

Is ChatGPT Any Good at Legal Research – and Should We be Wary or Supportive of it?

Abstract: We have all heard horror stories of ChatGPT making information professionals redundant and taking over the world, but just how proficient is it at good old legal research? **Greg Bennett**, a law librarian at BPP, invested in the premium version of the package to put it to the test.

Keywords: Artificial intelligence (AI); legal research; legal technology

Like many of us, I had heard that ChatGPT was fast becoming a very proficient legal mind. It had been able to pass the American Bar exam,¹ and almost passed the Solicitors Qualifying Examination (SQE).² And since it gets higher marks on the Bar exam with every new iteration,³ it seems like it will be only a matter of time before it can likewise pass the SQE. But passing an exam is one thing (I have passed several exams in my life with shockingly poor knowledge of the subject matter), I wanted to test it with some genuine legal research questions: the kind of things we law librarians might be asked to research ourselves.

So, I paid for the premium version of ChatGPT (\$24 per month, including tax), which allowed me to use GPT-4 – the most advanced iteration at the time. Having the premium access also allowed me to install various plug-ins to my account. I added a plug-in called ‘KeyMate.AI Search’, which allows users to include live Google searches in ChatGPT-4’s research. Without such a plug-in, any responses ChatGPT would give to my questions would only be based on its historic Large Language Model (LLM), which only covers content from the internet up to September 2021.

Armed with a correct answer to my first question – “Who won the 2022/23 English Premier League?”⁴ – I knew that my version of ChatGPT had access to the live internet, so I was all set to interrogate it. (NB: you can read the full transcript of every conversation I had with ChatGPT by going to the links in the references.)

RESEARCHING LEGISLATION

I decided to start asking it about legislation. All UK legislation is available for free through legislation.gov.uk, so I thought that this would be ChatGPT’s best chance at successful legal research. (ChatGPT cannot access – at least officially – the internet that sits behind a paywall.⁵) I asked it several questions relating to s.224 of the

Sentencing Act 2020, since this was a section that I knew had been amended twice since September 2021.

My first question simply asked: “what is the text of that section?”⁶ ChatGPT duly responded with the text as enacted, not amended. And it gave the following warning at the end of its reply: “Please note that this is the original version of the section as it was originally enacted. Any changes or amendments to the section after the enactment of the Sentencing Act 2020 are not included in this text.”⁷

I thought I might have to coax the current text of the section out of it, so I asked it a question relating to the actual amendments that had taken place since 2021. I asked it: “Up to how long can a magistrate detain a young offender for an either way offence?”⁸ and it detailed, with complete accuracy, the current correct answer and what had been the correct answer before it changed (and the dates of the changes). So it knew of both amendments to the section since 2021.

I then simply asked it to provide me with the text of the section incorporating any amendments. And it successfully gave me the full text of the section as amended. However, although the wording was all correct, it had got slightly out of sync with the numbering of the subsections (it had successfully added subsection 1A, but then it labelled subsection 2 as subsection 1).⁹ Nevertheless, apart from this slight numbering issue it had come up with the goods.

Section 224 of the Sentencing Act 2020 happens to be a section that has received a lot of discussion (the amendments were quite significant changes). I thought that this discussion might have helped ChatGPT to find the correct answer. So, I decided to ask it some questions relating to a less noteworthy piece of legislation.

I asked it the question: “What is the relevant UK law that covers unfair dismissal during maternity leave?”¹⁰ The answer it gave was not high-quality legal research. It had pulled the information primarily from a charity website that is concerned with working parents.¹¹



Some believe that ChatGPT might one day take over from information professionals, but just how good is it at basic legal research?

(When using GPT-4 and the live internet plug-in, ChatGPT informs you which website the answer comes from). And, although it is a useful resource for non-legal readers, it had no statutory references, and so the ChatGPT answer was very limited.

