MENLO SECURITY

# The continued impact of generative AI on security posture

**Revisiting how the usage of Generative AI platforms like ChatGPT is transforming how we work and posing new security risks to organizations.**


**Report Volume 2**

# Generative AI: diversification, concerns, and unanswered questions



On November 30th, 2023, OpenAI's ChatGPT celebrated their one-year anniversary. While the landscape looks quite different, the security risks around generative AI usage are still a frequent concern for security and IT teams. The last 12 months have exposed many of the risks of using generative AI, including exposing sensitive data, such as customer data, trade secrets, classified information, and even intellectual property when using these platforms.

Organizations are looking for the right balance of enabling productivity and innovation with generative AI, while stopping the potential loss of proprietary data or other intellectual property. To find the right balance, organizations must have an updated understanding of the market.

In the last 6 months, from July 2023 to December 2023, the following changes around generative AI have occurred:

- **Diversified and specialized:** As funding to the market increases, new platforms have launched, leading to a significant increase in the number of platforms and specializations.
- **Data privacy concerns:** Not only are organizations concerned about data loss from the end-user, but there are also data privacy concerns of the platform themselves. In March 2023, OpenAI experienced its first documented breach where data of around 1.2 million subscribers was exposed. This data potentially included the user's first and last name, email address, payment address, credit card type, partial credit card number, and expiration date. This breach started to raise questions about platform security and how personal data is used to train their models.

- **Private versions of AI:** There's been a growing trend within organizations that train algorithms for specific needs. Instead of hiring an in-house team to build and support an AI model, which is extremely resource intensive, organizations can use a platform, then fine-tune the algorithm with private data, and ensure that data is not used in widely-available algorithms. While there's less of a data loss concern with private versions of AI, there are still privacy and compliance concerns that organizations should be aware of. This may modify the rate of adoption of public generative AI platforms within the business environment.
- **Decreasing rate of growth:** ChatGPT became one of the fastest-growing platforms in history, collecting over 100 million users in just two months. However, the exponential usage of generative AI is no longer continuing. While there is still growth, there's a decreasing rate of growth.

Even with all of these changes, it's clear that generative AI is here to stay but it will evolve. That means Security and IT teams still need to make sure they have technology and policies in place that protect and evolve as the landscape evolves.

## The impact of AI on phishing scams:

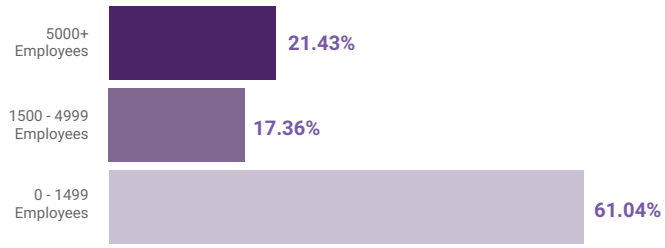
While this report focuses on data loss, there is an increased concern around AI-generated phishing scams. Just as millions of users are employing these platforms to improve their day to day, bad actors can be using these platforms to bolster their phishing campaigns. Instead of phishing scams with misspellings, poor grammar and awkward phrasing, platforms like ChatGPT help fix all those mistakes. Additionally, ChatGPT could potentially be manipulated into generating hacking code. It's important to understand the various ways bad actors might be using generative AI to target organizations and even more importantly, have the protections in place to stop these types of threats.

# Methodology

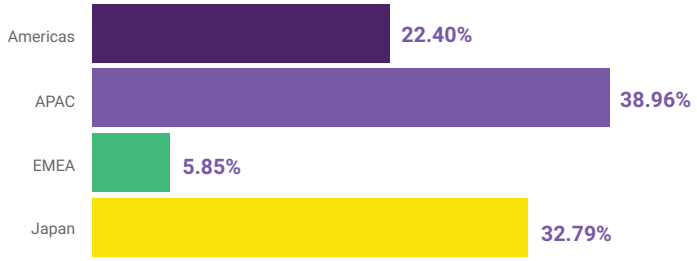
In June 2023, Menlo Security analyzed generative AI interactions from a sample size of 500 global organizations. In this second volume, the report will focus on any changes to how cybersecurity has been impacted by employee usage of generative AI.

In order to provide an accurate comparison, the data will compare the same 6 generative AI domains as the first volume unless otherwise stated. The report will also look at generative AI domains as a broader category.

