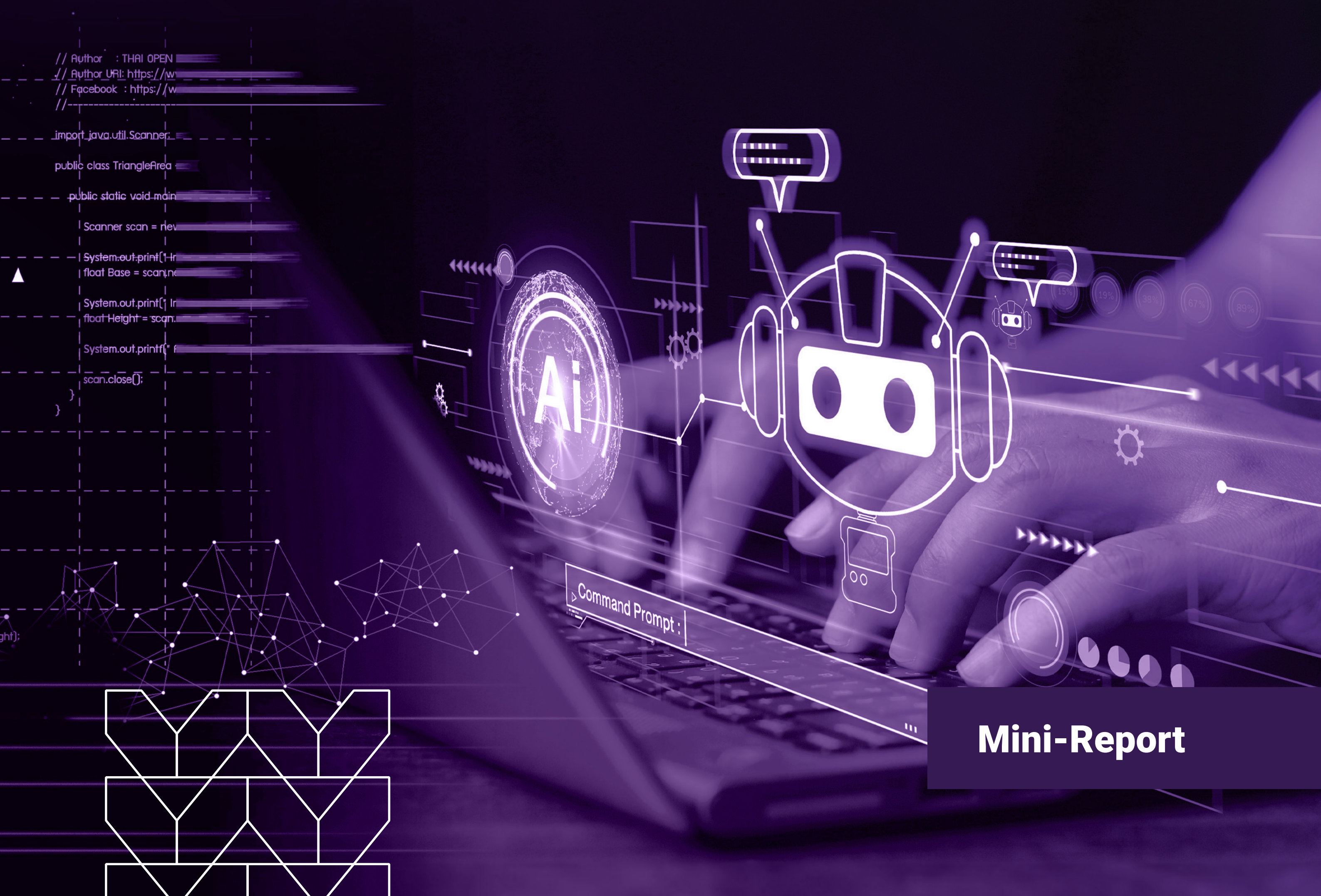
MENLO SECURITY

How employee usage of generative AI is impacting organizational security

Generative AI platforms like ChatGPT are transforming the way we work while posing new risks to organizations.



