ECHELON RISK + CYBER

The Security Risks of Generative AI

Introduction

In recent years, artificial intelligence tools such as ChatGPT have skyrocketed in popularity. Companies across the world are using AI tools to write code, draft emails, brainstorm, analyze data, and many other tasks.

Unfortunately, many companies are not taking a step back and truly asking if the use of AI in their company is worth the risks it introduces.

In certain instances, generative AI has the power to maximize employee efficiency, saving money and allowing employees to avoid tedious and monotonous tasks. According to a study conducted by Stanford and MIT, the use of generative AI increased worker productivity by an average of 14%.

It should be noted for novice and low-skilled workers, artificial intelligence increased productivity by a staggering 34%, but made minimal impact for highly skilled workers. Unfortunately, these tools come with some risks and complications that companies should be aware of.

Security Risks for Generative AI

Personally Identifiable Information (PII)

In recent years, artificial intelligence tools such as ChatGPT have skyrocketed in popularity. Companies across the world are using AI tools to write code, draft emails, brainstorm, analyze data, and many other tasks.

Unfortunately, many companies are not taking a step back and truly asking if the use of AI in their company is worth the risks it introduces.

In certain instances, generative AI has the power to maximize employee efficiency, saving money and allowing employees to avoid tedious and monotonous tasks. According to a study conducted by Stanford and MIT, the use of generative AI increased worker productivity by an average of 14%.

It should be noted for novice and low-skilled workers, artificial intelligence increased productivity by a staggering 34%, but made minimal impact for highly skilled workers. Unfortunately, these tools come with some risks and complications that companies should be aware of.

Reputational Damage

Reputational damage is another major risk. Output from AI chatbots and tools can often be incorrect, citing fabricated sources or even producing racist or sexist language. Many AI users are over-reliant on it, and don't properly verify and fact-check the output. If this output is going to be put in a public document, or seen by clients, this can cause irreversible reputational damage that could easily have been prevented.

Managing the Security Risks of AI Tools

Trying to manage the risks of new technology such as AI can be challenging, especially when there is such a large push to utilize these tools. In April of this year, NIST released the AI Risk Management Framework (RMF).

Part 1 of this framework covers AI risks, trustworthiness, as well as how to address risks, impacts, and harm from AI.

Part 2 of the framework defines the AI RMF Core which is broken up into 4 sections (Govern, Map, Measure, and Manage). The AI RMF Core provides actions to help communicate, understand, and perform activities to help manage AI risk. Utilizing this framework will help any organization analyze risk, manage AI, and protect themselves from harm.

Tips for Improving Security when Using Generative AI

So how should businesses decide whether AI should be used in their environment? Currently, over 75% of companies are considering banning the use of generative AI, citing data security and company reputation as the two largest concerns. Some industries benefit greatly from its use, especially companies with a large amount of novice and low-skilled workers, and a lack of intellectual property or PII.

If using AI does make sense, outline clear guidelines for employees and train them on how to safely use these tools. This can be done by issuing data protection training, enforcing an acceptable use policy, or even through introducing a standalone artificial intelligence policy. If companies want a boost in efficiency from AI, with fewer risks, there are also some paid solutions to consider that address these problems. Many paid generative AI tools encrypt data-at-rest, and some even store data in a private cloud to ensure that it cannot be accessed by the public. Many enterprise-level tools reduce the risk of data breaches, but it has proved difficult to verify and control the output of these large language models.

75% Companies are considering banning generative AI

Guidelines for Reducing Risk when Using AI in Your Environment

Ensure that your work force is trained on the risks of generative AI. Implement required training through your organization's training platform as soon as possible.

Outline clear language in your Acceptable Use Policy (or create a separate AI policy) documenting the limitations and proper use in your company. This language should restrict AI use with confidential, sensitive, or proprietary information. It should also state that all approved AI tools must comply with company policies and applicable data protection regulations.

Evaluate the risks of each AI tool before selecting one and implementing it. Use the NIST AI Risk Management Framework (RMF) to understand the risks that AI poses to your organization.

If necessary, separate confidential information systems from non-sensitive information systems. This lowers the risk of users accidentally leaking confidential data, and prevents LLMs from using your sensitive data in public outputs or for training.

Bottom Line on Improving Security with Generative AI

Overall, AI introduces risks that may not be suitable for most companies. But if a company wants to use generative AI tools, proper policy, guidelines, and training should be in place to ensure that employees are educated on the risks they present.

Visit the Echelon website to explore our range of comprehensive cyber services and safeguard your organization's future.

ECHELONCYBER.COM