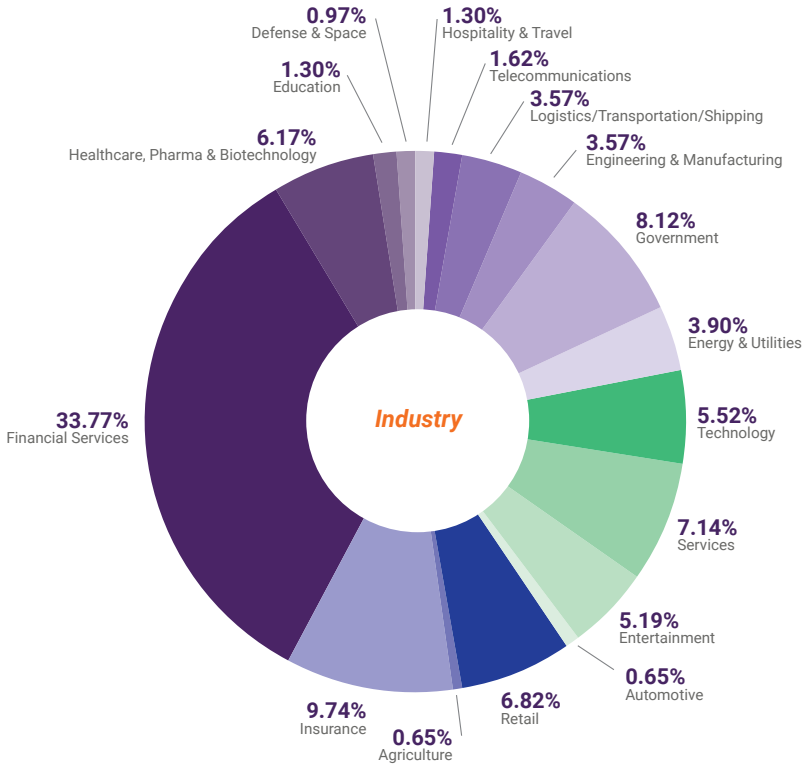
***Business Size***




***Regions***



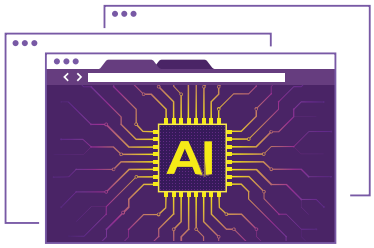

sample size of 500 global organizations



INSIGHT ONE:

## There's been consistent growth in generative AI site visits and power users in the enterprise

Unlike the decreasing rate of growth of generative AI usage that has been reported, usage and growth within the enterprise continue to be significant. This insight might highlight the differences between business usage versus personal usage. In a business setting, generative AI could help create new ideas, improve emails, create content, and check for spelling and grammar mistakes.




**>100%**

**increase in visits to generative AI sites**

**+64%**

**increase in users that are visiting generative AI sites**

## What sites are users visiting?

We're beginning to see a shift in the popularity of certain generative AI platforms. Unsurprisingly, OpenAI's ChatGPT takes the majority of traffic, however there's been an increase of visits and usage of generative AI platforms with a more specific function. For example, QuillBot AI, a rewording tool, and other grammar-focused AI platforms have started to become increasingly popular.

INSIGHT TWO:

# Organizations are applying more security-focused technology policies to generative AI sites — however, the majority are doing so on a domain basis vs a group basis

There has been a 26% increase of security policies for generative AI traffic. As generative AI continues to be top of mind, security and IT leaders have seen the need to enact security-focused policies around the usage of generative AI within their organizations. However, the current approach by the majority of organizations is not scalable.

### When looking at organizations that have security policies on a domain basis:


**92%**

**have security-focused policies in place around generative AI usage**


**8%**

**are allowing unrestricted generative AI usage**

### When looking at organizations that have security policies on a group basis:


**79%**

**have security-focused security policies in place around generative AI usage**


**21%**

**are allowing unrestricted generative AI usage**

As stated earlier in the report, the generative AI landscape is constantly shifting, especially with new platforms and new functionality. For security and IT teams that apply policies on a domain-by-domain basis, they must revisit that list frequently to ensure that users are not accessing and, potentially, exposing sensitive data on a more obscure platform. This process can be time consuming and ultimately will not scale. Organizations need to adopt security technology that enables policy management on a generative AI group level, providing protection against a broader cross-section of generative AI sites.

**Please note that organizations might have other technologies in place to protect or block generative AI usage.*

INSIGHT THREE:

# Your employees are inputting data in multiple ways

Most users continue to input their questions through typing, but there are two other common sources of data loss—file uploads and copy & paste. Interestingly, the incidence of data loss through file uploads is on the rise. Previously, most solutions did not natively allow file uploads, but as new versions of generative AI platforms are released, new features are added, such as the ability to upload a file.

