graduates may be expected to expertly prompt GenAI to assist their research and writing. Simple assessments, such as the GenAI prompt-based essay and reflection used in the Law, Lawyers and Society subject considered in this case study, demonstrate that it is relatively simple to design an assessment that provides students with significant insight into the operation and risks of GenAI. As this case study shows, simple empirical research into these student insights can inform legal academics to design improved GenAI based assessments.

In conclusion, legal academics need considerable expertise in GenAI, including knowledge of its ethical risks, how to utilise it optimally given its strengths and limitations, and how to ensure it does not usurp students' comprehension of core legal principles. Not every assessment in every subject needs to meet all three knowledge requirements, but they must all be met during each student's law degree. The burden of achieving this will likely fall disproportionately on those teaching compulsory law subjects. Unfortunately, due to the recency of GenAI, most legal academics lack GenAI expertise. The challenge for universities is to provide the necessary resources and support for their staff to gain the GenAI skills they urgently require to teach and assess law students in a GenAI disrupted world. The sooner legal academics engage with GenAI in their legal teaching, and the more research is undertaken on those engagements, such as the case study discussed in this article, the more likely universities will be sufficiently equipped to produce a new generation of competent, knowledgeable lawyers.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## References

- Adamopoulou, E. & Moussiades, L. (2020) Chatbots: history, technology, and applications, *Machine Learning with Applications*. Available at: <https://www.sciencedirect.com/science/article/pii/S2666827020300062>, accessed 15 August 2023.

- Ajeviski, M., Barker, K., Gilbert, A., Hardie, L. & Ryan F. (2023) ChatGPT and the future of legal education and practice, *The Law Teacher*, 57, pp. 352–364. doi:10.1080/03069400.2023.2207426, accessed 19 November 2023.
- Alarie, B., Niblett, A. & Yoon, A. H. (2018) How artificial intelligence will affect the practice of law, *University of Toronto Law Journal*, 68, pp. 106–124.
- Allen & Overy LLP. (2023) Artificial Intelligence, *Allen & Overy*. Available at: <https://www.allenoverly.com/en-gb/global/expertise/practices/artificial-intelligence>, accessed 28 December 2023.
- American Bar Association (ABA) (2023) Rule 1.1 Competence – Comment. Available at: [https://www.americanbar.org/groups/professional\\_responsibility/publications/model\\_rules\\_of\\_professional\\_conduct/rule\\_1\\_1\\_competence/comment\\_on\\_rule\\_1\\_1/](https://www.americanbar.org/groups/professional_responsibility/publications/model_rules_of_professional_conduct/rule_1_1_competence/comment_on_rule_1_1/), accessed 1 November 2023.
- An, Q. (2023) Challenges and responses: reflection on legal education in the age of artificial intelligence, *Advances in Education, Humanities and Social Science Research*, 6(1), pp. 279–282.
- Australian Bureau of Statistics (ABS). (2023) National, state and territory population, *Australian Bureau of Statistics*. Available at: <https://www.abs.gov.au/statistics/people/population/national-state-and-territory-population/mar-2023>, accessed 6 December 2023.
- Brescia, R. H. (2023) Teaching to the tech: law schools and the duty of technology competence, *Washburn Law Journal*, 62(3), pp. 507–540.
- Casey, T. (2014) Reflective practice in legal education: the stages of reflection, *Clinical Law Review*, 20(2), pp. 317–354.
- Chan, C. K. Y. (2023) A comprehensive AI policy education framework for university teaching and learning, *International Journal of Educational Technology in Higher Education*, 20, pp. 38. doi:10.1186/s41239-023-00408-3
- Chan, C. K. Y. & Hu, W. (2023) Students' voices on generative AI: perceptions, benefits, and challenges in higher education, *International Journal of Educational Technology in Higher Education*, 20, pp. 43. doi:10.1186/s41239-023-00411-8
- Choi, J. H. & Schwarcz, D. (2023) AI Assistance in legal analysis: an empirical study, *Minnesota Legal Studies Research Paper No. 23-22*. Available at SSRN: <https://ssrn.com/abstract=4539836>, accessed 1 November 2023.
- Cotton, D. R. E., Cotton, P. A. & Shipway, J. R. (2024) Chatting and cheating: Ensuring academic integrity in the era of ChatGPT, *Innovations in Education and Teaching International*, pp. 228–239. Available at: doi:10.1080/14703297.2023.2190148, accessed 1 November 2023.
- Croft, L. (2023) Authentic Law School assessments to combat use of ChatGPT to cheat, *Lawyers Weekly*. Available at: [https://www.lawyersweekly.com.au/newLaw/36513-authentic-law-school-assessments-to-combat-use-of-chatgpt-to-cheat](https://www.lawyersweekly.com.au/newlaw/36513-authentic-law-school-assessments-to-combat-use-of-chatgpt-to-cheat), accessed 15 August 2023.
- Dal Pont, G. E. (2021) *Lawyers Professional Responsibility* (Pymont, Thomson Reuters).
- De Poloni, G. (2023) Lawyers in the United States blame ChatGPT for tricking them into citing fake court cases, *ABC News*, 9 June.
- Farazouli, A., Cerratto-Pargman, T., Bolander-Laksov, K. & McGrath, C. (2023) Hello GPT! Goodbye home examination? An exploratory study of AI chatbots impact on university teachers' assessment practices, *Assessment & Evaluation in Higher Education*, 49, pp. 363–375. Available at: doi:10.1080/02602938.2023.2241676, accessed 1 November 2023.
- Farhi, F., Jeljeli, R., Aburezeq, I., Dweikat, F. F., Al-shami, S. A. & Slamene, R. (2023) Analyzing the students' views, concerns, and perceived ethics about chat GPT usage. *Computers and Education: Artificial Intelligence*, 5: 100180.

