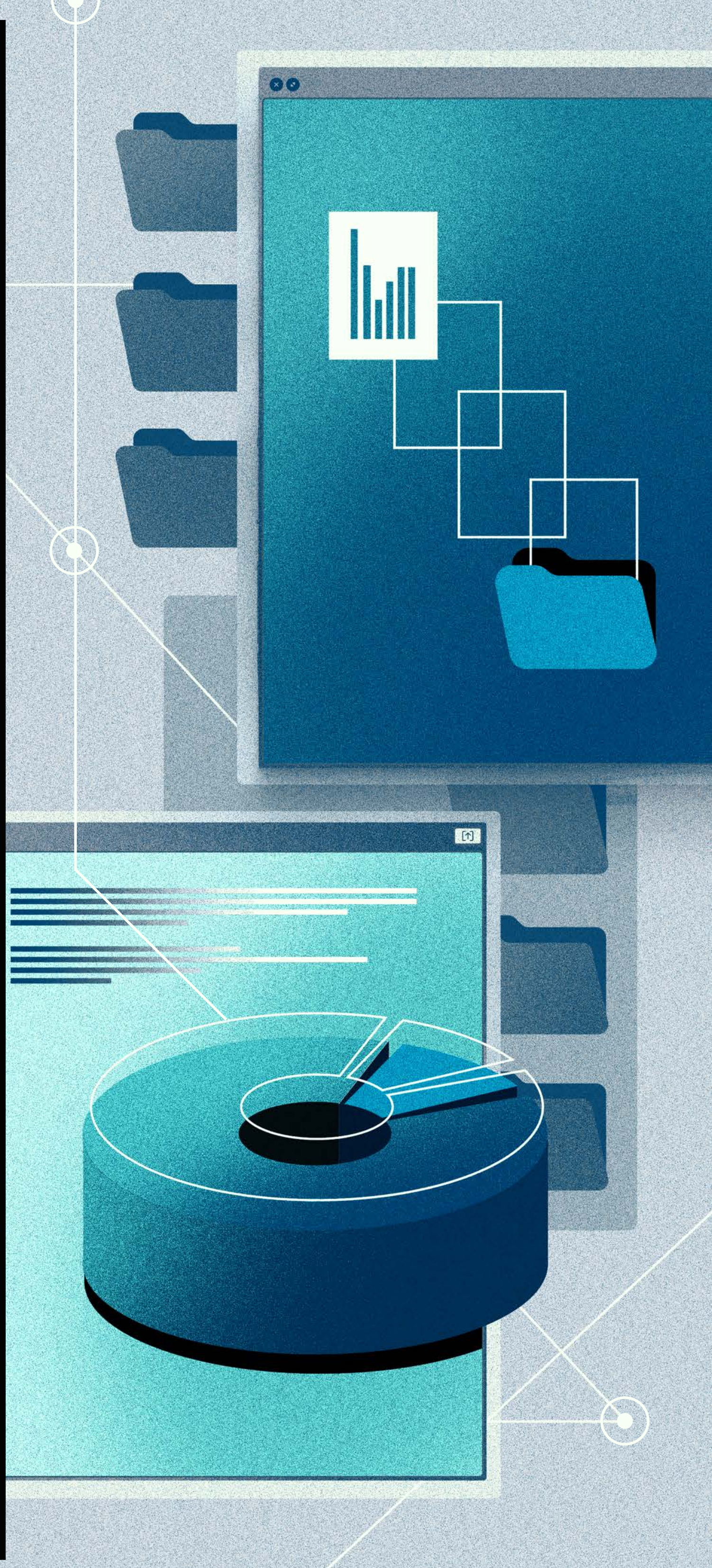
THE GREAT DATA HEIST

cyberhaven

Insider Risk Report

2022



THE GREAT DATA HEIST

Table of contents

Introduction

Every day, there are headlines about a hacking group conducting an attack, a familiar brand whose customers' information is for sale on the darknet, a new variant of malware, or a critical vulnerability in a popular product or service you urgently need to update to the latest version.

What all these things have in common is they involve threats from outside the organization. But some of the most common threats to an organization's data come from trusted insiders.

It's never been a better time to use your employer's data for personal gain. A record **47 million Americans quit their jobs in 2021**, many of them taking their knowledge (and files) with them to a competitor. Cyber criminals increasingly contact employees and offer them cash in exchange for information. Even leaking files to the media can be monetized. Former Facebook manager Francis Haugen inked a book deal shortly after she shared thousands of company documents.

NOT EVERY INSIDER RISK BECOMES AN INSIDER THREAT; HOWEVER, EVERY INSIDER THREAT STARTED AS AN INSIDER RISK.

Gartner

To better understand the risks to corporate data, we performed the analysis shared in this report. What makes this report unique is that unlike surveys that ask IT security professionals what they think about insider threats, we've compiled our findings using anonymized usage data precisely tracking how over 1.4 million people handle sensitive information at work.

The incidents we found cover many forms of data exfiltration but broadly they involve an employee either taking sensitive information outside the organization or sending it to someone outside of the organization they shouldn't. These insider risks, the culmination of multiple malicious or careless actions, are one step away from becoming an insider threat.

Once the information is outside the organization's control, anything can happen to it.

Key Findings

01 THE FREQUENCY OF INSIDER RISKS

Just **2.5%** of employees exfiltrate sensitive information in an average month, but in a six-month period nearly one in ten **(9.4%)** of employees do so.

Among employees that exfiltrated data, the top **1%** most prolific “super stealers” were responsible for **7.7% of incidents**, and the **top 10% were responsible for 34.9%.**

02 WHAT SENSITIVE DATA IS EXFILTRATED

44.6% of sensitive data that’s exfiltrated is client or customer data, **13.8%** is source code, and **8.0%** is regulated personally identifiable information (PII).

