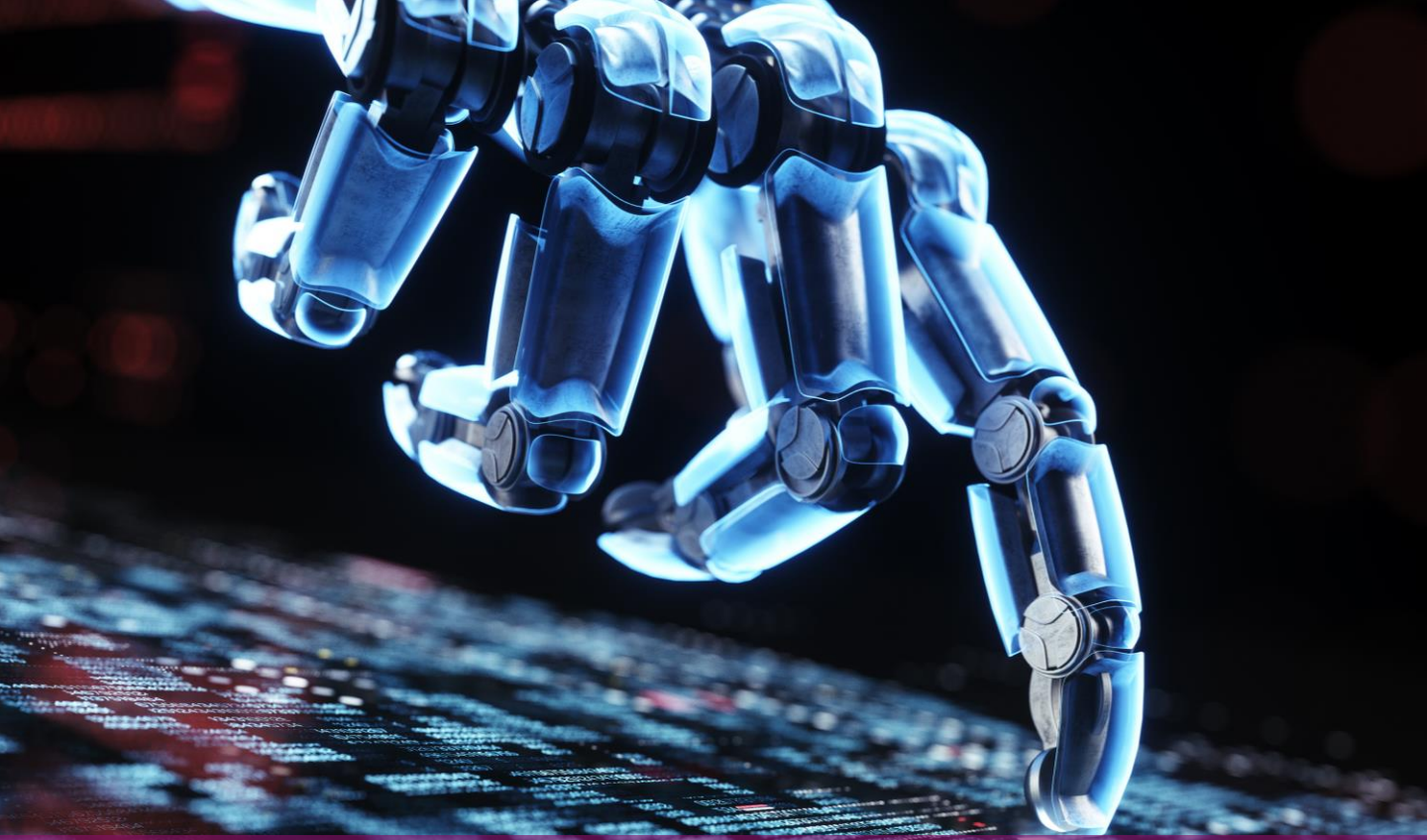


ChatGPT: Will it pass its probation?

The risks and issues in using AI chatbot auto-generative systems

Introduction

If you had any doubts about the disruptive nature of artificial intelligence, the arrival of ChatGPT should have well and truly put them to bed. The language processing tool has been a runaway hit since its launch by OpenAI last November, and has sparked conversations about how businesses and institutions could be impacted by its increased use.

The premise of ChatGPT is as simple as it is extraordinary. Trained on a massive amount of text data, it's capable of understanding and generating human-like prose. It answers questions and can assist on tasks like composing essays, job applications or letters.

But despite its powerful capabilities, the use of ChatGPT and other generative AI pose several legal and practical risks that should be considered before allowing its use within your organisation. We discuss some of those risks in this article.

