Layer

REVEALING THE TRUE GENAI DATA EXPOSURE RISK

RESEARCH REPORT

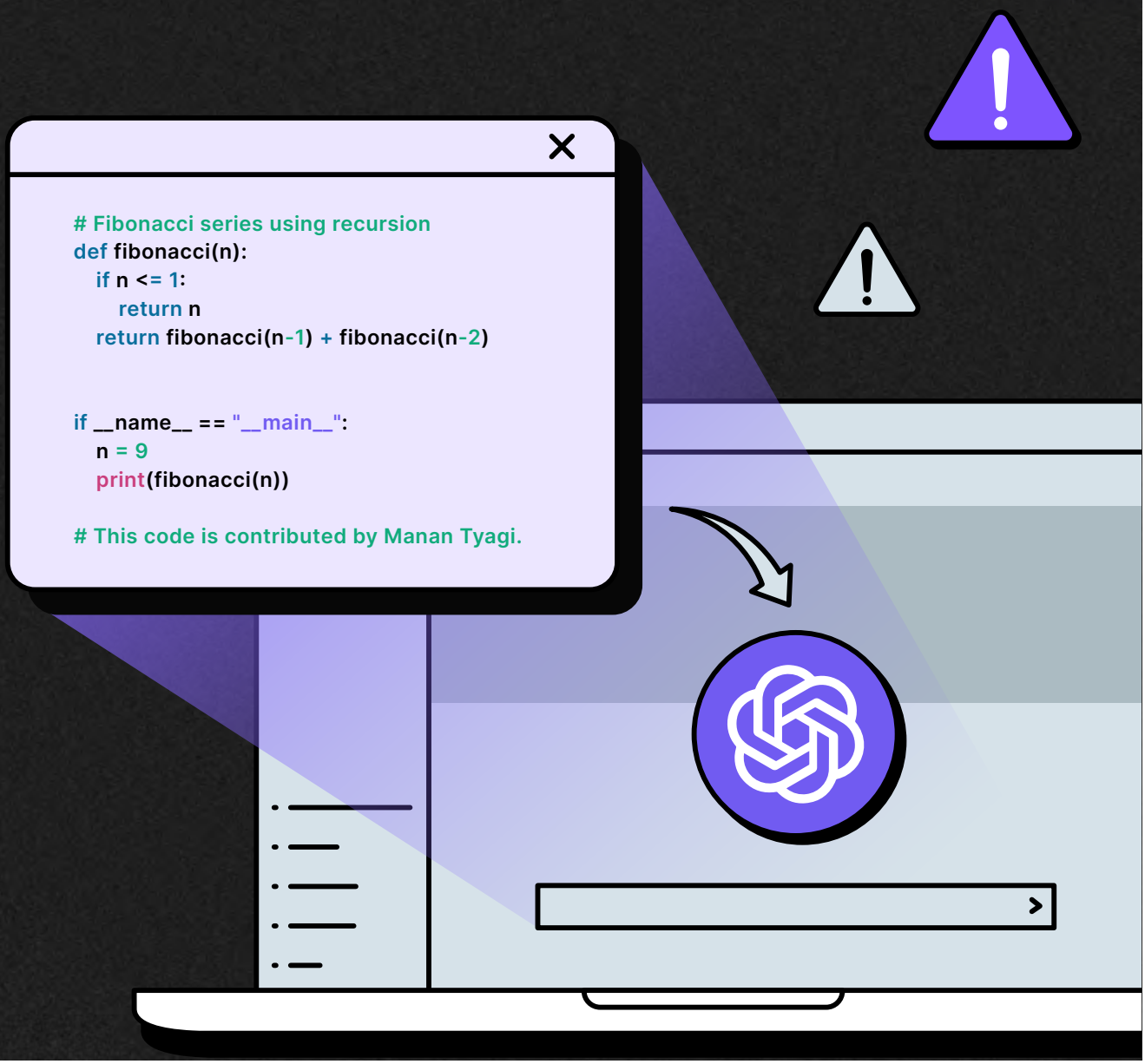


TABLE OF CONTENTS

EXECUTIVE SUMMARY:
SENSITIVE DATA IS POURING OUT IN INCREASING VOLUMES

ChatGPT and other GenAI tools are enthusiastically being embraced by the global workforce across all verticals and geolocations. However, **ungoverned pasting of data into GenAI tools could expose it to the entire GenAI user community**. While such general warnings have been made numerous times lately, there is still a gap in concrete data on the volume, trends, and type of exposed data.

This research was initiated to fill this gap, and provide, for the first time, the true numbers behind the GenAI data exposure FUD. To do that we've analyzed ChatGPT and other generative AI app usage for 10,000 employees (using data from devices with the LayerX browser extension installed).

Key Findings:

- **15% of employees have pasted data into GenAI.** Pasting is the riskiest action taken on GenAI because it is beyond the reach of existing data protection solutions.
- **6% of employees have pasted sensitive data into GenAI.** This behavior is putting their organization at risk of data exfiltration.
- **4% of employees paste sensitive data into GenAI on a weekly basis.** The risk is recurring, increasing the chances of sensitive data exposure.