Just **17.9%** of exfiltrated data are the classic regulated data like PII, PCI, and PHI. Over **80%** of sensitive data exfiltrated is harder to identify intellectual property (IP).

03 THE PATH TO DATA BREACH

The majority of incidents **(53.8%)** involve data taking multiple steps across multiple people through the organization, not a user accessing it directly and then exfiltrating it.

Some sensitive information took as many as 42 steps, as it was shared, reshared, and moved between multiple systems before someone exfiltrated it.

04 HOW EMPLOYEES EXFILTRATE DATA

The most common exfiltration vectors are personal cloud storage **(27.5% of incidents)**, personal webmail **(18.7%)**, and corporate email **(14.4%)** to an inappropriate recipient.

The most popular personal cloud storage services used to exfiltrate data are **Dropbox (44.8% of exfiltration incidents using cloud storage)** and **Google Drive (25.5%).**

05 WHEN DEPARTING EMPLOYEES EXFILTRATE DATA

When an employee voluntarily quits, **68.7%** of the increased exfiltration before their departure occurs before they give notice when they are less likely to be monitored.

Employees who are fired are **23.1%** more likely to exfiltrate data the day before they were fired and **109.3%** more likely to exfiltrate data the day they are fired.

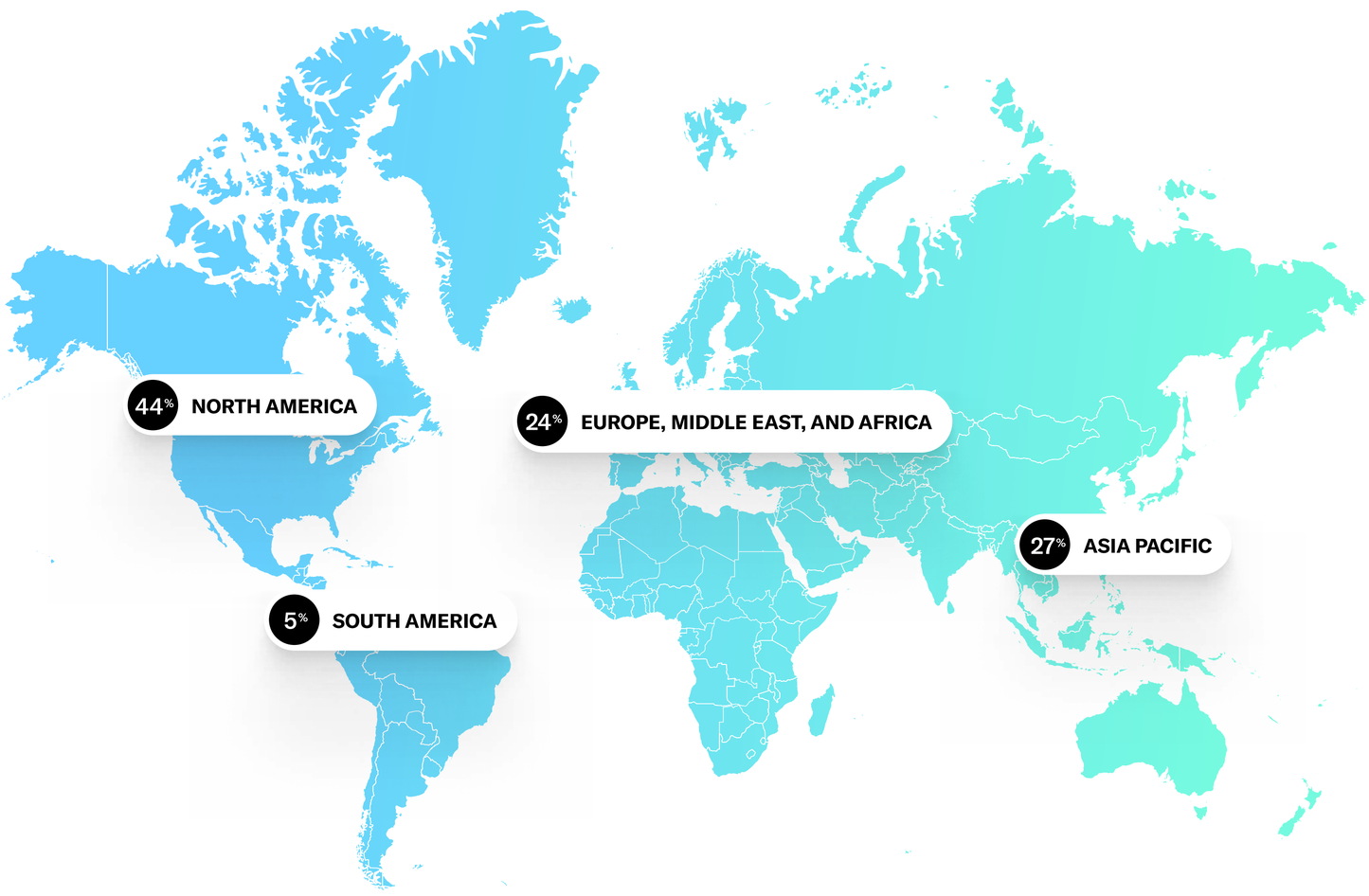
Methodology

To compile the findings in this report, we analyzed 372,000 data exfiltration incidents and anonymized usage data for 1.4 million workers from January 1, 2022 to June 30, 2022 using the Cyberhaven product. The companies in our sample represent 11% of the Fortune 100 and span multiple industries.

Top industries represented in the report

Biotech and pharma	10%
Chemicals	5%
Engineering	4%
Financial services	9%
Healthcare	11%
Legal	8%
Manufacturing	15%
Technology	13%

Geographic distribution in our sample



HOW WE CLASSIFIED DATA AS SENSITIVE

We used multiple signals to determine whether data is sensitive or not. Certain types of sensitive data such credit card numbers and Social Security numbers can be identified using pattern matching. Other forms of sensitive data are difficult or impossible to classify by looking at the content alone. We combined content inspection with contextual factors such as where the data originated from, and how it moved throughout the organization.