So I asked it a more specific question: “Can you give me statutory references for unfair dismissal while on maternity leave?”¹² Its answer was very useful, and it pointed me towards appropriate legislation, including section 99 of the Employment Rights Act. At this point, with my question answered, I could have looked up that section myself. But I chose to ask it to “Give me the current [i.e as amended] text of section 99.”¹³

HALLUCINATIONS

ChatGPT’s answer was confusing. It claimed to have taken the text directly from the [legislation.gov.uk](https://www.legislation.gov.uk) page,¹⁴ but, even though that page on [legislation.gov.uk](https://www.legislation.gov.uk) did have a fully updated version of the section (which isn’t always the case for [legislation.gov.uk](https://www.legislation.gov.uk)), the text that ChatGPT gave me was not the same text.¹⁵ It had lots of similar words and phrases, but was nevertheless a very different piece of text. I then compared the ChatGPT version with all the different historic versions of the section on

Westlaw,¹⁶ but I just could not work out why ChatGPT had produced the text it had. ChatGPT’s wording had the ring of accuracy, but in fact it wasn’t correct at all. I could see no sense to it.

If this was indeed incorrect information, this would not be the first time that ChatGPT had produced it in a legal setting. It has produced entirely fictional citations for a litigant in person.¹⁷ And ‘hallucinations’ are occurring regularly. These hallucinations are when ChatGPT makes “mistakes in the generated text that are semantically or syntactically plausible but are in fact incorrect or nonsensical”.¹⁸

Nevertheless, in answer to the previous question I had asked it, ChatGPT had told me what was the relevant legislation, and so perhaps, if I was actually carrying out research (rather than just testing ChatGPT), I would have looked up the section myself. In which case, ChatGPT had been helpful to me. In addition, OpenAI, which owns ChatGPT, states that it is working towards eradicating hallucinations.¹⁹ So, presumably they see a future ChatGPT without hallucinations.

REFERENCING IN OSCOLA

I thought I would test it for its referencing skills at that point, so I asked it to give me an OSCOLA reference for section 99. Its reference was perfect. However, an act is pretty much the easiest thing to reference in OSCOLA. So, I then asked it to reference a book and an article. The references it produced were very poor: no author for the book, no italics for its title; no volume, part or page numbers for the article.²⁰ So, perhaps ChatGPT is not great at referencing in OSCOLA yet. While frustrating for the law school student, not being able to reference accurately in OSCOLA is perhaps not a hugely major failing for ChatGPT in a legal context.

RESEARCHING CASES



What about cases? Up until now I had tested it on things that are in the public domain. Both legislation and OSCOLA’s guidelines are on the free-to-use internet. What if I asked it a question which would normally require access to paid-for resources to answer? I asked it a question that I have often used in teaching: “What are the key cases that determine whether a barge moored on a piece of land is annexed to that land as a fixture, or whether it is merely a chattel?”²¹ In my classes I demonstrate how to find answers to this by looking within practitioner texts, which themselves point me towards relevant cases. And those practitioner texts all sit behind a paywall.

The initial answer I got could best be described as the kind of answer a schoolboy gives when he doesn’t know the specific answer, but he knows a bit about the general topic.²² We used to call it waffle at school. At the end of the answer though, ChatGPT said “I can look up more specific cases related to this topic if you’d like. Would

UK Maternity Dismissal Law

August 18, 2023

 **Anonymous**
What is the relevant UK law that covers unfair dismissal during maternity leave?

 **ChatGPT**
Used unknown plugin 
The UK law that covers unfair dismissal during maternity leave is complex and multifaceted. Here are some key points:

1. **Fair Dismissal:** An employee can be dismissed during pregnancy or family-related leave (maternity, adoption, or shared parental leave), but the dismissal should be fair. The employee should not be dismissed because of their pregnancy or family-related leave. If an employer does not have a potentially fair reason for dismissal or does not use a fair and reasonable procedure when dismissing an employee, the affected employee may have a claim for unfair dismissal if they qualify. To qualify, the employee needs to have been employed for at least 2 years, unless the dismissal is for a reason that makes it "automatically unfair". In automatic unfair dismissal, the employee does not need to have the 2 years' continuous service ¹.
2. **Dismissal During Pregnancy:** A pregnant woman can be fairly dismissed if the main reason for dismissal is unconnected to her pregnancy. If the reason for dismissal is related to her pregnancy, this is likely to amount to unfair dismissal and discrimination ¹.
3. **Dismissal During or Shortly After Family-Related Leave:** It is not necessarily unlawful to dismiss an employee while they are on maternity, paternity, adoption, or shared parental leave. However, the dismissal must be for a fair reason and following a fair procedure. If an employee is dismissed because they took or intended to take family-

A screenshot showing the author's interaction with ChatGPT

you like me to do that?"²³ To which I said, "Yes, please. Could you look up the key UK cases?"