STEPHENSON HARWOOD

Accuracy and bias

Like any machine learning model, ChatGPT is subject to certain accuracy and bias risks. In particular, if the training data contains errors, inaccuracies, or biases, these will be reflected in the model's responses. For example, if the training data contains mostly examples of one type of person or group, the model may not be able to generate accurate or appropriate responses for other types of people or groups (and may perpetuate existing stereotypes and discrimination). The output may also be skewed if the model is trained on limited or narrow data sets, or if the algorithm used to train it is biased.

Users of ChatGPT have no control over the data sets used to train it, or the algorithms it runs on. This means users are completely reliant on how OpenAI developed and trained the AI, but don't have any real recourse should it produce inaccurate or biased outputs (more on this later).

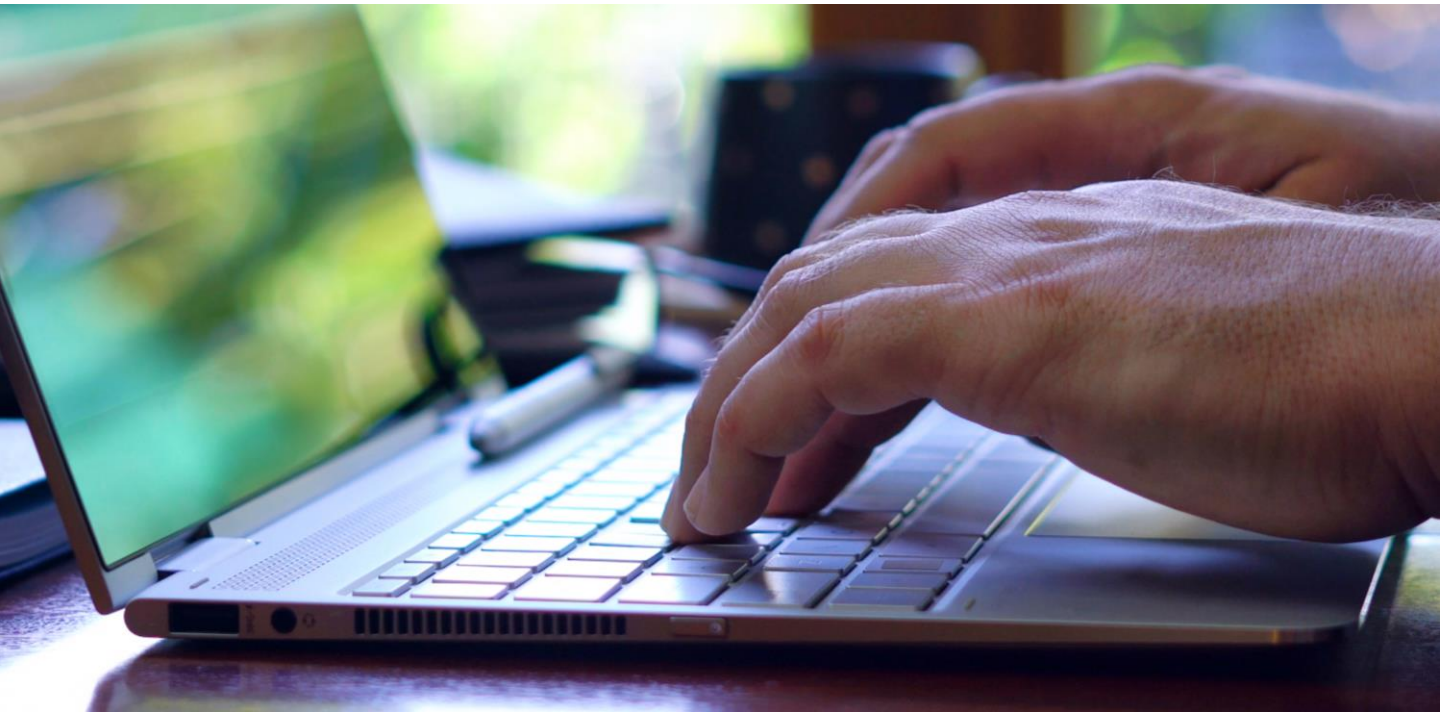
It's also worth keeping in mind that ChatGPT's training data is a snapshot of the internet at a certain point in time (currently in 2021!), and the output will be based on statistical patterns on the data it was trained on and not on a deep understanding of the subject.

So, whilst it is a hugely powerful tool, it should not in our opinion be relied on as the oracle of truth and accuracy or a pillar of good standing.

Ownership of output

Since 1988 UK law has allowed copyright to subsist in machine-generated content. ChatGPT's terms of use state that all output generated by the model for a user is owned by the user. As such, and at least under UK law, the user will own the copyright in the output.

This means that the user will be able to prevent the copying of the output. While on the surface this might seem simple enough, those rights won't stop the output infringing the rights of others.



Infringement risks

Just because content is openly available doesn't mean it isn't protected by copyright or other rights. With systems like ChatGPT, there’s a risk that the content generated could infringe the copyright and/or database right of the owner of rights in the materials it was trained on.