HOW WE DEFINE AN INCIDENT

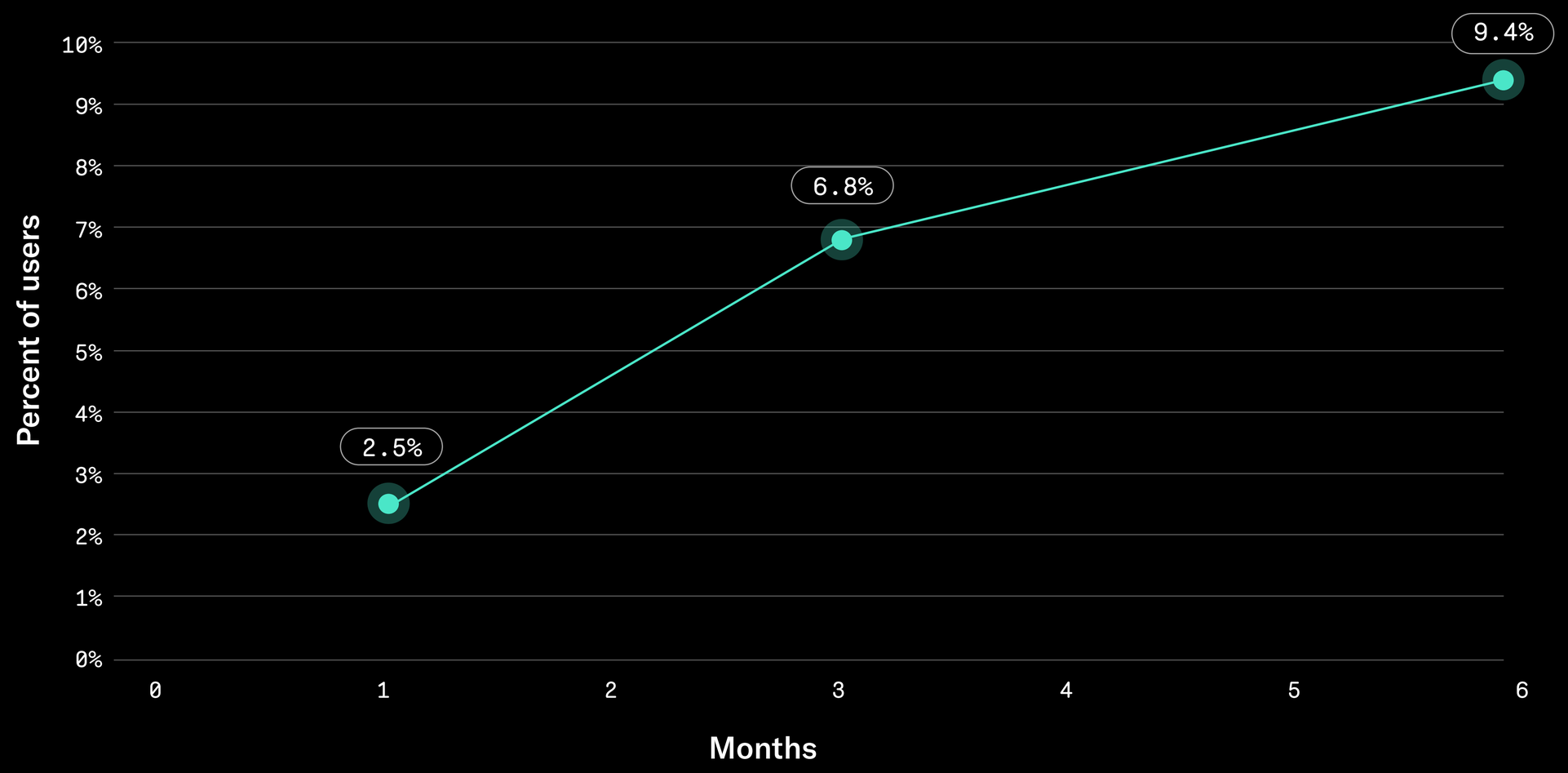
In multiple sections of this report we discuss the number of data exfiltration incidents. Data exfiltration, simply put, is transferring data outside the organization in unapproved ways. Exfiltrating sensitive data is risky, but it doesn't always turn into an insider threat. And not all insider threats are intentional or malicious, either. For example, employees sometimes email sensitive information to the wrong person outside the company by mistake, or they copy a sensitive document to a USB drive to work on it at home, only to lose the drive and its data.

The frequency of insider risks

A single data exfiltration incident can lead to an insider threat, but employees with multiple repeat offenses could also indicate malicious intent rather than simple mistakes.

Users are more likely to engage in insider risks over time

(Users with one or more incidents during time period)



We found that in any given month, 71% of employees who exfiltrated data did so only once. Among employees that exfiltrated data, the top 1% most prolific “super stealers” were responsible for 7.7% of incidents, and the top 10% were responsible for 34.9% of incidents.

The frequency of insider risks

We found that organizations experience, on average, 0.045 data exfiltration incidents per employee per month.

This number alone isn't very meaningful or easy to understand, but it gives organizations a way to estimate the number of incidents they might have. For example, extrapolating from this number, a 1,000 employee company experiences on average 45 data exfiltration incidents each month.

