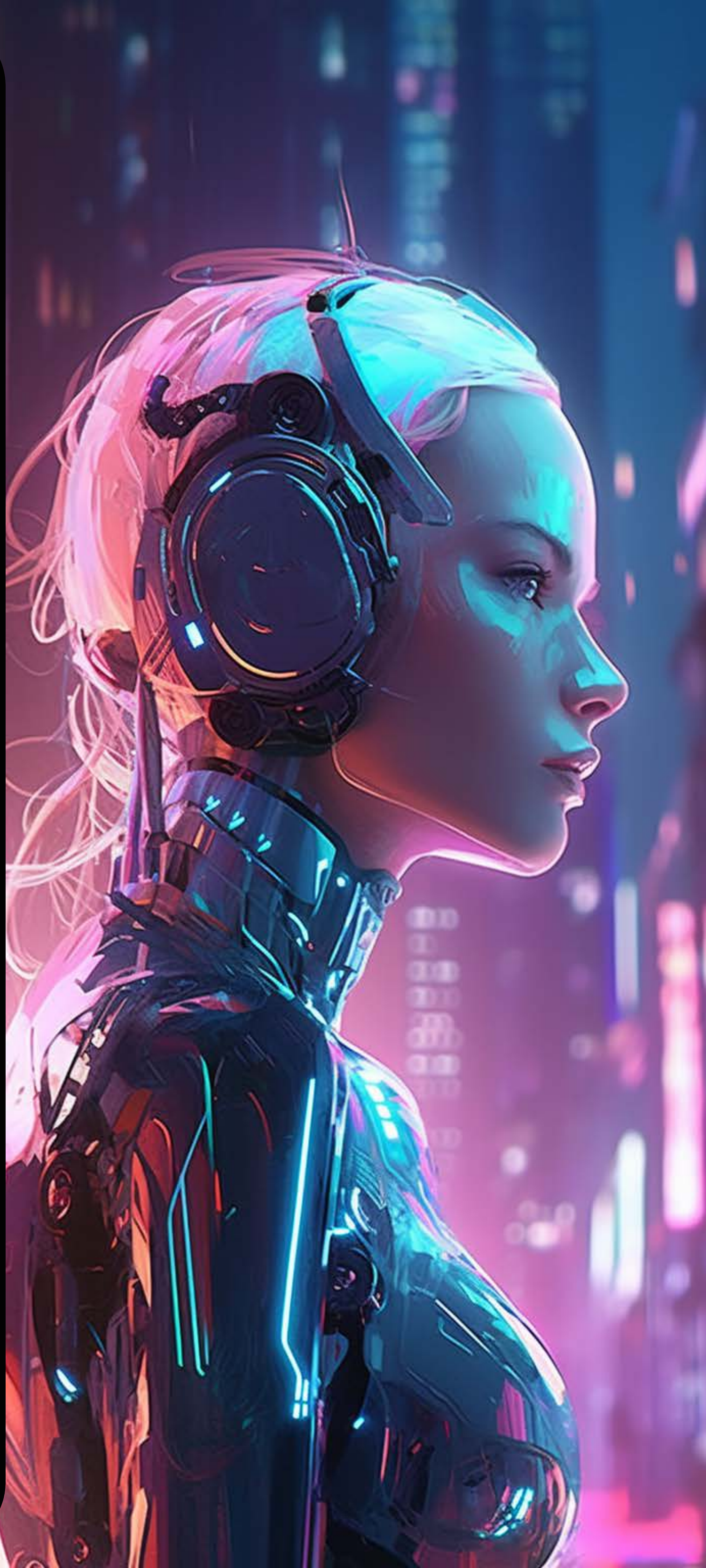
CHATGPT AT WORK

cyberhaven

Research Report

Q2 2023



Introduction

Since ChatGPT launched On November 30, 2022 it's taken the world by storm. People are using it to create poems, essays for school, and song lyrics. It's also making inroads in the workplace.

Some knowledge workers say that using the tool makes them 10 times more productive. But companies like JP Morgan and Verizon are blocking ChatGPT over risks to confidential data.

According to data from Cyberhaven's product, as of April 19, 9.3% of employees have used ChatGPT in the workplace and 7.5% have pasted company data into it since it launched. Our analysis shows that 4.0% of employees have pasted confidential data into ChatGPT.

The risk of company data and ChatGPT

OpenAI uses the content people put into ChatGPT as training data to improve its technology. This is problematic because employees are copying and pasting all kinds of confidential data into ChatGPT to have the tool rewrite it, from source code to patient medical records.

Recently, an attorney at Amazon warned employees not to put confidential data into ChatGPT, noting, "we wouldn't want [ChatGPT] output to include or resemble our confidential information (and I've already seen instances where its output closely matches existing material)."

Consider a few examples:

- A doctor inputs a patient's name and details of their condition into ChatGPT to have it draft a letter to the patient's insurance company justifying the need for a medical procedure. In the future, if a third party asks ChatGPT "what medical problem does [patient name] have?" ChatGPT could answer based what the doctor provided.