The answer was great, bringing back several key cases. It dealt with a case research question very well indeed. ChatGPT had very successfully passed this harder test. However, I wanted answers to two questions: why didn't it give this answer in the beginning, i.e. why did it first produce the waffle? And where had it found this better information from?

The answer to the first question might simply be that ChatGPT isn't able to discern what information is of a high legal quality, and what is of less high quality, since its LLM is not focused solely on high-quality legal resources (and so it includes many lower-quality resources). In a recent demonstration to BIALL's Academic Special Interest Group, a representative from vLex suggested that they would be making their own AI focus solely on high-quality legal documents for its large language model. This

will apparently produce far more trustworthy legal results than ChatGPT, because poorer quality documents (i.e. the rest of the internet) would be excluded.

The second question then follows on from this. In its answer²⁴ ChatGPT said that it had drawn from a particular article that happens to be behind a paywall.²⁵ Most high-quality legal commentary is held behind a paywall. And, in theory, ChatGPT cannot access information behind a paywall. So how did it access this article?

The article has an open-access pre-print version available on Durham University's repository.²⁶ Had ChatGPT accessed the text of the article there? If so, that raises some important questions about open-access resources that I will discuss later. But if it did access it there, why did it say it had accessed it on the journal's website?

I couldn't help but think that ChatGPT might be accessing things behind a paywall. And this has certainly happened before.²⁷ Seemingly, tech-savvy people have



Anonymous

Can you give me statutory references for unfair dismissal while on maternity leave?



ChatGPT

The statutory references for unfair dismissal while on maternity leave in the UK are primarily found in the Employment Rights Act 1996 and the Equality Act 2010.

1. **Employment Rights Act 1996:** This Act provides the statutory rights for employees in the UK. Specifically, Section 99 of the Employment Rights Act 1996 states that an employee is unfairly dismissed if the reason or principal reason for the dismissal is connected with her pregnancy or childbirth, or if she was exercising her right to maternity leave.
2. **Equality Act 2010:** This Act provides the statutory rights against discrimination, including pregnancy and maternity discrimination. Under this Act, it is unlawful to discriminate against an employee because she is exercising her right to maternity leave.

Please note that this is a simplified explanation and the actual laws are more complex. For a full understanding of your rights and obligations, you should consult the actual text of the laws or seek legal advice.

Interrogating ChatGPT for statutory references

been able to access documents behind a paywall using ChatGPT.²⁸ And if individuals can use ChatGPT to access things behind paywalls, then surely ChatGPT can itself access those same documents. As Emily Dreibelis of *PCMag* says, “We can now add bypassing paywalls to the list of ways AI threatens the existing security and legal measures that govern the web.”²⁹

We law librarians are happy to spend the vast bulk of our budgets on subscriptions to databases like Westlaw and Lexis+ because we know that they contain information that we cannot access in other ways. That information (and to a lesser extent, the way it is organised) is the selling point of those databases. If ChatGPT can access it all, then the threat to Thomson Reuters and LexisNexis is a great one. I assume this threat is why I was recently prompted, on logging in to Lexis+, to click to acknowledge that its “terms do not permit the uploading of [its] content into third party applications, including artificial intelligence technologies such as large language models and generative AI”.

OpenAI has stated that it “want[s] to do right by content owners”³⁰ And so, even though it can access things behind a paywall, it chooses to remove that option from users. Nevertheless, it can clearly access it itself. However, presumably it is easier to access the text of an article behind a paywall than content deep within a database. So, perhaps the threat against Westlaw and Lexis+ is less than it is against a journal website.

SHOULD LIBRARIANS CONTINUE TO CHAMPION OPEN-ACCESS RESOURCES?