Employee count	Avg. incidents per month
1,000	45
5,000	225
10,000	451
20,000	902
50,000	2,254

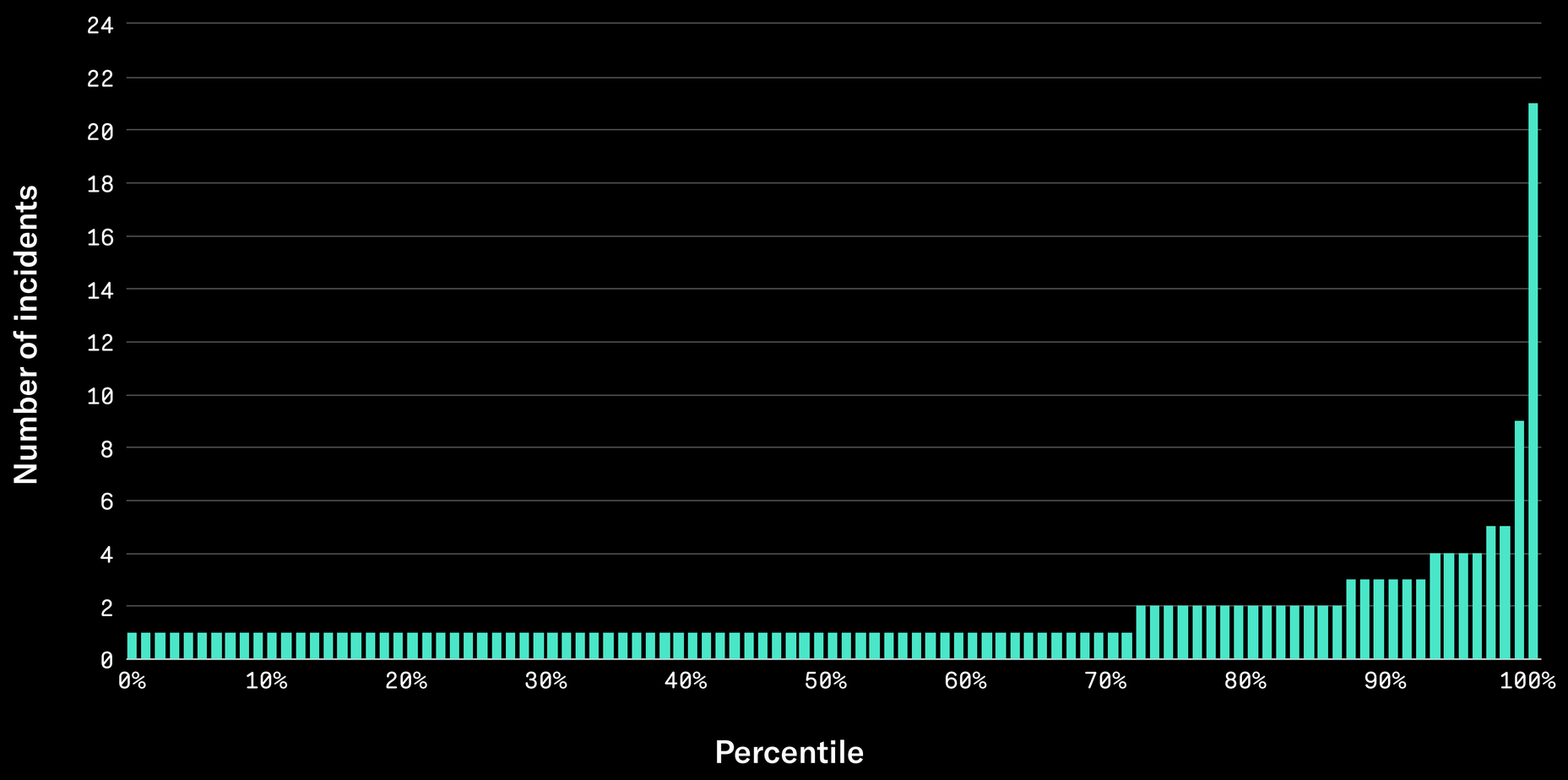
A relatively small number of employees exfiltrate sensitive information—just 2.5% of employees at an average organization are responsible for one or more incidents during a one-month period. But across a longer time horizon, we found data exfiltration is more widespread across a broader group of employees. In a three-month period, 6.8% of employees exfiltrate sensitive data and in six months the number grows to 9.4% or nearly one in ten.

The frequency of insider risks (cont.d)

Remember that only 2.5% of all employees exfiltrate data each month, so across all employees just 0.025% are responsible for 7.7% of incidents and 0.25% are responsible for 34.9% of incidents.

Distribution of data exfiltration incidents across users

(Per user per month normalized by user percentile)



At a 50,000 person company that works out to 12 employees and 125 employees, respectively. The number of serial offenders is small enough that an organization could investigate each one, while relying on automation to coach users who make one-off mistakes.

What sensitive data is exfiltrated

The top 10 types of sensitive data employees exfiltrate includes:

1

CLIENT/CUSTOMER DATA

Example: a spreadsheet exported from NetSuite showing all customers and the dollar amount they've paid over the past year, an M&A plan that a publicly traded client of an investment bank shared with the bank's deal team, etc.

2

SOURCE CODE

Example: the code used in a social media app to determine what content to show a user in their feed, the algorithm a buy-now-pay-later company uses to determine whether to extend credit to a customer at checkout, etc.

3

REGULATED PERSONAL DATA (PII)

Example: a California customer's name and mailing address stored by an e-commerce company as part of their order tracking system, a German user's date of birth stored by a social media company, etc.

4

DESIGN FILES AND PRODUCT FORMULAS

Example: a 3D CAD file with the parts and assembly of a LiDAR sensor package for a self-driving car in development, the recipe and production process for a popular candy bar, etc.

5

REGULATED HEALTH DATA (PHI)

Example: the medical record of a celebrity who was checked into the hospital following a serious car accident, a CSV file downloaded from an insurance billing application containing patient names and diagnostic codes, etc.

6

REGULATED FINANCIAL DATA (PCI)

Example: a customer's credit card number stored in a billing application for a water utility in order to process recurring payments, a folder with new customer signup forms containing bank account and routing numbers, etc.