ChatGPT is, at its core, trained on very large datasets of text and other materials from the internet — although the precise source of such training materials is unclear. These datasets will include materials subject to copyright protection. As a result, the output generated by ChatGPT might be similar or even identical to works already in existence — giving rise to a real risk that the use of the output, without permission, could constitute copyright infringement. As well as legal action being taken against OpenAI, it could also be taken against individual users, too.

This risk of output infringing the rights of a third party, or at least of action being taken, is not just theoretical.

In January 2023 proceedings were started in the High Court in London by Getty Images against Stability AI, an AI business. Getty Images claims that Stability AI has copied millions of images from its database to train its image generation model. Commentators have pointed to the fact that Stability AI's images have a tendency to include Getty Images' watermark to suggest that its images have been used to train Stability AI's model.

It isn't clear whether ChatGPT's model will alter the original works sufficiently to avoid infringement, but the Stability AI case raises a real risk to the use of AI trained on openly available works and perhaps even an existential threat to some models completely. Also of note is the government's statement on 2 February 2023 that a proposed general text and data mining exception to copyright and database right infringement will not be proceeding, suggesting that no end is in sight in the UK to the risk of infringement in this area.

This infringement risk may make some users, particularly enterprise users, nervous to deploy the use of ChatGPT or similar AI in its business. As noted above, the infringement risk extends to the user and not just the platform.

Data protection risks

Because ChatGPT is trained on a large dataset of text from the internet, this may include personal information of individuals. When using ChatGPT, there is a risk that personal information may be inadvertently processed, which could be in violation of data protection laws. However, users have no control over this data, or any means to ensure that it was collected and processed legitimately.

According to ChatGPT (when we raised the question), it does not have the ability to check if it has the right to process personal data in its training data, and OpenAI's terms of use put the responsibility on the user to ensure that it has the right to process any personal data.

It is important to note that, if you are using it to process personal data, users have to opt in to execute a Data Processing Addendum to cover the requirements under applicable data protection laws. This doesn’t apply automatically.

If users are inputting personal data into ChatGPT, they will be responsible for their processing obligations as a data controller, and need to be transparent to individuals about their use of ChatGPT. The OpenAI terms also provide that, unless you specifically opt out, users "agree and instruct" OpenAI to use any input data and the output it provides to "develop and improve the Services". This places the risk on the user to ensure that such transfer to OpenAI is permitted under data protection law.

Liability

The terms of use for ChatGPT, state that OpenAI is not liable for any damages arising from the use of the model, and that the model is provided "as is" without any warranty that the output will be accurate or capable of being used for a particular purpose. Furthermore, OpenAI's liability is capped at the greater of $100 or the amount paid for the service in the previous 12 months. So, if any of the risks highlighted in this article play out, there is little or no recourse against OpenAI.

So, while we acknowledge the advantages and economies that can be obtained from using ChatGPT, we advise that any output is treated with a degree of caution.