I have, like many other librarians, tried very hard throughout my career to encourage academics to make their papers available (perhaps in pre-print form) for free (perhaps via an institutional repository like Durham’s). The dream for many librarians was that at some stage in the future all academic publishing would be made available for free, and we would no longer be in hock to large, powerful and expensive publishers. However, the more that is made available for free, the more powerful tools like ChatGPT will surely become. They can seemingly access paid-for information anyway, but they can only justify its use if free versions of it exist online.

To avoid a ChatGPT (or equivalent) monopoly on information, might we need to reconsider our desire for open-source resources? Might librarians and content providers become unlikely bedfellows? If the role of a law librarian is no longer needed because of tools like ChatGPT, would we be considered Luddites to resist that change? Or would monopoly-avoidance be a good enough reason in itself for librarians to side with content providers? I don’t claim to know the answers to these questions, but I do believe they are questions we should be asking.

CONCLUSIONS

It seems that ChatGPT is already very capable at legal research, albeit with occasional errors. These errors can normally be smoothed out when the user of ChatGPT is someone who knows how to carry out legal research already. ChatGPT will likely continue on its trajectory, and get better at legal research, especially if OpenAI chooses to create a version of ChatGPT that is trained on an LLM that contains only high-quality legal information. In addition, if hallucinations can be eradicated from the research generated by ChatGPT, it will become substantially more reliable. In which case, the need for a knowledgeable user will likely diminish. As a result, this could ultimately lead to a large decrease in law librarian jobs.

One major obstacle in the way of ChatGPT becoming dominant in the field of legal research is that it cannot

access things behind a paywall (at least officially). So, law librarians might, to protect their jobs and, more importantly, to protect against a ChatGPT monopoly, become champions of the pay-for-access legal databases in a way that makes them less supportive of open-access materials than they have historically been. But perhaps law librarians will choose rather to encourage the use of ChatGPT precisely because it brings to the masses the ability to carry out legal research.

At the moment, Westlaw, Lexis+ and other databases contain information that is unavailable to ChatGPT. However, it seems that ChatGPT has the technological capability to access a lot of this information, albeit potentially in a way that infringes content providers' rights. So, the future may well involve legal battles between organisations like OpenAI and Thomson Reuters and LexisNexis.

