Generative AI Has Its Risks, but the Sky Isn't Falling

The threat organizations face with GenAI is not new, but it could speed how quickly private data reaches a wider audience.

Generative artificial intelligence (GenAI) and large language models (LLMs) are the disruptive technologies *du jour*, redefining how enterprises do business and spurring the debate on just how much AI will change the way civilization interacts with computers in the future. The hyperbole is reaching epic proportions as social scientists and pundits debate the End Times facing civilization due to smarter and potentially proactive computing. Perhaps some perspective is in order.

A recent report from Israeli venture firm Team8, "Generative AI and ChatGPT Enterprise Risk," addresses some of the realistic technical, compliance, and legal risks GenAI and LLMs place on corporate boards, C-suites, and cybersecurity personnel. The report underscored the potential operational and regulatory vulnerabilities of GenAI, but cautioned that some concerns might be premature. One concern — that exposing private data submitted to a GenAI application such as ChatGPT might make that data available to others in near real time — is discredited in the report.

"As of this writing, Large Language Models (LLMs) cannot update themselves in real-time and therefore cannot return one's inputs to another's response, effectively debunking this concern," the report states. "However, this is not necessarily true for the training of future versions of these models."

The report identifies several potential threats by risk category. Among the high risks are:

- Data privacy and confidentiality of nonpublic enterprise and private data.
- Enterprise, software-as-a-service (SaaS), and third-party security of nonpublic and enterprise data.
- AI behavioral vulnerabilities, such as prompt interjection, of enterprise data.
- Legal and regulatory compliance.

Among the threats that fall into the medium risk category are:

- Threat actor evolution for attacks, such as phishing, fraud, and social engineering.
- Copyright and ownership vulnerabilities leading to an organization's legal exposure.
- Insecure code generation.
- Bias and discrimination.
- Trust and corporate reputation.

"The CISO is in a position to have the technical knowledge to support processes not necessarily under their umbrella, which might affect their role," says Gadi Evron, CISO-in-residence at Team8 and one of the report's authors. "Some interpretations of upcoming European Union regulation may push the CISO into a position of responsibility when it comes to AI. This may elevate the CISO to a position of responsibility where they are 'ambassadors of trust,' and this is a positive thing for the CISO role."

Chris Hetner, cybersecurity advisor at Nasdaq and chair of Panzura's Customer Security Advisory Council, says an initial risk assessment would identify potential issues with who has access, what can be done, and how the technology will interact with existing applications and data stores.

"You need to determine who has access to the platform, what level of data and code are they going to introduce, [and does] that data and code introduce any proprietary exposure to the enterprise," he notes. "Once those decisions are made, there's a process to proceed forward."

The threat organizations face with GenAI is not new, but it could speed how quickly private data reaches a wider audience.

"When it comes to security, it is clear that most companies are in much worse condition to mitigate the risks of their corporate and customer being stolen or leaked than they were just six months ago," opines Richard Bird, chief security officer at Traceable AI. "If we're being intellectually and historically honest with ourselves, the vast majority of companies were already struggling to keep that same data safe before the rise of generative AI.

"The ease of use and access that employees are already taking advantage of with AI technologies with little to no security controls is already showing the increased risk to companies."

The Human Element

Bird takes a pragmatic approach to GenAI, adding that companies are going to move fast and not wait on compliance demands to protect their data, customers, supply chains, and technologies. Users have shown "a complete lack of restraint combined with no awareness of the unintended security consequences of using AI," he notes. "This toxic combination is what companies must work quickly to address. AI isn't the key threat here. Human behavior is."

One issue that has yet to be fully analyzed is how users interact with GenAI based on their existing habits and experience. Andrew Obadiaru, CISO for Cobalt Labs, notes that iPhone users, for example, already have native experience with AI through Siri and thus will adapt quicker than users who do not have that experience. Like Bird, Obadiaru thinks those habits might make those users more susceptible to misusing the applications by inputting data that should not get out of an organization's direct control.

"The concerns are, 'What are the additional risks?' Everyone has the ability to tap into [GenAI technology] without necessarily going through a security review," he says. "And you can just do that on their personal device."

If employees use personal devices outside of the IT department's control to conduct business, or if employees use GenAI as they use Siri or similar applications, this could pose a security risk, Obadiaru adds. Using GenAI like a personal digital assistant potentially could put confidential data at risk.

Network Risks

Sagar Samtani, assistant professor in the Data Science and Artificial Intelligence Lab at the Indiana University Kelley School of Business, cautions that AI models are extensively shared via the open source software landscape. These models contain significant quantities of vulnerabilities within them, some of which CISOs need to be aware of.

"These vulnerabilities place an impetus on organizations to understand what models they are using that are open source, what vulnerabilities they contain, and how their software development workflows should be updated to reflect these vulnerabilities," Samtani says.

Asset management is a critical aspect to any strong cybersecurity program, he adds.

"[It's] not the exciting answer, but an essential ingredient," Samtani says. "Automated tools for detecting and categorizing data and assets can play a pivotal role in mapping out corporate networks. ... Generative AI could help provide layouts of corporate networks for possible asset management and vulnerability management tasks. Creating inventory lists, priority lists, vulnerability management strategies, [and] incident response plans can all be more [easily] done with LLMs in particular."