CLIENT ALERTS & INSIGHTS

ChatGPT Generates More than Data Outputs; Data Security and Privacy Concerns Grow as Artificial Intelligence Technology Rapidly Advances

MAY 17, 2023

Authors: Megan Faska

One of the primary concerns surrounding ChatGPT and AI is the security of the data used to train and operate these machines and the potential privacy implications if the data is mishandled or intercepted by malicious actors.

With an abundance of news headlines heralding ChatGPT and other similar AI machines as the "new era of technology and intelligence," AI *appears* as a seemingly new development. Not so- AI began in the 1950s. The work of modern computer scientists, mathematicians, and technicians rapidly improved and grew the AI industry as experts strove to develop a machine comparable to the human brain. Despite their efforts, however, support for the field ebbed and flowed for over half a century, until another surge of interest sprouted in 2012. This surge persisted through current day and propelled breakthroughs in AI that many thought were elusive.

Today, one AI has come to forefront-ChatGPT is a free chatbot trained by OpenAI that was released to the public in November 2022. ChatGPT stands for "generative pretrained transformer." It is an AI chatbot designed and trained by large amounts of textual (documents and words) and reinforcement (human feedback) learning data to hold "natural" conversations. To use ChatGPT, a person simply enters a written prompt or question and ChatGPT responds in a sentence or more. It appears as if two humans are chatting online.

Since its public release, there have been various versions of ChatGPT as the AI continues to learn and improve. The most recent version, ChatGPT-4, was officially announced on March 13, 2023, but is currently only available in the ChatGPT Plus paid subscription. ChatGPT-4 supposedly challenges human creativity by generating collegiate level essays, moving poetry, visual art, and even new languages. But lurking in the otherwise whimsical landscape of possibility lies troubling data security concerns and legal uncertainty.

What is Generative AI?

There are four main types of AI: (1) Reactive; (2) Limited Memory/Generative; (3) Theory of Mind; and (4) Self-Aware. Reactive and Limited Memory/Generative AI fall under the "Narrow AI" umbrella, which is the type of AI that exists today. These narrow AI machines can only perform what they are programmed to perform, which involves a single or "narrow" task. Theory of Mind and Self-Aware AI will have the ability to learn, perceive, and function like human beings, but are currently unrealized. For the purposes of this article, it is important to understand the first two types of AI:

1. **Reactive AI:** The first AI algorithms were purely reactional (hence the name). Reactive AI includes machines that have no memory-based functionality, meaning they cannot learn from previously gained experiences or prior data. Mathematicians created these models to digest large amounts of seemingly non-sensical data to produce statistical compilations. Reactive AI is only useful for responding to a limited combination of inputs. Unlike humans that rely on past experiences to make decisions, Reactive AI does not have a storage bank of scenarios to formulate its outputs. Therefore, developers pushed on to create the next level of AI-Limited Memory/Generative AI.
2. **Limited Memory/Generative AI:** Limited Memory/Generative AI algorithms were designed to imitate human brain receptors and connectors so that the AI machines could absorb and retain "training data" to improve its outputs over time. Nearly all present-day AI machines use and store an insurmountable amount of training data in their memory for future problem solving (think chat boxes, Siri and Alexa, Netflix recommendations, etc.).

Although far advanced from their predecessors, ChatGPT and other market competitors (Amazon Bedrock, Google's Bard AI, DeepMind's Chinchilla AI) are still considered Limited Memory/Generative AI machines. But, as developers keep improving the AI and releasing new generations, the Limited Memory/Generative AI capabilities feel less and less "narrow."

Data Security and Privacy Concerns

One of the primary concerns surrounding ChatGPT and AI is the security of the data used to train and operate these machines and the potential privacy implications if such data is mistreated. In addition to the data that is manually input to the system, ChatGPT scrapes data from the web. Data scraping is a process of importing data from other websites without permissions or consent. This scraped data is then accessible by any user if the data is responsive for the user's query or prompt.