Endnotes

- ¹ Daniel Martin Katz and others, 'GPT-4 Passes the Bar Exam' (15 March 2023) <<https://papers.ssrn.com/abstract=4389233>> accessed 3 August 2023
- ² Andrew Gilbert, 'Revealed: Could ChatGPT Pass the SQE?' (*Law Society Gazette*, 19 January 2023) <www.lawgazette.co.uk/commentary-and-opinion/revealed-could-chatgpt-pass-the-sqe/5114840.article> accessed 17 July 2023
- ³ Katz and others (n 1)
- ⁴ OpenAI, 'ChatGPT Response to "Who Won the 2022/23 English Premier League?"' (4 July 2023) <<https://chat.openai.com/share/3b9be5db-bd0f-4046-b916-c1ae3a732f59>>
- ⁵ Natalie Byrom, 'AI Risks Deepening Unequal Access to Legal Information' *Financial Times* (London, 17 July 2023) <www.ft.com/content/2aba82c0-a24b-4b5f-82d9-eed72d2b1011> accessed 20 July 2023
- ⁶ OpenAI, 'ChatGPT Response to "What Does the Text of Section 224 of the Sentencing Act 2020 Say?"' (4 July 2023) <<https://chat.openai.com/share/71d130c4-b208-4f7a-a2d6-2b38b9338f40>>
- ⁷ *ibid.*
- ⁸ *ibid.*
- ⁹ *ibid.*
- ¹⁰ OpenAI, 'ChatGPT Response to "What Is the Relevant UK Law That Covers Unfair Dismissal during Maternity Leave?"' (4 July 2023) <<https://chat.openai.com/share/47e3a9bb-9ab7-471d-ab60-ceed79e4b4e8>>.
- ¹¹ 'Dismissal during Pregnancy, or Maternity, Adoption or Shared Parental Leave' (*Working Families*, 3 May 2023) <<https://workingfamilies.org.uk/articles/dismissal-during-pregnancy-maternity-leave-or-soon-after/>> accessed 4 July 2023
- ¹² OpenAI, 'ChatGPT Response to "What Is the Relevant UK Law That Covers Unfair Dismissal during Maternity Leave?"' (n 10)
- ¹³ *ibid.*
- ¹⁴ legislation.gov.uk, 'Employment Rights Act 1996, Section 99' <www.legislation.gov.uk/ukpga/1996/18/section/99> accessed 18 August 2023
- ¹⁵ OpenAI, 'ChatGPT Response to "What Is the Relevant UK Law That Covers Unfair Dismissal during Maternity Leave?"' (n 10)
- ¹⁶ Westlaw, 'Employment Rights Act 1996, Section 99' <<https://uk.westlaw.com/Document/I51C5301059BD11DB899B8284D0D7430E/View/FullText.html>>
- ¹⁷ John Hyde 29 May 2023, 'LiP Presents False Citations to Court after Asking ChatGPT' (*Law Society Gazette*) <www.lawgazette.co.uk/news/lip-presents-false-citations-to-court-after-asking-chatgpt/5116143.article> accessed 17 July 2023
- ¹⁸ Craig Smith, 'Hallucinations Could Blunt ChatGPT's Success - IEEE Spectrum' (13 March 2023) <<https://spectrum.ieee.org/ai-hallucination>> accessed 11 September 2023
- ¹⁹ OpenAI, 'GPT-4' <<https://openai.com/gpt-4>> accessed 11 September 2023
- ²⁰ OpenAI, 'ChatGPT Response to "What Is the Relevant UK Law That Covers Unfair Dismissal during Maternity Leave?"' (n 10)
- ²¹ OpenAI, 'ChatGPT Response to "What Are the Key Cases That Determine Whether a Barge Moored on a Piece of Land Is Annexed to That Land as a Fixture, or Whether It Is Merely a Chattel?"' (4 July 2023) <<https://chat.openai.com/share/a61c448d-984d-4fd6-91ea-0819f4bc129d>>
- ²² *ibid.*
- ²³ *ibid.*
- ²⁴ *ibid.*

²⁵ Chris Bevan, 'The Law of Fixtures and Chattels: Recalibration, Rationalisation and Reform' (2022) 42 *Legal Studies* 358

²⁶ Chris Bevan, 'The Law of Fixtures and Chattels: Recalibration, Rationalisation and Reform - Pre-Print Version' (*Durham Research Online*, 2022) <<https://durham-repository.worktribe.com/preview/1220717/34765.pdf>> accessed 18 August 2023

²⁷ Emily Dreibelbis, "'Browse With Bing" Disabled on ChatGPT Plus Because It Bypassed Paywalls' (*PCMag UK*, 5 July 2023) <<https://uk.pcmag.com/ai/147625/browse-with-bing-disabled-on-chatgpt-plus-because-it-bypassed-paywalls>> accessed 24 July 2023

²⁸ Maria Diaz, 'ChatGPT Browsing Feature Deactivated Only a Week after Roll out - Here's Why' (*ZD Net*, 5 July 2023) <www.zdnet.com/article/chatgpt-browsing-feature-paused-after-users-bypassed-paywalls> accessed 24 July 2023

²⁹ Dreibelbis (n 27)

³⁰ OpenAI, 'We've Learned That ChatGPT's "Browse" Beta Can Occasionally Display Content in Ways We Don't Want...' <<https://twitter.com/OpenAI/status/1676072388436594688>> accessed 31 August 2023

Biography

Greg Bennett retrained to become a librarian after a brief career in pensions in the City of London. His trainee year was at the Institute of Advanced Legal Studies, where he discovered a love for specifically law librarianship. He came back to work at IALS after his MA, and then worked at the law firm Slaughter and May before returning to the academic sector at the University of Essex. From 2018 to 2021 he worked for Goldsmiths, University of London, and during that time he completed an LLM by distance learning. He now works at BPP in London, where, in addition to being a law librarian, he is studying the LPC.