7

SENSITIVE PROJECT FILES

Example: a folder of images taken with the unannounced smartphone that could be analyzed to reveal specifications of the new camera, an unreleased movie stored on a share drive by a production house that makes movie trailers, etc.

8

COMPANY CONFIDENTIAL

Example: an internal report that found use of the company's service increases the rate of depression in teens, an email thread between executives discussing how to handle an upcoming regulatory action by the government, etc.

9

UNRELEASED OR SENSITIVE MARKETING

Example: an unreleased press release in Google Docs with details about the company's upcoming product announcement, ad creative being developed in Figma with imagery of the company's unannounced product, etc.

10

EMPLOYEE HR DATA

Example: a spreadsheet downloaded from Workday with the salaries of all employees in the department, a document containing details about the planned bonus payouts for executives, etc.

Client or customer data is most at risk

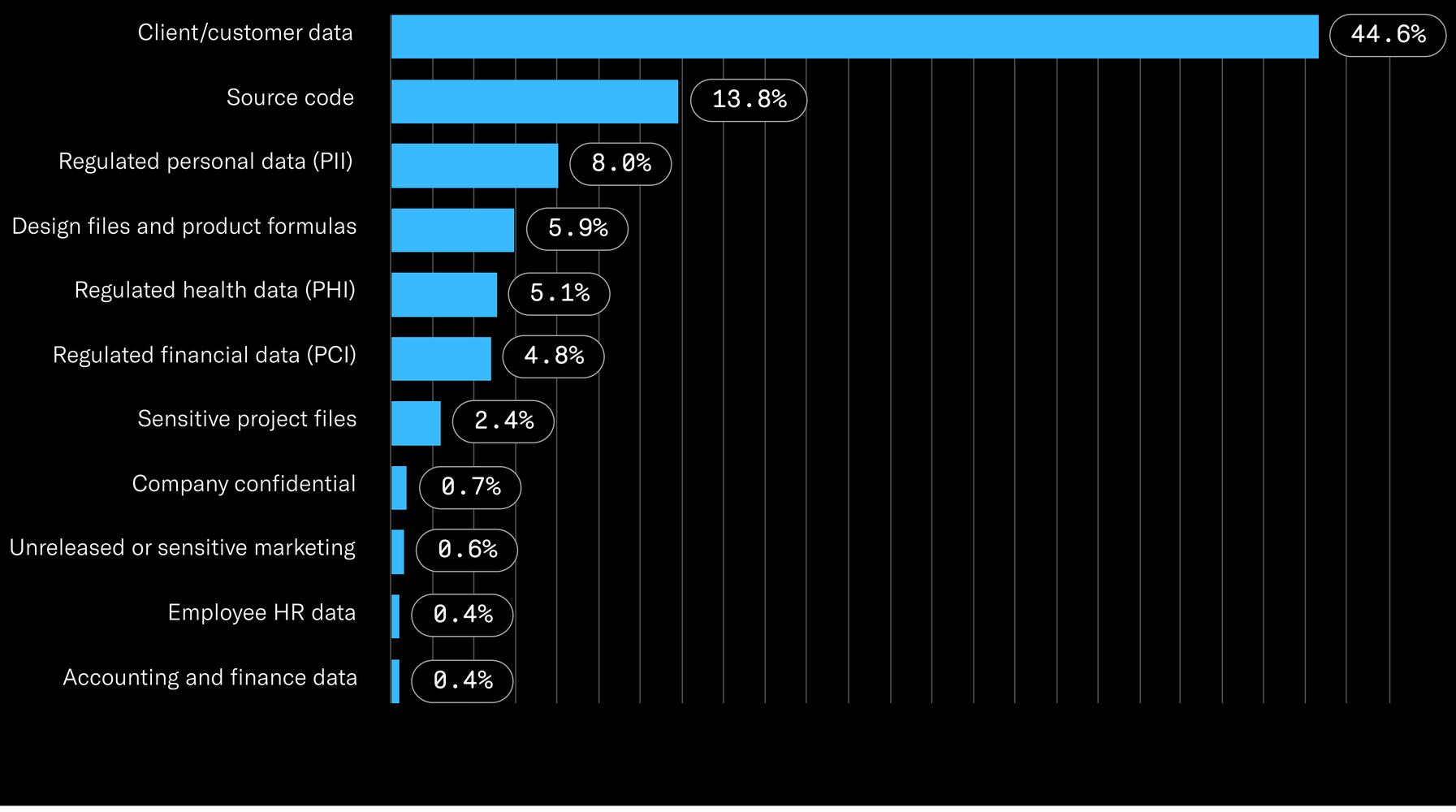
Comprising 44.6% of the sensitive data employees exfiltrate. Modern enterprises are awash in information about their customers and files from their customers. One possible explanation is that employees do not understand the sensitivity of this information in the same way they do for, say, a product formula or a medical record.

The second most at-risk data is source code

Source code accounts for 13.8% of exfiltrated data. These incidents don't just occur at software companies. Today, companies across verticals including airlines, retail, financial services, and manufacturing all develop their own applications and algorithms, which they use to gain a competitive advantage. Losing their source code to a competitor can have a material impact on their businesses.

