

HEA Symposium Birmingham 27 Feb 2024.

Are our law students ‘robot proof’? AI chatbots and the future of working with computer generated copyrighted works.

By Ursula Smartt, Associate Professor of Law, Northeastern University London.

Abstract

Artificial Intelligence (AI) and ChatGPT are now being used by university law students to write their coursework essays and cheat in examinations. Whilst there is tremendous scope of large language models to revolutionise the legal workspace and reduce tedious fact-checking by law trainees and paralegals, we need to teach our future lawyers to fact-check their AI-generated work for garbage, including sources which simply do not exist, such as President Biden’s counter terrorism strategy when writing about Lord Hoffmann’s dissenting judgment in the ‘Belmarsh’ case (2004). This paper advances that students need to learn how to use such large AI language model chatbots properly, how they will need to understand them in their future legal workplace and equally how they may well breach copyright which can amount to a criminal offence in the UK and the wider world. The opinion expressed in this paper includes examples from the author’s own law teaching practices and assessments. Given the acceptance and acknowledgement that our law students will use AI to ‘deceive’ their tutors and examiners, HE policy needs to address both, how to disrupt this practice by addressing academic misconduct as well as engender a way how we can educate our future lawyers to become robot-proof, where human potential includes reflection of reality and fact, whilst at the same time preparing legal trainees for the future legal workspace without cheating via AI.

Introduction: cheating in law essays with ChatGPT

Artificial intelligence (AI) is a transformative technology, which is already revolutionising many areas of our lives. In November 2023 we saw the official launch of ChatGPT, a large artificial intelligence (AI) language model, a chatbot from Microsoft-backed OpenAI that can parse text and write convincing answers to questions. There are now several AI-powered writing tools which paraphrase and generate custom text, based on student needs.

Innovative AI-technology now helps not only law students, but particularly those writing in subjects such as English, History, Politics and the Humanities, to adjust tone and formality while retaining meaning. Microsoft Bing is now an integrated AI tool with the Bing search engine; there is also Google Bard, Pi, Chatsonic, Jasper, GitHub, Perplexity and Amazon’s Codewhisperer. Such AI-technologies are disrupting a number of industries, dependent on generating large amounts of text, such as media, advertising, education and the legal marketplace.

Here are a couple of extracts from two of my first year Public Constitutional law students who submitted their summative coursework in May 2023. The task was to discuss and evaluate Lord Hoffmann’s dissenting judgment in a significant House of Lords Appeal Committee case, known as the ‘Belmarsh Prisoner Case’, following 9/11 in 2001, when some foreign suspects of the New York Twin Tower terrorism attacks were held without charge at Belmarsh Prison in East London.¹ In both student cases, the Turnitin AI tool had revealed a score of 85% and 90% AI -use, respectively.

Student J wrote:

“The Belmarsh case is a landmark case in the UK and internationally. It is a significant victory for civil liberties, as it set a legal precedent that indefinite detention without trial breaches the ECHR. In addition, the case further established a legal framework for the detention of suspected terrorists, which has been followed as a legal precedent in many countries” – citing in the footnotes and bibliography: Vidaschi, A., & Graziani, C. (2023). *President Biden’s Counter-Terrorism Strategy: Between Old and New Threats*. DPCE Online, 56(Sp 1), and Pratt Jr, W. F. (2022). *The Rhetoric of Judicial Review in the Supreme Courts of the United States and the United Kingdom*.

Student M wrote:

“The issue of discrimination in counter-terrorism measures is complex and controversial. Critics argue that counter-terrorism measures can disproportionately affect certain groups, such as Muslims and people of colour, and can lead to stigmatisation and marginalisation” - citing in the footnotes and bibliography: Richardson, J., et al., *Prevent, Statutory Duties and the Relationship Between the State and UK Universities* (2020) *Journal of Law and Society* 95, 97, and Sivanandan, A., *From resistance to rebellion: Asian and Afro-Caribbean struggles in Britain*. (2005) *Race & Class* 1.