- **Source code (31%), internal business information (43%), and Personal Identifiable Information (PII) (12%) are the leading types of pasted sensitive data.** Organizations might be unknowingly sharing their plans, product and customer data with competitors and attackers.
- Data was mostly pasted by users from the R&D, Sales & Marketing, and Finance departments.

We believe that the enclosed findings indeed substantiate the deep concern that has been expressed by security stakeholders in the recent months. It is also our hope that, armed with this knowledge, these stakeholders will find it easier to craft a sound strategy to better safeguard their sensitive data and mitigate the exposure risks.

FINDING #1:
Is GenAI Usage that Wide?

Key Insight

Usage has increased rapidly but is still practiced by a relatively small group of users.

Findings

44% ⬆

in GenAI API usage over the past 3 months. On the week of May 15th, our product recorded a record of 1,373 visits per 1,000 employees to GenAI platform pages. That's an average of more than 196 visits per day!

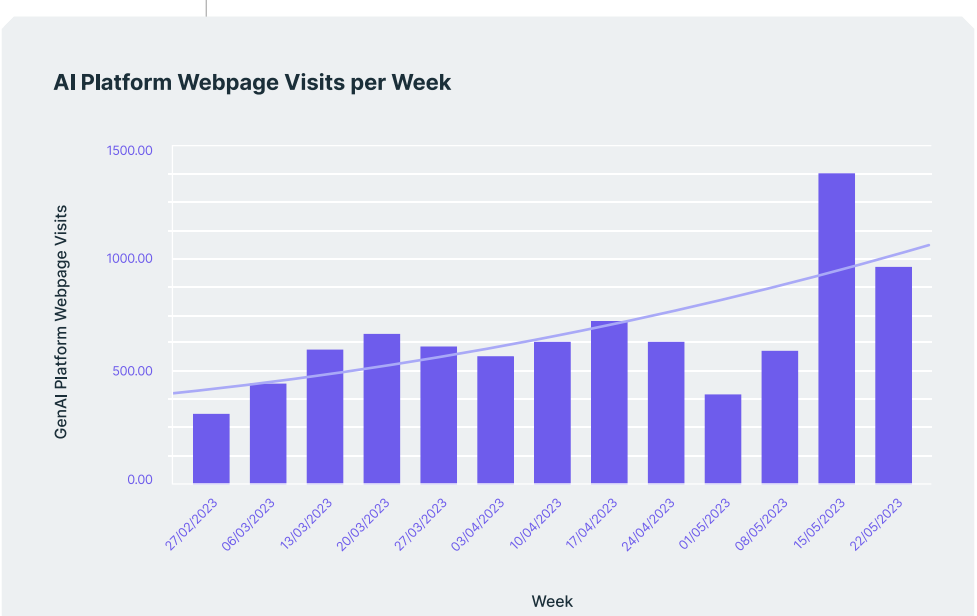


Diagram 1: 44% increase in GenAI webpage visits over three months (normalized for 1,000 employees)

19% OF EMPLOYEES

have visited a GenAI application page at least once in the past month.