Top types of sensitive data employees exfiltrate

(By volume of data)



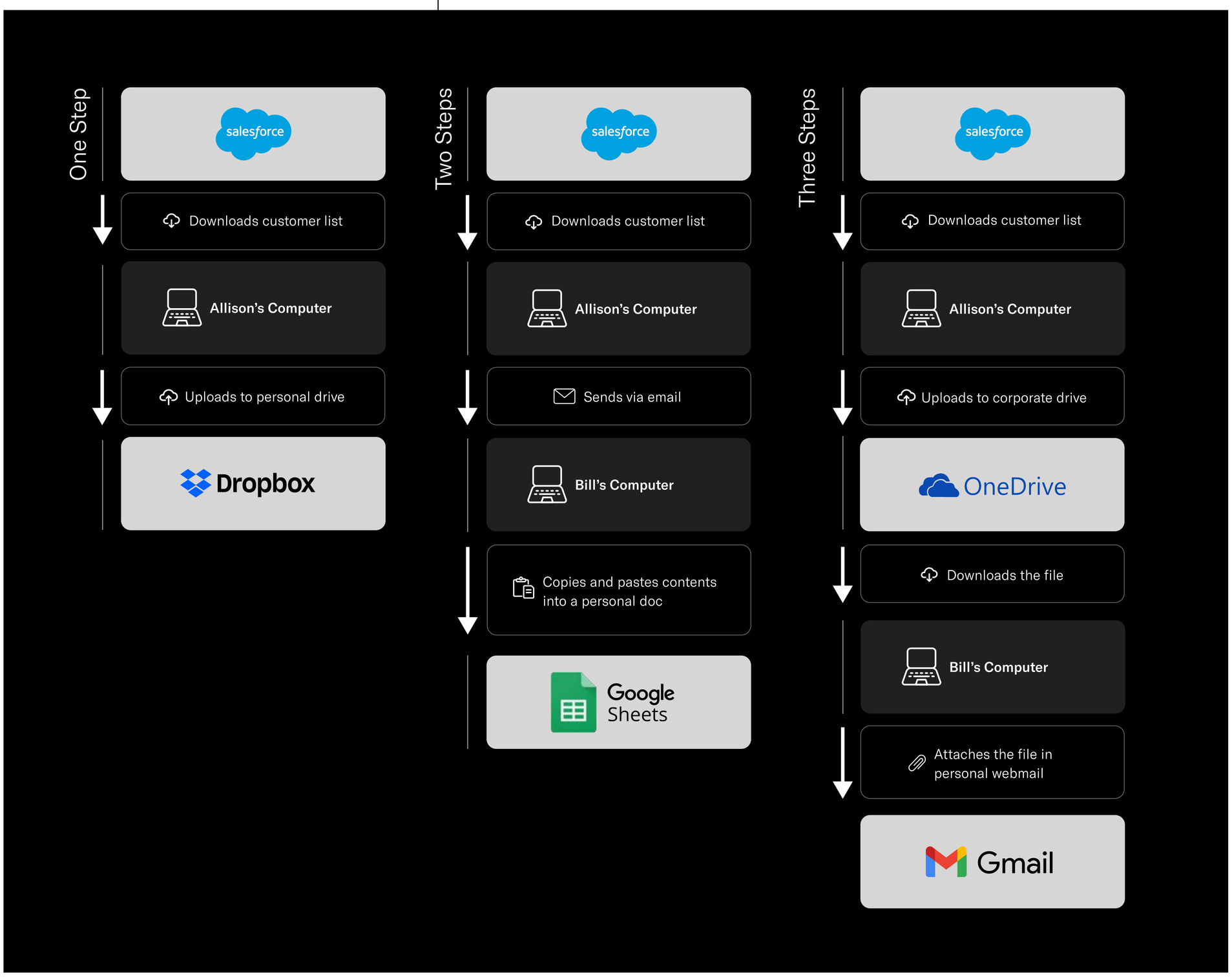
Regulated data, including personally identifiable information (PII), payment card information (PCI), and protected health information (PHI) collectively account for just 17.9% of exfiltrated data. This information, which often includes a standard alphanumeric pattern, has historically been easier to classify using software and therefore easier to protect. Our analysis finds that over 80% of exfiltrated data is harder-to-identify intellectual property (IP).

The path to a data breach

As seen in the examples above, most sensitive data originates in specific systems or locations.

However, it has a tendency to move throughout the organization, in many cases spreading beyond the people who have permission to access it at the source. Less than half of data exfiltration incidents involve an employee accessing data directly and then exfiltrating it. The majority of incidents (53.8%) involve data moving two or more steps before it is exfiltrated.

Let’s take a look at a few examples to see the path a customer list downloaded from Salesforce can take through the organization one, two, or three steps before being exfiltrated:



Some sensitive data takes as many as 42 steps as it circulates within an organization before someone exfiltrates it. As data naturally and constantly spreads, it moves to places it is less likely to be tracked and protected. The challenge facing security teams is they need to protect data as it moves beyond the systems where it is being protected today.

How employees exfiltrate data

Employees remove sensitive information from the organization in a variety of ways.