- An executive inputs bullet points from the company's 2023 strategy document into ChatGPT and asks it to rewrite it in the format of a PowerPoint slide deck. In the future, if a third party asks "what are [company name]'s strategic priorities this year," ChatGPT could answer based on the information the executive provided.

On March 21, 2023 OpenAI shut down ChatGPT due to a bug that mislabeled chats in user's history with the titles of chats from other users. To the extent that those titles contained sensitive or confidential information, they could have been exposed to other ChatGPT users.

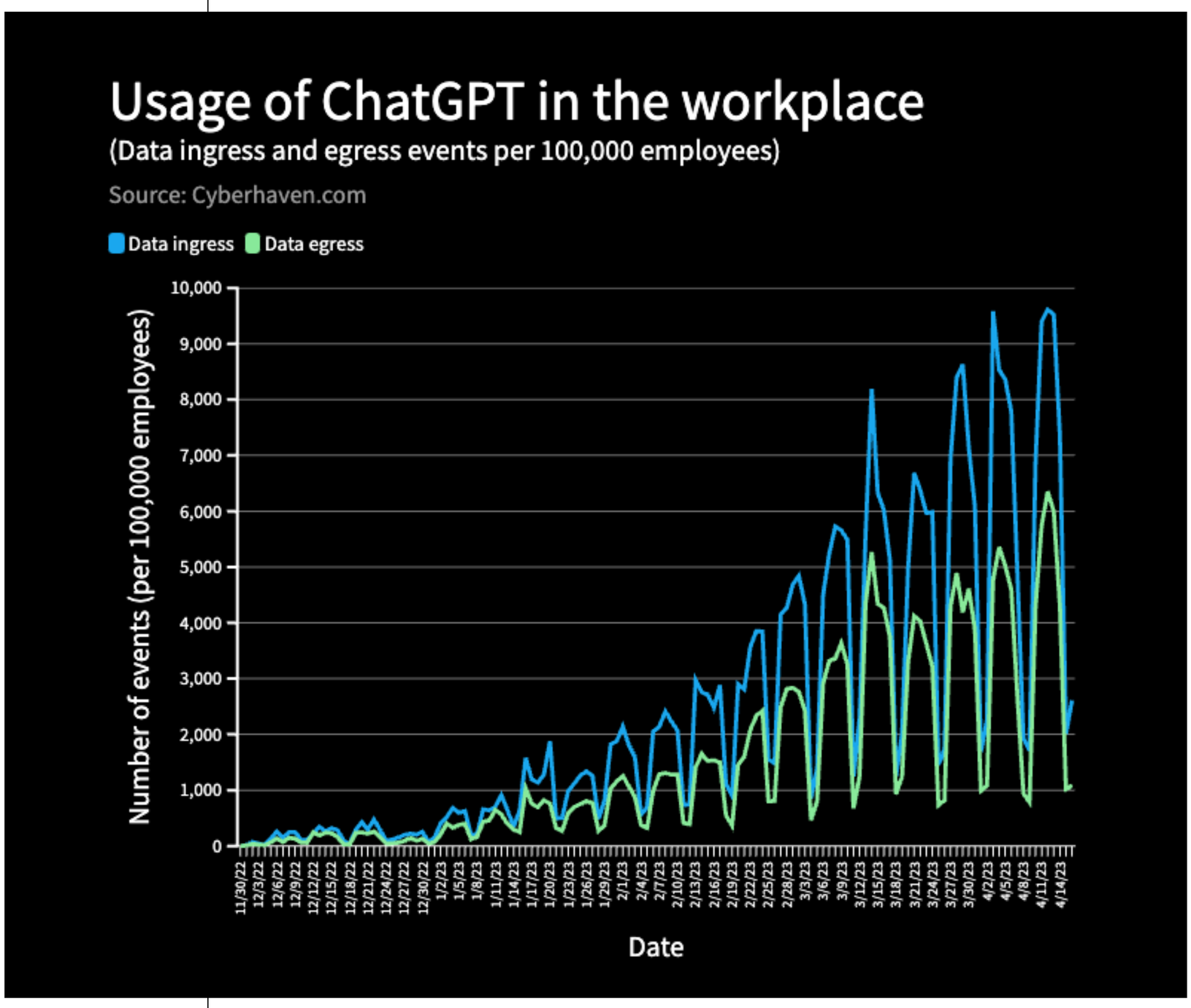
On April 6. 2023 news broke that Samsung discovered employees putting confidential data into ChatGPT including source code in order to debug it and transcripts of internal meetings to summarize them. As an emergency measure the company limited input to ChatGPT to 1024 bytes.

The rapid growth of ChatGPT Usage

Cyberhaven Labs analyzed ChatGPT usage for 1.6 million workers at companies across industries that use the Cyberhaven product.

Since ChatGPT launched publicly, 9.3% of knowledge workers have tried using it at least once in the workplace and 7.5% have pasted data into it.

Despite a growing number of companies outright blocking access to ChatGPT, usage continues to grow exponentially. On April 12, our product detected a record 6,352 attempts to paste corporate data into ChatGPT per 100,000 employees, defined as “data egress” events in the chart below.