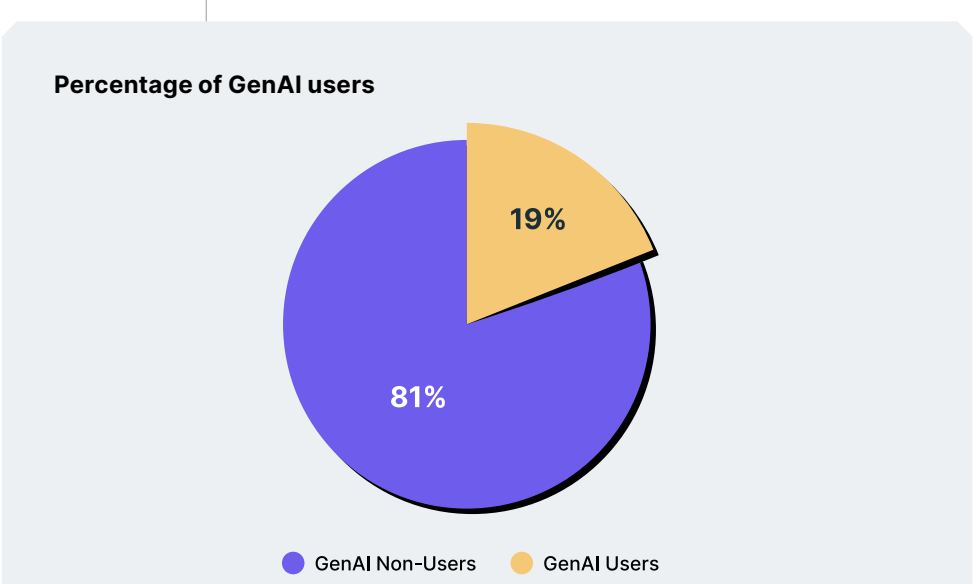


Diagram 2: GenAI usage among employees

5% OF GENAI USERS

are using GenAI 6 times more than all the rest (50 visits per month).