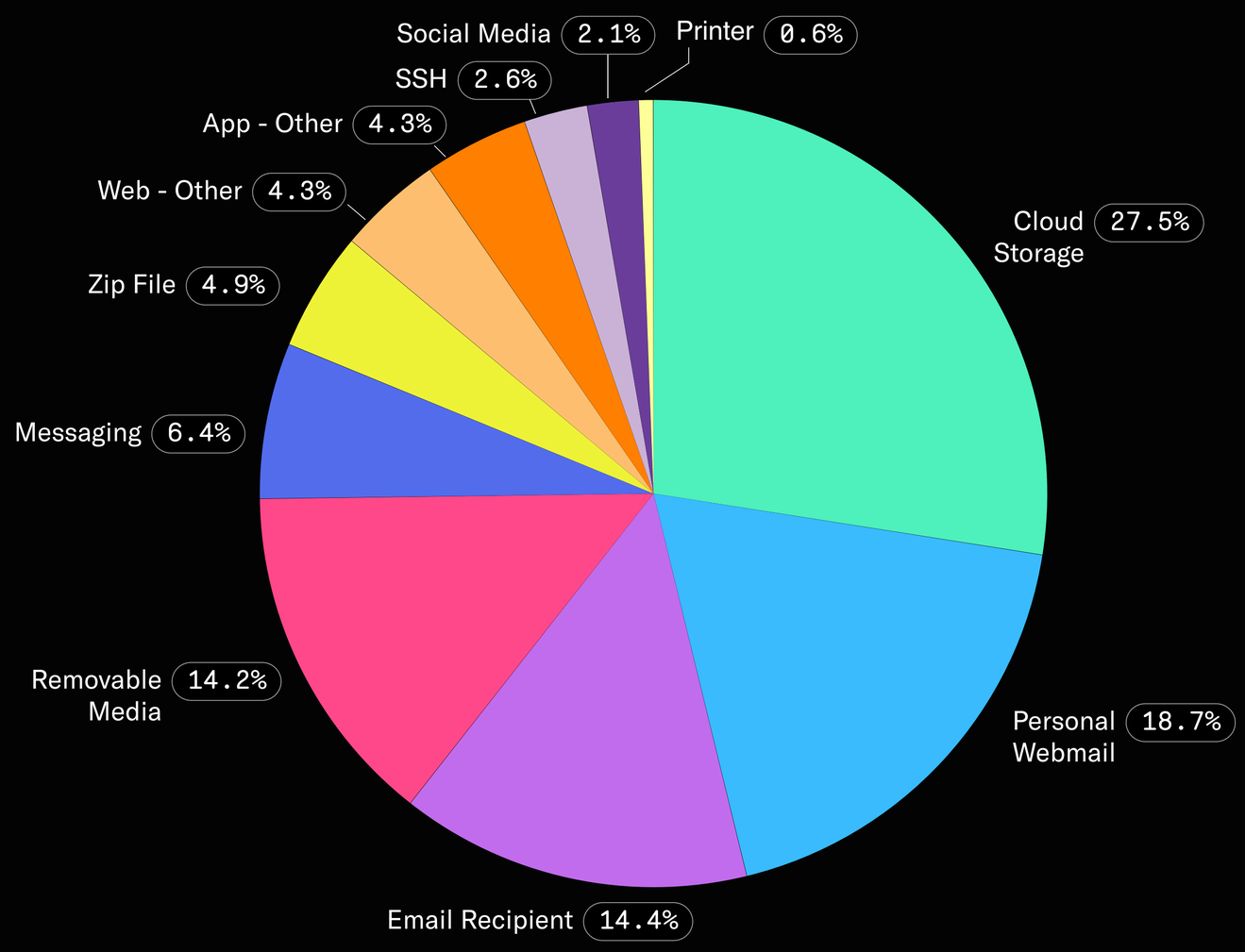
The top exfiltration vector is personal cloud storage, which accounts for 27.5% of all incidents. The next most common is personal webmail (18.7%), where the user attaches a sensitive file to an email or copy/pastes sensitive data into the email body or subject. Exfiltration via corporate email (14.4%) can include employees emailing sensitive data to their personal email addresses from their work account or employees accidentally sending sensitive information to the wrong recipient like when their email client auto-completes the addressee and in a rush they send it.

Along with personal cloud storage, removable media (14.2%) such as USB storage drives make it easy to exfiltrate a large amount of data at one time. For instance, some employees copy and paste the entire contents of their computer to an external hard drive. Messaging applications (6.4%) include apps like Whatsapp and Signal and they are a growing concern because their use of end-to-end encryption makes it difficult for organizations to know what's being sent with them.

The two most popular personal cloud storage services used to exfiltrate data are Dropbox (44.8%) and Google Drive (25.5%). The native sync clients of both services utilize certificate pinning, which is a security feature meant to prevent a malicious third party from intercepting and decrypting traffic to their clouds. Like encrypted messaging apps, it has the side effect of making corporate network security tools blind to what employees are uploading to them.

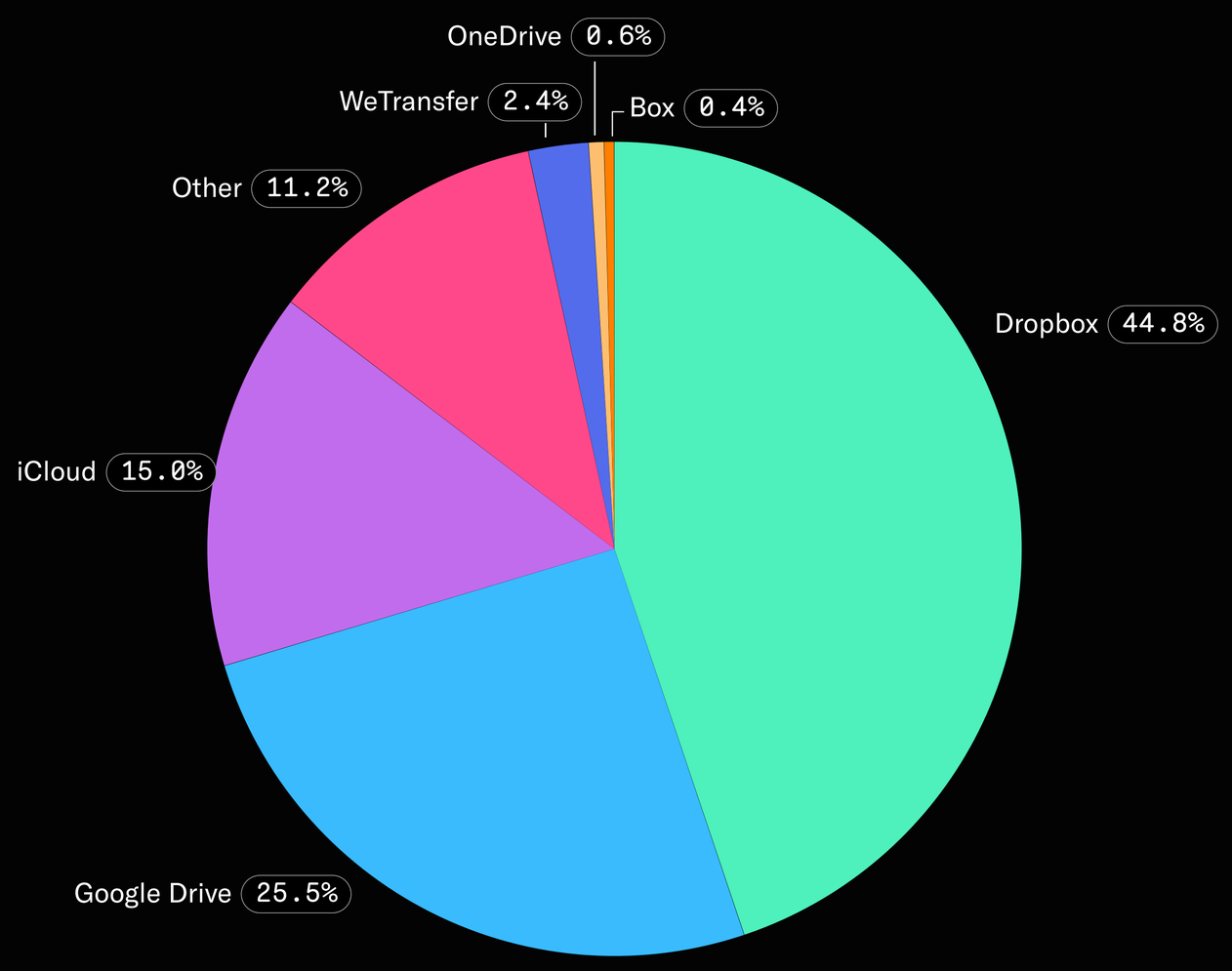
Methods used to exfiltrate sensitive data