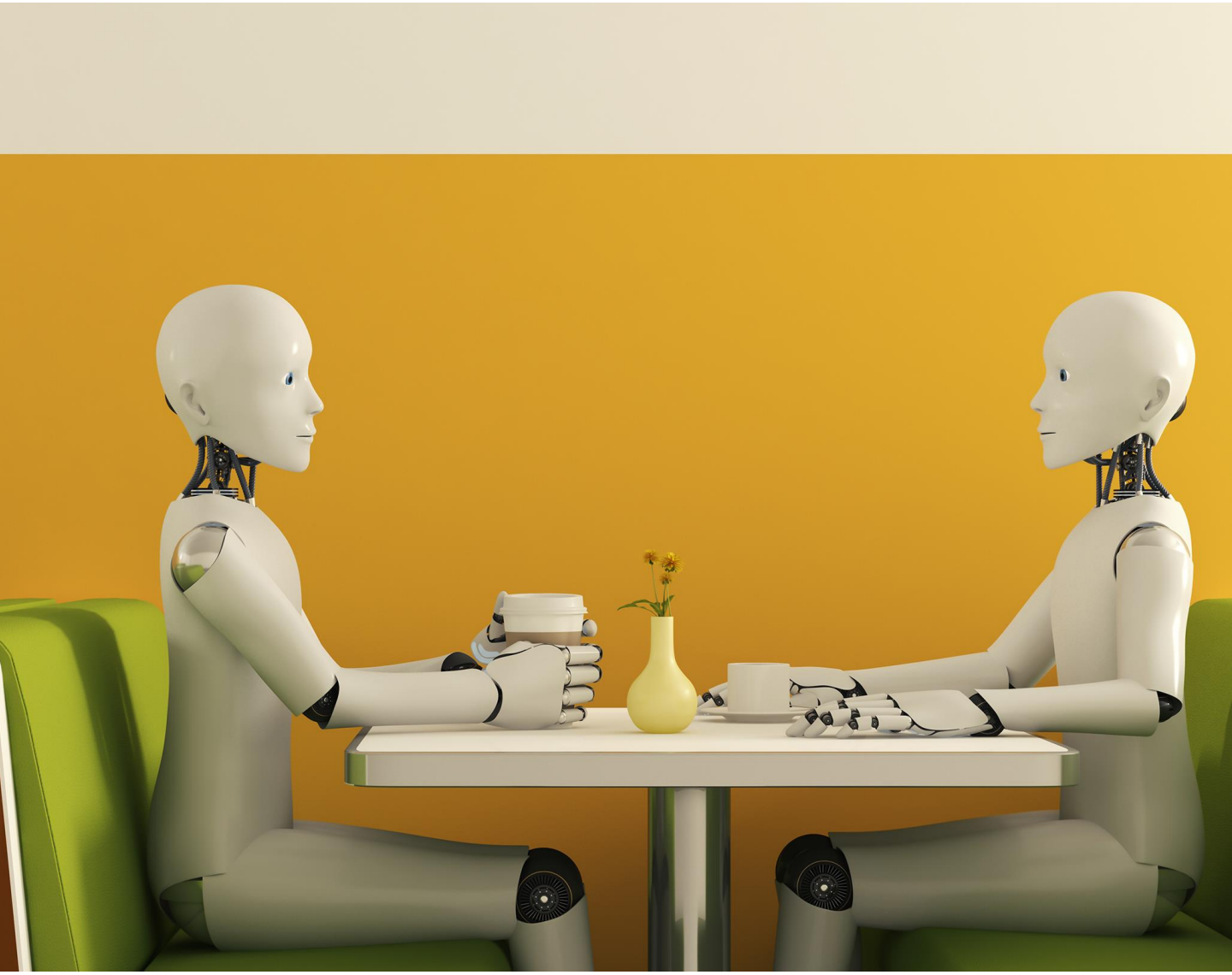
Use within the workplace

Organisations should consider the above highlighted risks when allowing their employees to use ChatGPT, and should put in place relevant controls to ensure that the use of ChatGPT is safe and compliant.

In particular, organisations should establish policies and guidelines for the use of ChatGPT, its permitted and prohibited use cases, and how any output should be subject to human "checks and balances". They should also establish a process for employees to report any concerns or issues related to its use. Alternatively, an employer may want to ban the use of ChatGPT for certain roles or types of work product where they do not feel it is suitable.

Putting aside the legal implications of using the tool to generate, for example, a speech or article that you intend to pass as your own, it is worth bearing in mind that there is a real risk that the same or similar content could be generated for another user! At the very least, that could be embarrassing and cause reputational damage.

The sooner that organisations grapple with these potential ramifications the better, thus enabling them to take a pro-active rather than reactive stance to issues that might occur. As with any new joiner, we would at least expect organisations to implement some form of probationary checks and balances on ChatGPT's performance, before embedding it into the business.



Innovative but not risk-free

While ChatGPT is an extraordinary innovation and has done a brilliant job at bringing AI to the forefront of public consciousness, it would be remiss to abandon all safeguarding principles when using it. Being mindful of its risks and setting up systems to protect against them is the best way to enjoy all the benefits of the AI tool without falling victim to its shortcomings.

But don’t just take our word for it, here it is from the horse's mouth:

“ChatGPT may be the next frontier in technology, but it is not without its legal perils. As always, knowledge is power and understanding the potential risks and liabilities associated with this technology is essential for any forward-thinking business or individual. As the famous legal maxim goes, 'ignorantia juris non excusat' - ignorance of the law is no excuse.” — ChatGPT

STEPHENSON HARWOOD

Key Contacts

Technology

Simon Bollans
Partner

T: +44 20 7809 2668
E: simon.bollans@shlegal.com

Nic McMaster
Associate

T: +44 20 7809 2661
M: +44 7920 431 106
E: nic.mcmaster@shlegal.com

Intellectual property

Rob Jacob
Partner

T: +44 20 7809 2072
M: +44 7825 601 925
E: rob.jacob@shlegal.com

Joshua Cunnington
Managing associate

T: +44 20 7809 2256
E: joshua.cunnington@shlegal.com

Employment

Anne Pritam
Partner

T: +44 20 7809 2925
M: +44 7946 647 238
E: anne.pritam@shlegal.com

Leanne Raven
Senior knowledge development lawyer

T: +44 20 7809 2560
M: +44 7827 353 108
E: leanne.raven@shlegal.com

For further details on our tech expertise, insights, and credentials please visit our technology hub can you embed a link to the tech hub: https://www.shlegal-technology.com/

STEPHENSON HARWOOD