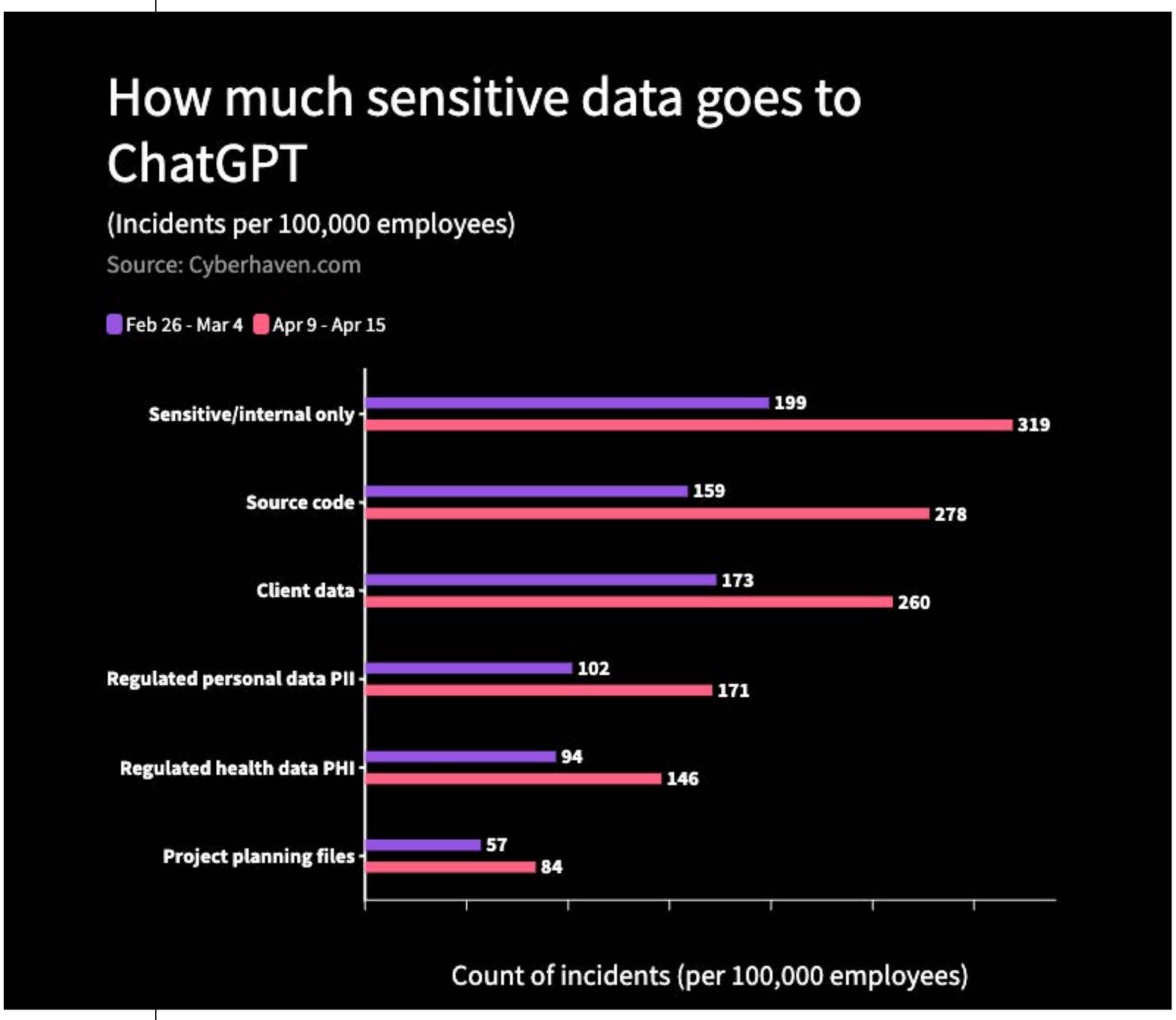


Cyberhaven also tracks data ingress such as employees copying data out of ChatGPT and pasting it elsewhere like a Google Doc, a company email, or their source code editor. Workers copy data out of ChatGPT more than they paste company data into ChatGPT at a nearly 2-to-1 ratio. This makes sense because in addition to asking ChatGPT to rewrite existing content, you can simply type a prompt and it will write a response from scratch.

Identifying the flow of sensitive data

Since ChatGPT launched, 4.0% of employees have pasted sensitive data into the tool at least once. Sensitive data makes up 11% of what employees paste into ChatGPT, but since usage of ChatGPT is so high and growing exponentially this turns out to be a lot of information.

Cyberhaven Labs calculated the number of incidents per 100,000 employees to understand how common they are across companies. You can apply this rate of incidents to the number of employees at any given company to estimate how much data employees are putting into ChatGPT.



Between the week of February 26 and the week of April 9, the number of incidents per 100,000 employees where confidential data went to ChatGPT increased by 60.4%. The most common types of confidential data leaking to ChatGPT are sensitive/internal only data (319 incidents per week per 100,000 employees), source code (278) and client data (260). During this time period source code eclipsed client data as the second most common type of sensitive data going to ChatGPT.

cyberhaven

To gain visibility and control over your data, contact us today.

www.cyberhaven.com