By its nature, ChatGPT retains substantial quantities of personally identifiable information and other sensitive information, including but not limited to a user's browsing history, social media activity, credit scores, medical records, trade secrets and financial data. Therefore, the use of ChatGPT can lead to: (1) disclosure of sensitive information if ChatGPT deems the data responsive to a user's question or request; (2) sensitive business information being exposed to competitors; and (3) even reputational damage if the information obtained through ChatGPT's data scraping is inaccurate.

Additionally, due to the sheer volume of its data storage, ChatGPT itself becomes a valuable target for cybercriminals, which, if not adequately protected, could lead to serious data breaches and violations of privacy. A cybercriminal who gains unauthorized access to ChatGPT's data pool could use it to generate more customized and convincing phishing messages or launch automated attacks against vulnerable scam targets. What's more, while the AI chatbot uses natural language to communicate as if it were another person, it can also create computer code - making it easier for bad actors with limited knowledge or coding skills to carry out computer network attacks. ChatGPT does have some general safeguards in place to prevent the AI from being used for an outright malicious purpose. For instance, the AI will not write a code for a user who prompts the AI to "write a code for a ransomware application." But, like anything, cybercriminals are quickly learning how to navigate around ChatGPT's limited safeguards - the simplest method: phrase the prompt a different way and avoid malicious trigger words.

But Who is Actually Responsible?

Chat GPT's Terms of Use does nothing to alleviate the potential data security and privacy concerns; rather, it bolsters these legal concerns for its users. Under the Terms of Use, users grant ChatGPT a license to use any input content for the purpose of training and improving ChatGPT. Based on the nature of Generative AI, such a license to use input content for training and learning does not feel like an overreach-if anything, it is bottom-line necessary for the AI to improve over time. However, if the content is not properly secured or stored (or is unknowingly provided without knowledge of this license), the content is vulnerable to unintended access and use, and data breach and cyberattacks (as discussed above).

Moreover, Section 3 of the Terms of Use states that users are solely and legally responsible for any content they input into ChatGPT. Therefore, although users do not have control over the use and storage of their input data (or even full knowledge of how it will be used), users will still be liable if any content input in the system violates any third-party rights, including intellectual property or privacy rights.

What's more, ChatGPT's indemnification provision places the burden on the *user* to assume liability for any claims or damages (including settlement costs and legal fees) that may arise from their use of the chatbot, even if the user is not at fault. This means that if OpenAI/ChatGPT gets sued for an individual's use of the chatbot, then the individual user is financially responsible for those costs and expenses. Whether such an indemnification provision will be enforceable remains to be seen (for example, where the user is not at fault), but it is clear that ChatGPT's Terms of Use places significant legal and financial burdens on its users in broad strokes.

Key Takeaways

While there are many potential benefits to using ChatGPT and other AI machines, it is important to proceed with caution as Generative AI becomes increasingly incorporated into our business, societal, and personal lives. Companies must be aware of the data security and privacy risks associated with using such AI machines as well as the legal liability users assume simply by using the system. To minimize and mitigate the risks discussed above, businesses must take a proactive approach and invest in solid, robust data and cybersecurity measures, including encryption, access controls (like dual authentications), and monitoring tools to protect their data. Additionally, businesses should provide clear privacy policies that explain how data will be used and protected, as well as obtain necessary consents and approvals in accordance with data privacy laws including the GDPR, CCPA as amended by the CPRA, and other state data privacy laws.

Although Limited Memory/Generative AI has been in existence for decades, the rapid advancements this field has seen this past year is nothing short of extraordinary. But this vast technological advancement also breeds legal uncertainty and risks that can expect to cultivate a new regulatory landscape in the coming months and years. For now, by putting the proper protections in place, leading organizations in this field can harness the many benefits of ChatGPT and other Generative AI machines while protecting themselves against potential malicious actors and mitigating legal liability.

As the regulatory climate surrounding AI machines like ChatGPT continues to grow and take shape, the Benesch Data Protection and Privacy team is committed to staying at the forefront

of knowledge and experience to assist our clients in mitigation and compliance efforts. We are available to assist you with any mitigation and compliance needs.

Megan Kern at mkern@beneschlaw.com or 216.363.4615.