AI Hallucinations, Bias & Inaccuracies – Why Your Business Needs AI Policies

This is the final article of our 3-part artificial intelligence (AI) safety series. Where we focus on a critical component often overlooked in the discussion of AI, the implementation of AI policies.

Our first article covers the Bletchley Park AI Safety Summit, its global discussions surrounding AI risks, and international cooperation to mitigate these. Our second article relates to cyber security and AI, specifically covering the consequences of cyber security breaches, relevant guidance and practical steps businesses should take to protect themselves in this landscape.

In keeping with the theme of AI safety, this article will now explore the risks relating to approved and frequent unapproved use of AI by employees within a company and the consequential need for AI policies to be implemented.

Introduction

50% of adults have used AI, and the uptake of this technology is only increasing, including within businesses.[1] However, it has been found that 70% of employees who use ChatGPT do not tell their boss that they are using the technology.[2]

The use of AI, where it has been formally implemented by a company, can carry risks which require careful, intuitive system control. This is even more important when a company has yet to implement AI formally, but employees use AI as a tool to carry out their work.

This article discusses this further, outlining exactly what issues are relevant here and identifying the critical role policies play in mitigating risk.

Use of AI within a company

A company may use AI where its use has already been pre-approved by its board or relevant team.

Such use may arise within the following activities of a company's business:

- cybersecurity;

- customer service;

- fraud management;

- inventory management;

- customer relationships management.

It is common for employees to use AI technology to assist them with their work, even where it has yet to be officially approved and formally adopted by their employer.

Within this context, employees may be using AI in the course of their work to complete the following tasks:

- carry out research;

- draft documents, emails and letters; and

- create presentations, blog posts and adverts.

Important considerations & risks

Whilst the implementation and use of AI is advantageous to businesses in that it can increase the speed of business processes, level up operations and improve the efficiency of employees enabling them to focus on more complex work, AI does come with associated risks. This is because this technology is not governed by specific AI laws in the UK, and the position continues to develop rapidly.

Two of these risks include AI hallucinations and AI bias, which are particularly relevant to a business's use of AI, whether approved or not. This article illustrates the risks below and why companies need to introduce AI policies.

AI hallucinations

AI hallucinations occur when a generative AI model outputs information as though it is correct but is, in fact, incorrect.

This happens as data used to train the AI may be low in quality, biased or lack context in its presentation or application. When the AI then assembles its output in response to a user's request, it produces a coherent piece of text, music, artwork, etc., by applying its learning of patterns and probabilities. It does not understand the underlying or general reason behind why it has produced this output.

AI chatbots, such as ChatGPT, may then respond to users' requests with information that appears to be factual and correct but which, in reality, is inaccurate and cannot be relied on to be correct (a hallucination).

AI bias

As AI operates in a statistical and data-driven way, this can lead to it producing biased work. In the context of a company, this can mean that AI systems that deal with people-related data could produce discriminatory work; this may be particularly relevant to an employer's HR department.

The need for AI policies

The dangers associated with a company's use of AI, where AI hallucinations and AI bias are relevant risks, could include giving advice to clients that is incorrect or causing a discriminatory rejection of candidates within the company's recruitment process.

One of the best ways these risks posed by AI hallucinations and AI bias can be avoided is through implementing policies. This is because the purpose of policies is to set standards, rules, and expectations for company employees to comply with.

By introducing AI policies, a company can make sure that there is a common approach concerning the implementation and use of AI, whereby risks will have been formally and appropriately evaluated, and the rules and standards under a policy will, as a result, be most effective in mitigating risk.

A company's AI policy might state that employees are only permitted to use AI chatbots, such as ChatGPT, for limited purposes, such as producing the formatting design of documents, such as presentation slides, letters, or web articles. The policy could then state that employees are not permitted to use it to conduct research, nor produce advice to a client without careful checking, emphasising that employees are encouraged to be critical of the work produced by AI.

The benefit of a policy providing rules such as these is that the company can still use AI to speed up processes and, therefore, improve employee efficiency while also mitigating the risk of providing clients with incorrect advice or the risk posed by applying a company's own internal analytics.

Our Corporate & Commercial team and Employment & HR team can discuss these AI risks with you further. They can assist your company in preparing and implementing AI policies to mitigate these risks.

Contact us today by emailing online.enquiries@la-law.com or calling 01202 786188.

[1] Office for National Statistics (ONS), 'Understanding AI uptake and sentiment among people and businesses in the UK: June 2023', 16 June 2023

[2] Fishbowl, '70% Of Workers Using ChatGPT At Work Are Not Telling Their Boss; Overall Usage Among Professionals Jumps To 43%', 1 February 2023

Tel: 0344 967 0793 | Email: info@LA-Law.com | www.lesteraldridge.com