- Firat, M. (2023) What ChatGPT means for universities: perceptions of scholars and students, *Journal of Applied Learning and Teaching*, 6(1), pp. 57–63.
- Galloway, K. (2020) Is legal education over-regulated or under-regulated?, *Griffith University Law School Research Paper 21-2*. Available at: <https://ssrn.com/abstract=3719652>, accessed 14 November 2023.
- Halaweh, M. (2023) ChatGPT in education: strategies for responsible implementation, *Contemporary Educational Technology*, 15(2), pp. 1–11.
- Hargreaves, S. (2023) ‘Words are flowing out like endless rain into a paper cup’: ChatGPT & Law School Assessments, *Legal Education Review*, 33(1), pp. 69–105.
- Herbert-Lowe, S. (2021) Solicitors’ duties in the digital era – Is there a duty of technological competence? *Law Society of NSW Journal Online*. Available at: <https://lsj.com.au/>, accessed 2 November 2023.
- Huallpa, J. J., Arocutipa, J. P. F., Panduro, W. D., Huete, L. C., Limo, F. A. F., Herrera, E. E., Callacna, R. A. A., Flores, V. A., Romero, M. A. M., Quispe, I. M. & Hernández, F. A. (2023) Exploring the ethical considerations of using Chat GPT in university education, *Periodicals of Engineering and Natural Sciences (PEN)*, 11(4), pp. 105–115.
- Hugging Face (2024) *Hugging Face: The AI community building the future*. Available at: <https://huggingface.co/>, accessed 3 February 2024.
- Iu, K. Y. & Wong, V. M. (2023) ChatGPT by OpenAI: the end of litigation lawyers? Available at SSRN: <https://ssrn.com/abstract=4339839>, accessed 2 November 2023.
- Katz, D. M., Bommarito, M. J., Gao, S. & Arredondo, P. D. (2023) *GPT-4 passes the bar exam*. Available at SSRN: <https://ssrn.com/abstract=4389233>, accessed 15 August 2023.
- Law Society of New South Wales, Professional Support Unit. (2023) A solicitor’s guide to the responsible use of artificial intelligence, *Law Society of NSW Journal Online*. Available at: <https://lsj.com.au/articles/a-solicitors-guide-to-responsible-use-of-artificial-intelligence/>, accessed 3 February 2024.
- Leiser, F., Eckhardt, S., Knaeble, M., Maedche, A., Schwabe, G. & Sunyaev, A. (2023) From ChatGPT to FactGPT: A participatory design study to mitigate the effects of large language model hallucinations on users, *Proceedings of Mensch und Computer 2023*, pp. 81–90.
- Lowe, M. S., Stone, S. M., Maxson, B. K., Snajdr, E. & Miller, W. (2020) Boolean redux: Performance of advanced versus simple boolean searches and implications for upper-level instruction. *The Journal of Academic Librarianship*, 46, pp. 102234. doi:10.1016/j.acalib.2020.102234
- Muñoz, S. A. S., Gayoso, G. G., Huambo, A. C., Tapia, R. D. C., Incaluque, J. L., Aguila, O. E. P., Cajamarca, J. C. R., Acevedo, J. E. R., Rivera, H. V. H. & Arias-González, J. L. (2023) Examining the impacts of ChatGPT on student motivation and engagement, *Social Space*, 23(1), pp. 1–27.
- Murray, M. D. (2023) Artificial Intelligence and the Practice of Law Part 1: Lawyers Must be Professional and Responsible Supervisors of AI. Available at SSRN: <https://ssrn.com/abstract=4478588>, accessed 7 December 2023.
- Perlman, A. (2022) The Implications of ChatGPT for Legal Services and Society, *Suffolk University Law School Research Paper No. 22-14*. Available at SSRN: <https://ssrn.com/abstract=4294197>, accessed 7 December 2023.
- Plata, S., De Guzman, M. A. & Quesada, A. (2023) Emerging research and policy themes on academic integrity in the age of Chat GPT and generative AI, *Asian Journal of University Education*, 19(4), pp. 743–758. Available at: doi:10.24191/ajue.v19i4.24697, accessed 6 December 2023.