Mini-Report

Generative AI is introducing both increased productivity and risk to organizations

Since its release in November 2022, ChatGPT has become one of the fastest-growing platforms in history, collecting over 100 million users in just two months. In comparison, TikTok took nine months and Instagram took 2.5 years to reach the same number of users. ChatGPT has captured global attention, and it's just one out of many generative AI sites being used daily. This utilization has led to increased productivity and innovation as employees use generative AI to create new ideas, improve their emails, develop content, check for spelling/grammar mistakes, and so much more.

While ChatGPT has led to increased productivity and innovation, it has also led to real concerns, particularly in the area of cybersecurity. Many are nervous that generative AI platforms will allow threat actors to develop evasive threats at an alarming scale. In addition, platforms like ChatGPT have lowered the barriers for hackers to launch more sophisticated and effective phishing attacks. While these concerns are justified, cybersecurity experts also need to consider the more immediate effects of widespread generative AI use — the potential loss of proprietary data or other intellectual property (IP).

As employees use generative AI tools, such as ChatGPT and Bard, they might inadvertently be sharing and exposing sensitive company data, including customer data, trade secrets, classified information, and even intellectual property. With generative AI, private data has the potential for much greater exposure than typical data loss through breaches, improper sharing, and other avenues, because generative AI platforms save data, such as chat history, to train and improve their models. That means any data that was input could be used to train the models and potentially be exposed later to other users.

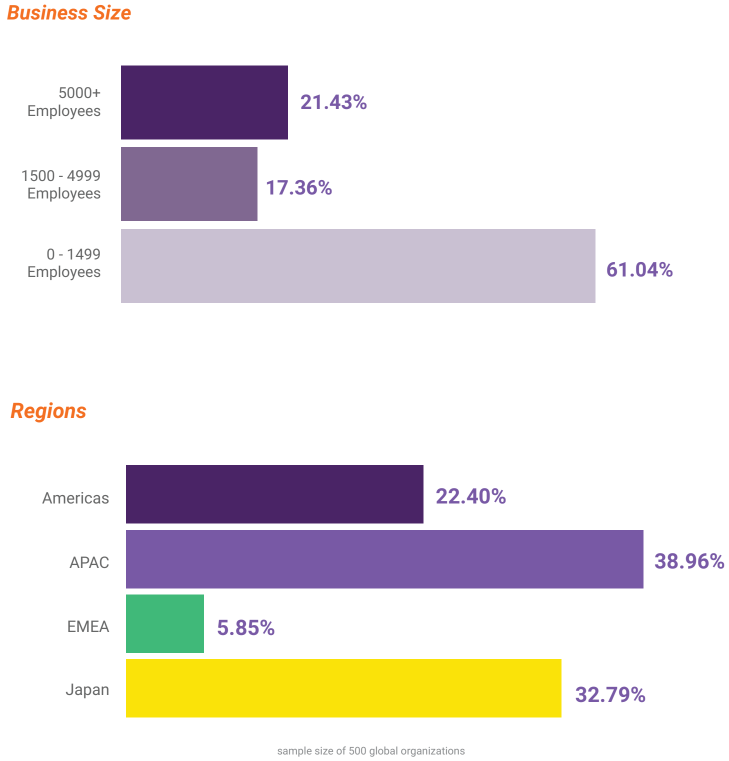
Real-world example:

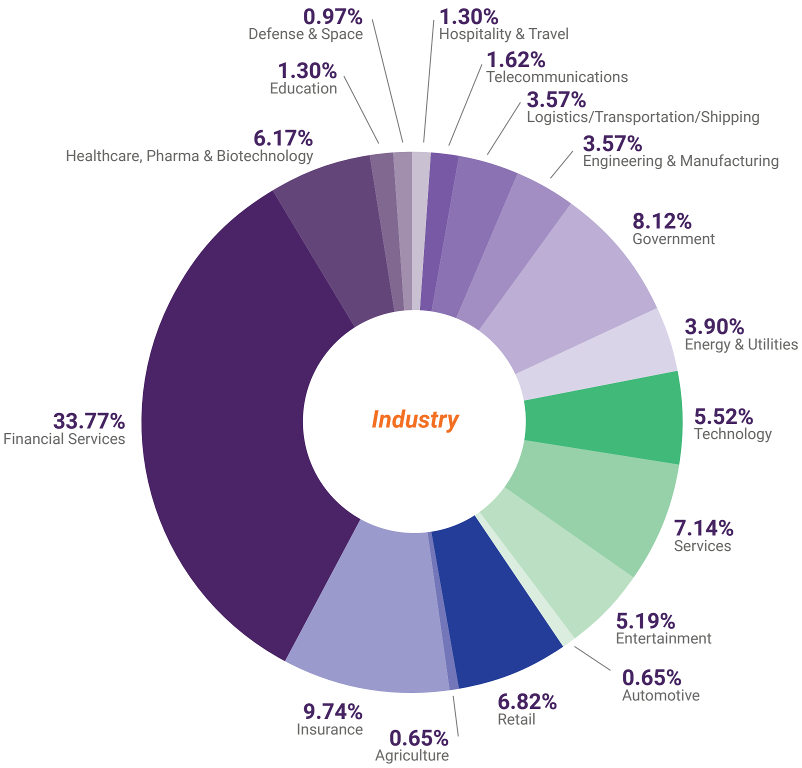
It was recently reported that a group of engineers from Samsung's semiconductor group input source code into ChatGPT to see if it could be made more efficient. Because generative AI platforms retain the data input by users to further train their models, the Samsung code can now be used to formulate a response to requests from other users. This could include a threat actor looking for vulnerabilities or a competitor looking for proprietary information.

To protect their organizations, some companies have outright banned generative AI sites. While preventing access to generative AI services may seem like a solution to potential security risks, it's just a quick fix. ChatGPT and the myriad of other generative AI platforms are powerful business tools that people can use to streamline business processes, automate tedious tasks, or get a head start on a writing, design, or coding project. Blocking these sites will also block productivity and business agility.

Methodology

In order to provide insight into the world of generative AI and its cybersecurity impact, Menlo Security analyzed generative AI interactions from a **sample size of 500 global organizations**. This snapshot focuses on how frequently employees are utilizing generative AI and the real impact it is having on data loss.





INSIGHT ONE:

Employees' use of generative AI is growing exponentially

Between November 2022 and May 2023, generative AI usage increased 1,200%, and this number is growing every day.

These employees are utilizing generative AI sites frequently — and each time is an opportunity for potential data loss.

Over 30 days, there were:

2.5 Million
visits to Generative AI sites for 500 organizations

78,825
Users at these organizations are utilizing Generative AI sites

That means...
32x
a month per user

What sites are users visiting?

While the list of generative AI sites is extensive, for the purposes of this snapshot we have decided to focus on the top six AI applications: chat.openai.com; bing.com; bard.google.com; writesonic.com; copy.ai; jasper.ai

The top three generative AI sites are:

OpenAI (ChatGPT)	1,977,607 visits
Microsoft (Bing)	410,850 visits
Google (Bard)	96,181 visits

Over a 30-day period, OpenAI accounted for more than **50% of the generative AI visits.**

INSIGHT TWO:

There isn't just one avenue of data loss

It's important to think about all the different ways employees are using generative AI. Are they entering their questions or queries into the search bar? Do they copy and paste information from a different source? Or are they uploading files along with their requests?

While most users are typing out their questions and queries, the other two avenues of data loss could have the largest impact. Using file uploads and copy and paste, employees could expose large amounts of sensitive data at a much faster rate. Examples include:

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

File upload—10,190 events
(In 30 days)

Copy/Paste—3,394 events
(In 30 days)

While the numbers of events are lower comparatively to the total number of generative AI visits, they have the highest impact in terms of potential data loss.

(Please note, some generative AI sites do not have a native file upload option, although users are able to upload through a plug-in. We have included file upload attempts that may have failed.)

File uploads by type over 30 days	
Unknown (ex. text files)	1,237
PDF	133
Word	48
Excel	22
PowerPoint	21
Script	13
WinEXE	3
ZIP	2

Employees are uploading a variety of file types into generative AI platforms. PDFs are the most frequently uploaded identifiable type.