When marking law essays or coursework, I always mark these pieces ‘in reverse’, that is, I check the bibliography and sources used in the footnotes as per OSCOLA.² Whilst it has so far been relatively easy to spot AI-generated law essays, the next generation of AI generated language models will have learnt how to cite more relevant sources, rather than citing President Biden’s counter-terrorism strategy, as Student J did in his ‘Belmarsh’ essay, a British legal case, dating back to 2004. Student M was either desperate or had not studied the

¹ *Re. A (A (FC) and others (FC) v Secretary of State for the Home Department (Conjoined Appeals)* [2004] UKHL 56) (‘The Belmarsh case’).

² Legal references have to adhere to ‘The Oxford University Standard for Citation’ – OSCOLA – which all law students are taught from the start of their legal studies.

case at all, taking the 21 bibliographical sources provided by AI as ‘true’ and ‘correct’. Most of the sources did not exist at all, though admittedly looked genuine.

What such chatbots tend to do, is pick up on key words (as per the Belmarsh-case), such as ‘terrorism’, ‘foreign’ etc. The word ‘dissenting’ in the essay title was then misinterpreted by ChatGPT to mean ‘discrimination’.

How do we prepare the next generation of (law) students to become more robot-proof?

In *Robot-Proof*, Joseph Aoun proposes a way to educate the next generation of higher education students to ‘invent, create, and discover’, to fill the needs in our society that even the most sophisticated AI agent cannot.³ So, what do we mean by being ‘robot-proof’ in higher education? Aoun argues, that our education should not solely be concerned with topping up students' minds with ‘high-octane facts’, but rather ‘calibrating them with a creative mindset and the mental elasticity to invent, discover, or create something valuable to society’. It has long been my aim to prepare our undergraduate law students for the challenge of working for global law firms whilst instilling in them that they can still make a difference in the competitive legal world as a human.

It is a fact that the commercial legal world has adopted generative AI software - such as Harvey (similar to ChatGPT) - to assist its lawyers draft contracts for example. This then ought to feature in modern law syllabuses at under- and postgraduate level to make our students fit for the AI lawyering reality.

Law firms’ use of AI chatbot-co pilots to increase ‘efficiency’

New AI systems in the commercial legal sector are causing concerns about the technology’s threat to millions of jobs. In February 2023, magic circle law firm, Allen & Overy, introduced an AI chatbot called ‘Harvey’ to help its lawyers draft contracts to find ‘efficiencies’ for its business practices.⁴ Former lawyer Winston Weinberg and AI researcher Gabriel Pereyra teamed up to form Harvey in 2022.⁵

³ Joseph E. Aoun (2018) *Robot-Proof. Higher Education in the Age of Artificial Intelligence*. MIT Press. ISBN 978-0262535977. Professor Aoun is President of Northeastern University.

⁴ ‘Allen & Overy introduces AI chatbot to lawyers in search of efficiencies. Magic circle law firm adopts much-hyped tech to help draft legal documents, but insists move will not replace jobs.’ *FT* 15 February 2023.

⁵ The startup Harvey was built using the underlying GPT technology created by OpenAI. Harvey, raised \$5m in 2022, led by the OpenAI Startup Fund in what the startup describes as a ‘co-pilot for lawyers’. Harvey was founded by Winston Weinberg, a former securities and antitrust litigator at law firm O’Melveny & Myers, and

Harvey is now available to around 3,500 individual lawyers at Allen & Overy which may well lead to cost reductions in billable hours and therein redundancies.⁶

The next cohort of our law graduates entering the legal workplace will no doubt find AI chatbots in place. This means that they have to be tech-savvy in order to be alert to fact-check any information generated by Harvey or similar such AI software. Instead of manually editing legal documents or performing legal research, Harvey enables lawyers to describe the task they wish to accomplish in simple instructions and receive the generated result. To enable this, Harvey leverages large language models to both understand users' intent and to generate the correct output. Similar to ChatGPT, Harvey can answer questions asked in natural language, such as, "tell me what the differences are between an employee and independent contractor?," and "tell me if this clause in a lease is in violation of California law, and if so, rewrite it so it is no longer in violation."