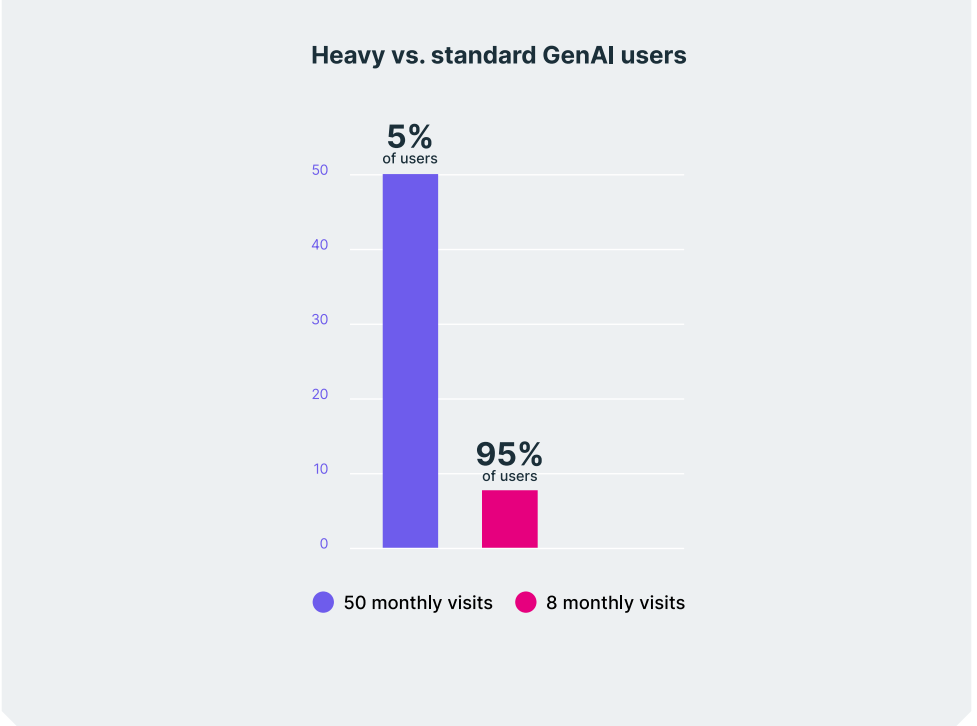


Diagram 3: Heavy vs. standard GenAI users

Analysis

GenAI usage is increasing steadily. As we're still at a relatively early stage, there's no way to know for sure if this trend will keep on. However, it's reasonable to assume we're not at the peak yet. This assumption might be reinforced by the fact that even in the last month GenAI users were still less than 20% of the entire workforce. Moreover, only 5% of these can be considered heavy GenAI users.

FINDING #2:
How Much Data is Pasted into GenAI?

Key Insight

Pasting data into GenAI is a prevalent action.

Findings

15% OF EMPLOYEES
have pasted data into GenAI.

36 TIMES\DAY
is the average number of data pasting occurrences per 1,000 employees.

23.5% OF VISITS
to GenAI apps include a data paste.

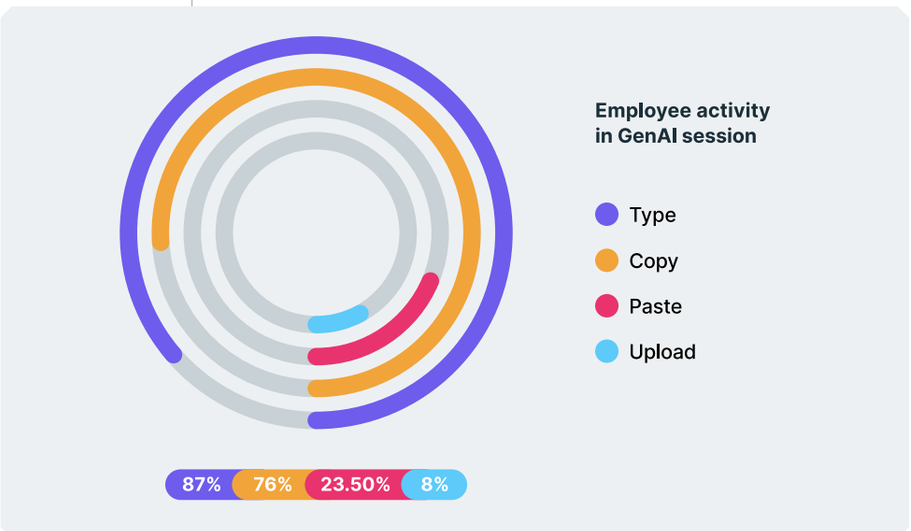


Diagram 4: GenAI session-activity breakdown

Analysis

It seems as if a significant portion of GenAI users don't rely on prompt instructions alone, but also paste data in their attempt to generate a desired text. It is reasonable to assume this behavior is practiced when the requested output should fulfill very specific requirements that can't be found in the publicly available data the GenAI tool has access to. This makes pasting actions a highly likely source of data exposure.

FINDING #3:
Who are the Top GenAI Users?

Key Insight

R&D, Marketing & Sales, and Finance are the heaviest GenAI users.