- Rothman, D. (2022) *Transformers for Natural Language Processing: Build, Train, and Fine-Tune Deep Neural Network Architectures for NLP with Python, Hugging Face, and OpenAI's GPT-3, ChatGPT, and GPT-4* (Birmingham, Packt Publishing).
- Rudolph, J., Tan, S. & Tan, S. (2023(a)) ChatGPT: bullshit spewer or the end of traditional assessments in higher education?, *Journal of Applied Learning and Teaching*, 6(1), pp. 342–362.
- Rudolph, J., Tan, S. & Tan, S. (2023(b)) War of the chatbots: bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education, *Journal of Applied Learning and Teaching*, 6(1), pp. 364–389.
- Ryznar, M. (2023) Exams in the time of ChatGPT, *Developments Washington and Lee Law Review Online*, 80(5), pp. 305–322.
- Savelka, J., Ashley, K. D., Gray, M. A., Westermann, H. & Xu, H. (2023) Explaining legal concepts with augmented large language models (GPT-4). *arXiv preprint arXiv:2306.09525*.
- Sharma, A. (2023) The escalation of ChatGPT: how ChatGPT will exert influence on the legal profession?, *Jus Corpus Law Journal*, 3(3), pp. 106–118.
- Shoufan, A. (2023) Exploring students' perceptions of CHATGPT: thematic analysis and follow-up survey, *IEEE Access*, 11, pp. 38805–38818.
- Smolansky, A., Cram, A., Radulescu, C., Zeivots, S., Huber, E. & Kizilcec, R. F. (2023) Educator and student perspectives on the impact of generative AI on assessments in higher education, *Proceedings of the tenth ACM conference on Learning@ Scale*, pp. 378–382.
- Sullivan, M., Kelly, A. & McLaughlan, P. (2023) ChatGPT in higher education: considerations for academic integrity and student learning, *Journal of Applied Learning and Teaching*, 6(1), pp. 31–40.
- Tarves, T. K. (2023) Technology competence instruction and assessment under the principles and standards of legal research competency, *Legal Reference Services Quarterly*, 42(2), pp. 56–70.
- Tertiary Education Quality and Standards Agency (TEQSA). (2023) *Assessment reform for the age of artificial intelligence* (Canberra, Australian Government). Available at: <https://www.teqsa.gov.au/sites/default/files/2023-09/assessment-reform-age-artificial-intelligence-discussion-paper.pdf>
- Universities Australia (2017) *UA academic integrity best practice principles*. Available at: <https://universitiesaustralia.edu.au/wp-content/uploads/2019/06/UA-Academic-Integrity-Best-Practice-Principles.pdf>
- Yim, N. (2023) Australia's biggest law firm, MinterEllison, is using a version of ChatGPT for its first draft of some legal advice, *The Australian*, 4 December.
- Yu, Y. (2023) Discussion on the reform of higher legal education in China based on the application and limitation of artificial intelligence in law represented by ChatGPT, *Journal of Education, Humanities and Social Sciences*, 14, pp. 220–228.