How to teach future law students with AI

Should we stop teaching contract law at undergraduate level completely and instead teach our students how to use AI effectively by encouraging coding and how to make use of Harvey? Such AI legal tools are very powerful but can also be fraught with difficulties and fake information. Just like Wikipedia is created by human information input and not necessarily fact-checked, AI chatbots use language models, hoovering up text, photos, music and copyrighted materials indiscriminately, without fact-checking or asking the author for permission of copyright. Many of the sources in referencing a bibliography are simply made up. When using such bots in the legal workplace, some documents such as contracts can become toxic and would not stand up in court. The small print of the Harvey software has a disclaimer which reads: "The tool isn't meant to provide legal advice to non-lawyers and should be used under the supervision of licensed attorneys." That said, we cannot forbid our students to use ChatGPT or Harvey – we just need to make them aware that they are a reasonably good starting point for editing and improvement of a final document.

Gabriel Pereyra, previously a research scientist at DeepMind, Google Brain (another of Google's AI groups) and Meta AI. Also participating have been Jeff Dean, the lead of Google AI, and Mixer Labs co-founder Elad Gil, among other angel backers.

⁶ 'A&O announces exclusive launch partnership with Harvey', Allen & Overy website announcement 15.2.2023: <https://www.allenoverly.com/en-gb/global/news-and-insights/news/ao-announces-exclusive-launch-partnership-with-harvey>

That said, law students ought to be aware of ethical issues and concerns of using the technology in legal settings, where accuracy and data protection is paramount. Yet, there is no doubt that AI bots can help generate insights, recommendations and predictions based on large volumes of data, enabling lawyers to deliver faster, smarter and more cost-effective solutions to their clients, leaving more time for human skills, such as negotiation and alternative dispute resolution and out of court settlements.

Does the use of computer generated works infringe intellectual property rights?

Yes and no. Intellectual property (IP) gives researchers, inventors, creators, and businesses the confidence to invest their time, energy and money in doing something new. UK business invests more than £130 billion a year in knowledge assets. IP rights protect around £63 billion of this. These assets are vital to the industries that bring us the innovation and products that add value to our lives.⁷

The UK Government is currently consulting in three specific areas:

- (1) copyright protection for computer-generated works (CGWs) without a human author;
- (2) licensing or exceptions to copyright for text and data mining (TDM), which is often significant in AI use and development and
- (3) patent protection for AI-devised inventions.

Computer-generated works (CGWs) are copyright works without a human author. They are currently protected in UK copyright law and the UK Government plans no changes to the law for CGWs. There is no evidence at present that protection for CGWs is harmful because it is believed the use of AI is still in its early stages.

Text and data mining (TDM) means using computational techniques to analyse large amounts of information to identify patterns, trends and other useful information. TDM is used for training AI systems, amongst other uses. It also has uses in research, journalism, marketing, business analytics and by cultural heritage organisations. Although factual data, trends and concepts are not protected by copyright, they are often embedded in copyright works. Data mining systems copy works to extract and analyse the data they contain. Unless permitted

⁷ For further reading see: Ursula Smartt (2023) Media and Entertainment Law 5th ed. Routledge. Chapters 9 and 10.

under licence or an exception, we need to tell our students that making such copies will constitute copyright infringement, a potentially criminal offence or in civil law, there will be financial penalties for people using data mining software indiscriminately.⁸ Several countries have introduced copyright exceptions for TDM. These encourage AI development and other services to locate in such territories, including the EU, Japan and Singapore. TDM may also be fair use under US law, depending on the facts. Our students should also be made aware of ethical questions relating to ‘mining’ personal data and what data should be used for training AI and other TDM uses. Data ethics should be considered in relation to scientific research for example.