Findings

Out of the 5% group that use GenAI extensively:

50%
R&D

23.8%
Marketing and Sales

14.3%
Finance

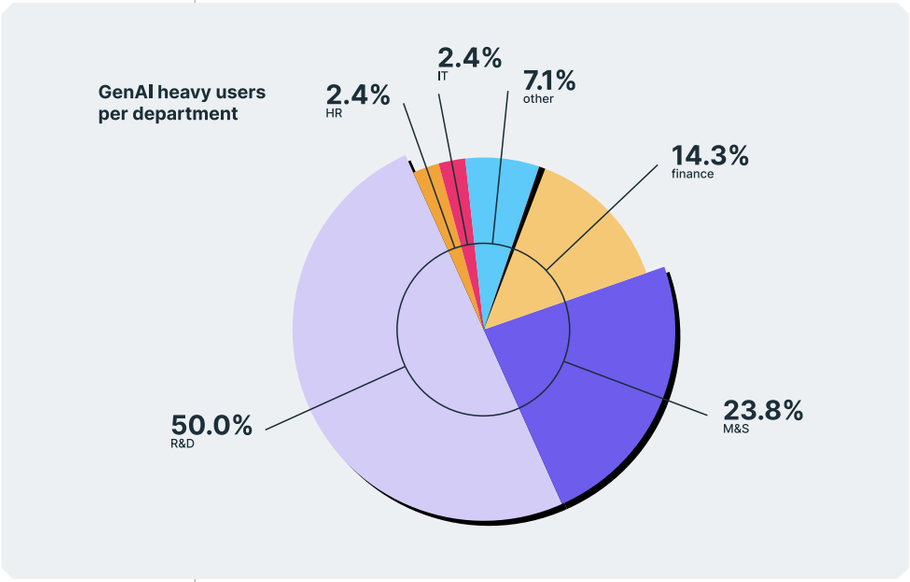


Diagram 5: GenAI heavy users per department

Analysis

This distribution clearly implies that the highest risk exposure is to data that's related to these three departments. This could include source code and sensitive business data such as planning, pricing, strategy and others. We can also assume that using GenAI tools by members of these departments inevitably entails the pasting of internal data in order for it to provide any value. For example, a Sales manager using GenAI to produce an executive summary of their quarterly performance would have to provide the GenAI tool with the actual sales results data

FINDING #4:
How Much Sensitive Data is Pasted into GenAI?

Key Insight

Pasting sensitive data into GenAI is a common occurrence.

Findings

6% OF EMPLOYEES
have pasted sensitive data into GenAI.

4% OF EMPLOYEES
paste sensitive data into GenAI on a weekly basis.

0.7% OF EMPLOYEES
paste sensitive data into GenAI multiple times a week (including over the weekend).