Copy & pasting and file uploads could have the largest impact on data loss due to the amount of data that is quickly uploaded or inputted. Examples include:

- Copy & pasting source code, customer lists, or roadmap plans
- Uploading a spreadsheet with hundreds of columns

## When looking at organizations that have security policies on a domain basis:

**6%**

**decrease in the amount of copy/paste events that were blocked to generative AI sites**

**80%**

**increase in the amount of attempted file uploads to generative AI sites**

While the majority of traffic is directed towards the main 6 sites, when looking at generative AI as a category, file uploads are 70% higher. This highlights the importance of enabling security policies on a group level vs on a domain-by-domain basis.

INSIGHT FOUR

# Employees are still attempting to input sensitive data into generative AI

The data loss implications around generative AI are well documented and while many organizations have implemented corporate policies we're still seeing data loss events on generative AI platforms. Most organizations have sent out policies to their employees on responsible use of generative AI, however, this data illustrates how employees knowingly or unknowingly still attempt to input sensitive information into these platforms. This highlights the need to supplement with appropriate cybersecurity technology.

| Type of data employees are inputting into gen AI | | |
|---|---|---|
| **PCI** | 0.03% | We analyzed how often employees were attempting to input sensitive and confidential information into generative AI platforms. In the past 30 days, there were DLP events with data pertaining to the following categories:<br><br>The most frequent potential exposure was personally identifiable information.<br><br>These organizations have Menlo Security in place that blocked these instances. |
| **PII** | 55.11% | |
| **Confidential Documents** | 39.86% | |
| **Medical Information** | 0.90% | |
| **Restricted Information** | 3.44% | |
| **Other** | 0.67% | |

## Generative AI in the Federal Space

Generative artificial intelligence (AI) has captured the attention of millions across the world, however, the public sector is still in its early stages of understanding how to utilize generative AI within various functions.

The attention on AI has led to an Executive Order on the "Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence" which placed the highest urgency around safe and responsible use of AI. While this executive order focuses generally on artificial intelligence, there are important aspects specific to generative AI that agencies must consider. The order states:

> *"Agencies are discouraged from imposing broad general bans or blocks on agency use of generative AI" but instead are urged to put appropriate safeguards in place to utilize generative AI "at least for the purposes of experimentation and routine tasks that carry a low risk of impacting Americans' rights."*

Agencies must ensure they are able to balance the positive impact that generative AI can have on citizens, businesses, and government, while stopping the potential security risks.

THE SOLUTION

# Adopt technology that safeguards generative AI without impacting the user experience

Organizations must adopt the right technology to enable the safe usage of generative AI. Outright blocking of generative AI platforms is not a workable policy, and is even discouraged within the public sector. Organizations need a layered approach. Instead of a single technology, like Data Loss Prevention (DLP), organizations need capabilities that address the various ways employees are using these platforms.

For example, employees are copying and pasting large amounts of data and uploading files into generative AI platforms. With controls around these avenues, organizations can provide protections in a way that doesn't prevent employees' use of these helpful tools. The solution is copy & paste controls with character limits. Organizations can protect against mass data loss by limiting what can be pasted into input fields – either restricting character counts or blocking known code. No one is going to manually type in thousands of lines of source code, so limiting paste functions effectively prevents this type of data loss. It also would make users think twice about the information they were trying to input.

Organizations should not see generative AI as an unknown. Organizations can also apply security policies that trigger additional controls – such as event logging or initiating session monitoring – to aid in resolution and post event analysis. It's important to remember that investigations into breaches caused by insiders must provide proof of intent. Recording events and browsing sessions could provide visibility and insight into whether users were malicious or just negligent.

Lastly, organizations should adopt technology that enables security controls on a generative AI group level. As illustrated in the report, new generative AI sites are being used frequently and increase the possibility of data exposure. If policies are applied on domain-by-domain basis, organizations must either constantly update their list or risk gaps in safeguards to generative AI sites that employees are using.

www.menlosecurity.com

## About Menlo Security

Menlo Security eliminates evasive threats and protects productivity with the Menlo Secure Cloud Browser. Menlo delivers on the promise of cloud-based security—enabling Zero Trust access that is simple to deploy. The Menlo Secure Cloud Browser prevents attacks and makes cyber defenses invisible to end users while they work online, reducing the operational burden on security teams.

Menlo protects your users and secures access to applications, providing a complete enterprise browser solution. With Menlo, you can deploy browser security policies in a single click, secure SaaS and private application access, and protect enterprise data down to the last mile. Secure your digital transformation with trusted and proven cyber defenses, on any browser.

Work without worry and move business forward with Menlo Security.

Contact us today to learn how to safely enable generative AI while protecting against data loss.