For TDM, new copyright and database legislation is planned in the UK. Rights holders will no longer be able to charge for UK licences for TDM and will not be able to contract or opt-out of the exception. The new legal provision may also affect those who have built partial business models around data licensing. However, rights holders will still have safeguards to protect their content. The main safeguard will be the requirement for lawful access. That is, rights holders can choose the platform where they make their works available, including charging for access via subscription or single charge. They will also be able to take measures to ensure the integrity and security of their systems.

For AI-devised inventions there is currently no change planned to UK patent law. It is the present government’s aim to make the UK a global centre for AI innovation, ensuring that the UK’s copyright laws are among the most innovation-friendly in the world. All users of data mining technology will benefit, with rights holders having safeguards to protect their content.

However, the imminent judgment by the UK Supreme Court in the *Thaler*⁹ case may well force the UK Parliament to change the current Patents Act of 1977. The question is: can an AI machine called *Dabus* in the Thaler case, qualify as an ‘inventor’ for the purposes of ss. 7 and 13 of the *Patents Act 1977*? This would make it a ‘person’. Currently copyright and IP laws across the world need human input and authorship to receive intellectual property accreditation.¹⁰

⁸ Set out in section 29A of the Copyright, Designs and Patents Act 1988.

⁹ UKSC hearing on 2 March 2023 on appeal from: *Thaler v Comptroller General of Patents Trade Marks and Designs* [2021] EWCA Civ 1374.

¹⁰ For further discussion see: Alina Skiljic (2021) ‘When art meets technology or vice versa: key challenges at the crossroads of AI-generated artworks and copyright law’. In: *International Review of Intellectual Property and Competition Law*, IIC 2021, 52(10), 1338-1369.

Should future law syllabuses include AI tools?

There are now several ways to build AI systems. Each involves the creation of an algorithm that uses data to model some aspect of the world, and then applies this model to new data in order to make predictions about it. As information processing power has dramatically increased, it has become possible to expand the number of calculations AI models complete to effectively map a set of inputs into a set of outputs. This means that the correlations that AI models identify and use to produce classifications and predictions have also become more complex and less intrinsically understandable to human thinking. It is therefore important to educate our students about artificially created systems and their outputs to detect what is ‘fake’.¹¹

Should we then include broader teaching terms in our (law) syllabuses, including the use of algorithmic art or computer generated musical compositions for AI has become a particularly divisive topic in the music industry.

In April 2023, an AI-produced song called ‘Heart on My Sleeve’ went viral for simulating the voices of Drake and the Weeknd.¹² Universal Media Group successfully petitioned to have the song removed from streaming services – though the track sent shock waves of discourse surrounding ethics and intellectual property through the industry. In June 2023 Paul McCartney told BBC Radio 4’s Today programme how he used AI technology to extricate John Lennon’s voice from a cassette recording of a demo tape of a 1978 Lennon composition titled ‘Now and Then’. The cassette was labelled ‘For Paul’ that Lennon had recorded shortly before his death in 1980.¹³

Conclusion

AI can support innovation and creativity in a range of ways. It can be a tool for scientists, entrepreneurs, artists and lawyers, enabling new human inventions and creations. Some

¹¹ For Rights related to automated decision making including profiling, see the recommendations by the Information Commissioner’s Office (ICO) at: <https://rb.gy/rwa5r>

¹² Abel Makkonen Tesfaye, known professionally as the Weeknd, is a Canadian singer, songwriter, and record producer.

¹³ Source: ‘Paul McCartney says there’s nothing artificial in new Beatles song made using AI’, by Paul Sun, *The Guardian*, 23 June 2023.

believe that AI will soon be inventing and creating things in ways that make it impossible to identify the human intellectual input in the final invention or work.

So, for the time being, we still need a human lawyer to check for accuracy and for our law students to check the sources of their ChatGPT-generated essays before submitting their coursework or examinations. To make our law teaching more future and robot proof we could even test our students whether they recognise AI-generated covers of popular songs, replicating the voices of Harry Styles, Rihanna and Kanye West.