File types like Excel could lead to accidental data loss. While an employee may think the file is safe, there could be hidden rows or columns with sensitive or confidential data.

INSIGHT THREE:

The threat of data loss to generative AI is real

Data loss prevention (DLP) is not a new concept or technology for the cybersecurity world. Due to the potential repercussions of having data move outside an organization's control, attempts to limit and/or block data loss have been an ongoing concern and problem for organizations in all industries and of all sizes. With generative AI, this concern has increased.

The breadth of sensitive data being sent to generative AI sites shows that this is an issue that goes well beyond the publicized source code concerns.

The following table illustrates that potential data loss incidents are regularly happening at organizations. As employees are using generative AI, they are knowingly or unknowingly exposing sensitive or proprietary data.

We analyzed how often employees were attempting to input sensitive and confidential information into generative AI platforms. In our analysis, there were DLP events with data pertaining to the following categories:

Types of data employees are inputting into generative AI	
PCI	5.4%
PII	50.4%
Confidential Documents	24.6%
Medical Information	2.2%
Restricted Information	1.5%
Other	15.9%

The most frequent potential exposure involved personally identifiable information (PII). The organizations in our study have policies in place that blocked these instances.

THE SOLUTION

Empower safe usage of generative AI to properly manage the risk of data loss

Unfortunately, existing data loss prevention (DLP) solutions, cloud access security brokers (CASBs), and other insider threat solutions are not enough to deal with the nuances of this new technology. Using a detect-and-respond approach, these solutions look for keywords or phrasing amongst the enormous amount of traffic flowing out of the organization. These elements often have to be entered individually by security professionals and product owners in a manual and error-prone process. Even if a solution detects data exfiltration, it could already be too late. Once information has been sent to a generative AI platform, there is no way to take it back. The data lives in the platform, and may continue to inform responses to queries.

Organizations need to prevent information from being entered into generative AI platforms and chatbots in a way a way that doesn't prevent employees' use of these helpful tools. While DLP is useful, organizations need a layered approach instead of focusing on a one-size-fits-all solution.

First, organizations can limit what can be pasted into into input fields with policies that restrict character count, block known code, or the like. 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, and may make users think twice about the information they were trying to input.

Most importantly, however, organizations should direct interaction with ChatGPT and other generative AI platforms away from the web browser. Executing app commands in a secure browser in the cloud provides an extra layer of protection between the user and the internet, giving the organization an opportunity to stop malicious activity (whether it's purposeful or not) before data exfiltration occurs.

Organizations can also apply security policies that trigger additional security controls, such as event logging or initiating a browser recording, to aid in issue 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 browser sessions could provide visibility and insight into whether users were malicious or just negligent.