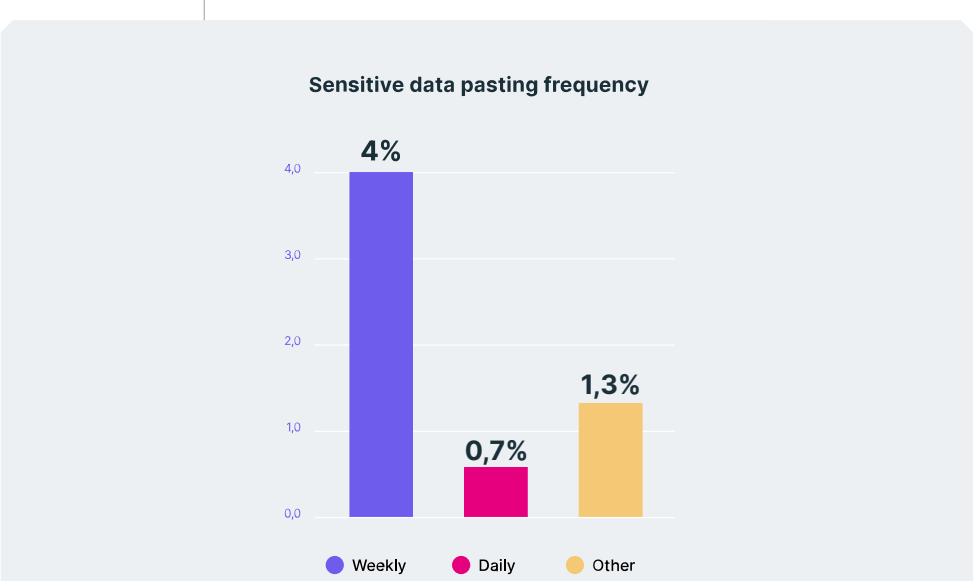


Diagram 6: Sensitive data pasting frequency

Diagram 7: Sensitive data pasting incidents in the first week of April (for 10,000 employees)

Analysis

There is a large percentage of GenAI users who are exposing sensitive company data into GenAI. This is most likely done innocently, with the goal of increasing their productivity by using GenAI to save time. However, this behavior is putting their organization at risk of data exfiltration. It also seems that such behavior is recurring, with many employees pasting sensitive data on a weekly, or even daily, basis. This goes to show that GenAI has become an inherent part of their daily workflows, raising the chances of a data exposure.

FINDING #5:
What Types of Data Are the Most Exposed in GenAI?

Key Insight

Source code and internal business data are the highest exposure risks.

Findings

From all the GenAI pasted input that was defined as sensitive data that shouldn't be exposed, these were the leading types:

43%
Internal business data

31%
Source code

12%
PII

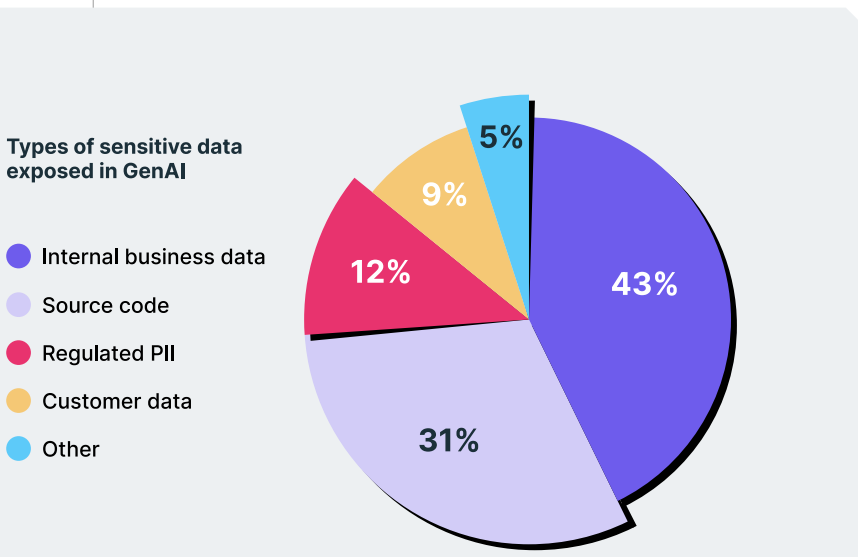


Diagram 8: Types of sensitive data exposed in GenAI

Analysis

The data type findings align with the department findings from the previous sections. They reinforce our former conclusion as to what is the data that is most prone to being exposed via pasting into GenAI. The relative low rate of PII is probably related to the fact that this type of data resides only within a subset of organizations.

CONCLUSIONS

GenAI has exploded. ChatGPT, the poster child of popular GenAI, has quickly accumulated a whopping 100 million active users in January 2023, and an even more whopping 173 million users in April 2023. Soon, we predict, employees will be using GenAI as part of their daily workflow, just like they use email, chats (Slack), video conferencing (Zoom, Teams), project management and other productivity tools.

GenAI is opening up a whole new horizon of opportunities, positively impacting productivity, creativity, and innovation. However, GenAI also poses significant risks to organizations, particularly concerning the security and privacy of sensitive data.