MENLO SECURITY

www.menlosecurity.com

About Menlo Security

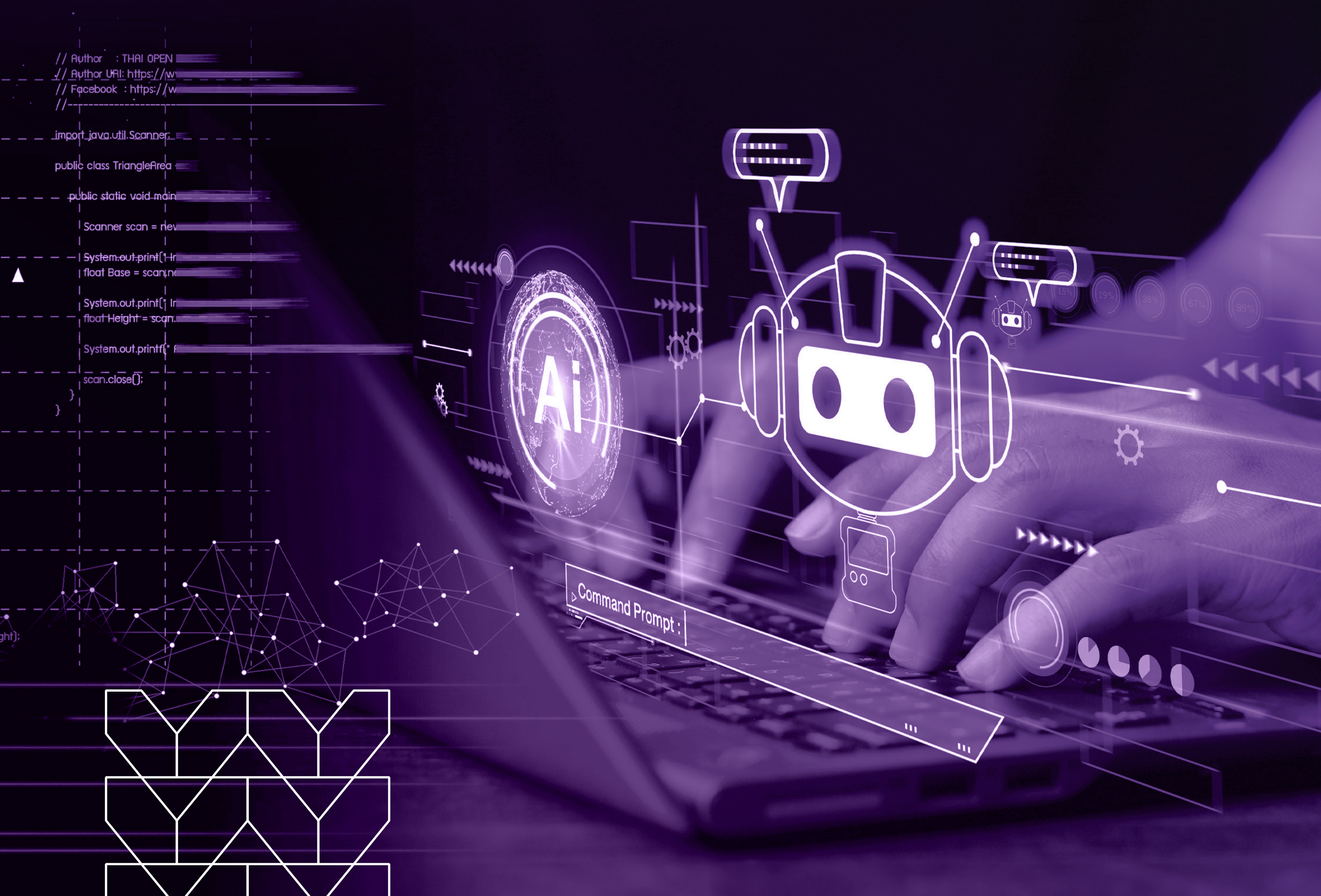
Menlo Security protects organizations from cyberattacks by eliminating the threat of malware from the web, documents, and email. It focuses on protecting the single biggest productivity driver for knowledge workers—the web browser.

Menlo's Cloud Security Platform prevents threats from entering an organization and secures data and application access in a single, global cloud-based offering. Our Elastic Isolation Core™ creates separation between the user, content and applications where security, policy and visibility are applied. With deep visibility inside the browser, adaptive policy enables the prevention of threats before they happen, as opposed to detecting and responding, organizations eliminate all threats, including Highly Evasive Adaptive Threats (HEAT) across web, email, SaaS applications and private applications.

HEATcheck

Menlo Security provides a lightweight penetration assessment to help organizations better understand any susceptibility to various HEAT attacks. The assessment leverages various real-world HEAT attacks currently being used by threat actors, safely allowing organizations to deduce their exposure. Menlo's HEATcheck tool does not deliver actual malicious content.

Contact us today to learn if your organization is susceptible to web threats, and discover how you can prevent them from putting the organization at risk.



```
// Author   : THAI OPEN
// Author URI: https://w
// Facebook  : https://w
//----------------------
import java.util.Scanner;
public class TriangleArea
  public static void main
    Scanner scan = new
    System.out.print("Ir
    float Base = scan.ne
    System.out.print(" Ir
    float Height = scan.
    System.out.printf(" f
    scan.close();
  }
}
ght);
```