This report has shed light on the potential dangers of pasting sensitive data associated with GenAI. As we've shown, a significant number of employees are pasting sensitive data, like business plans, source code and PII into public sources. These numbers are only expected to grow.

CISOs can ban the use of GenAI platforms in their organizations. However, sooner or later, this will be equivalent to blocking the use of email or Zoom. A more beneficial and forward-thinking approach is to find a security solution that addresses the risks and vulnerabilities themselves, rather than obliterating the use of the platforms themselves. Don't throw out the baby with the bathwater.

Since GenAI platforms operate in the browser, existing security solutions cannot address risks like pasting of sensitive data. Instead, organizations must prioritize browser security as a means for safeguarding their systems from potential GenAI-related threats.

Browser security solutions have deep visibility into the browser sessions themselves, with the ability to alert about or block risky actions. They enable configuring policies that enforce the company policy on GenAI platforms. For example, prohibiting pasting of code into ChatGPT.

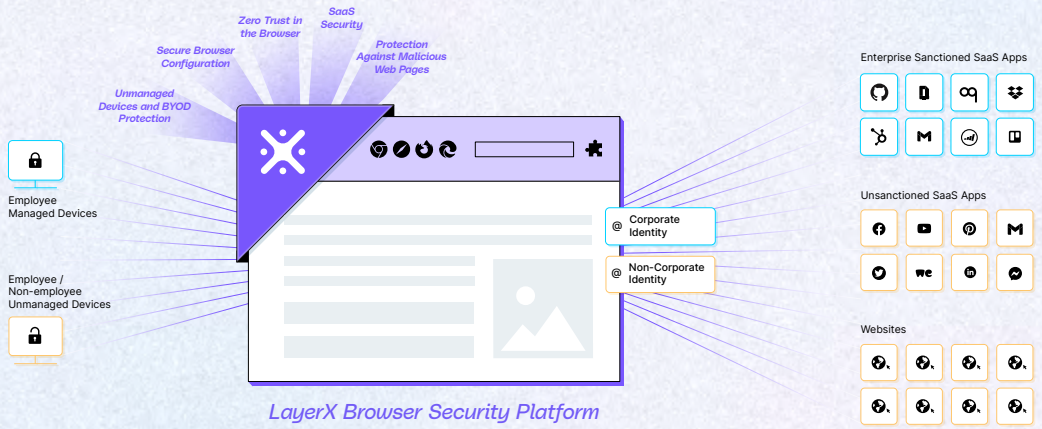
It is only through proactive and comprehensive browser security measures that organizations can confidently embrace the transformative power of GenAI while ensuring the privacy and integrity of their valuable information.

ABOUT LAYERX

LayerX provides a Browser Security Platform that's purpose-built to monitor, analyze, and protect against web-borne cyber threats and data risks. Delivered as a browser extension, LayerX natively integrates with any commercial browser, transforming it into the most secure and manageable workspace. Using LayerX, customers gain comprehensive protection against all threats that either target the browser directly or attempt to utilize it as a bridge to the organization's devices, apps, and data.

LayerX monitors every web-session at its most granular level to detect and disable risky activity at its utmost early stage with near-zero disruption to the user's browsing experience. With LayerX your workforce can securely browse anywhere.

Request Demo



KEY BENEFITS

Eliminate critical blind spots

Gain the most granular visibility into unsanctioned apps, shadow identities, DNS over HTTPS, SaaS apps, and dynamic websites.

GenAI DLP

LayerX browser security extension prevents employees from pasting of your sensitive data to ChatGPT and other GenAI platforms

High-precision risk detection

Multilayered AI analysis of every user activity and web session flags anomalies that can indicate risk in the browser session.

Unified browser management

Manage and configure your workforce's browsers from a single, centralized interface.

Bring your own browser

Enable your users to keep on using their browser of choice for both work and personal use.

Rapid deployment

Deploy across your entire environment and integrate with browser management tools and identity providers in a single click.