(By volume of data)



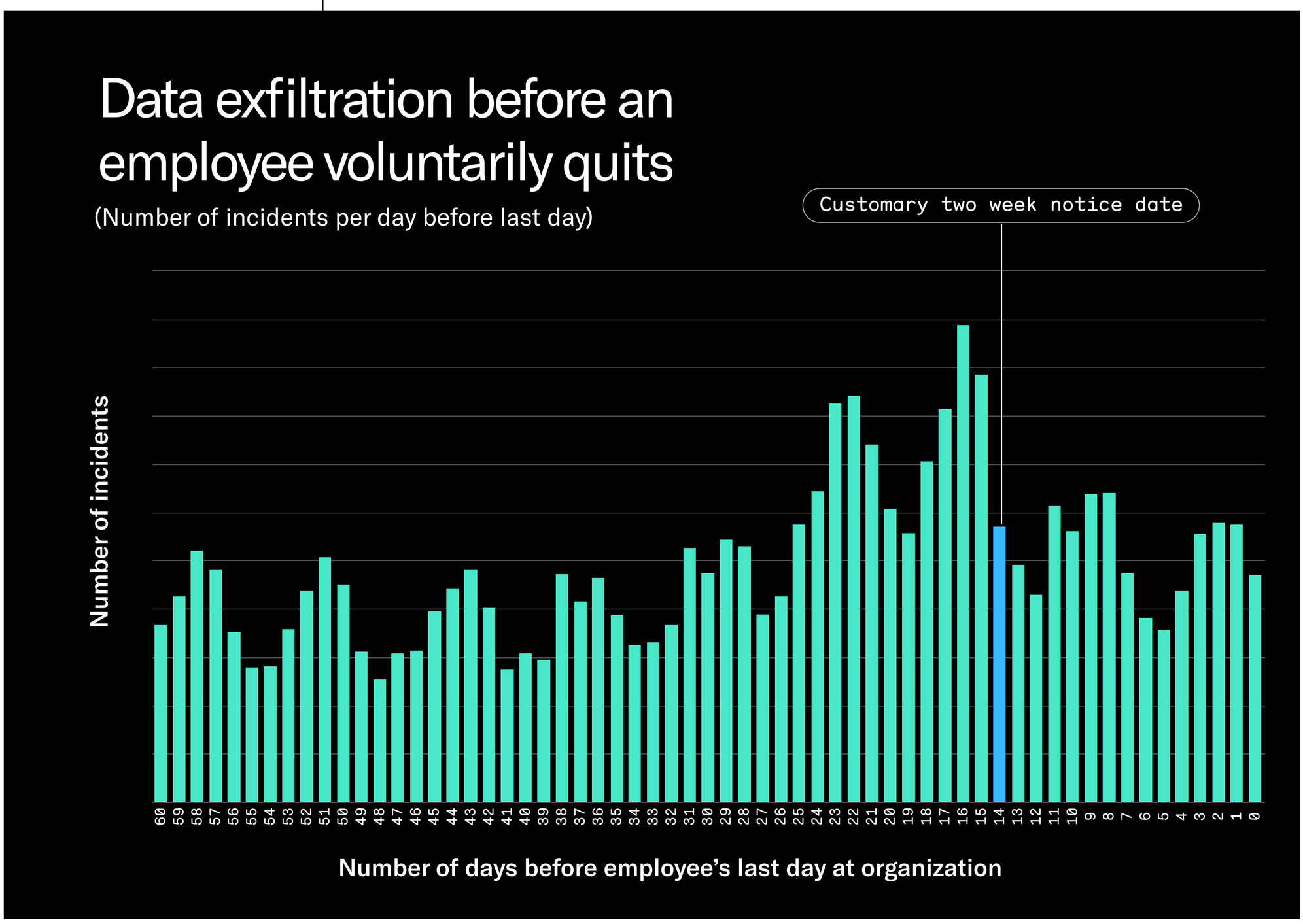
Cloud storage services used to exfiltrate sensitive data

(By volume of data)



When departing employees exfiltrate data

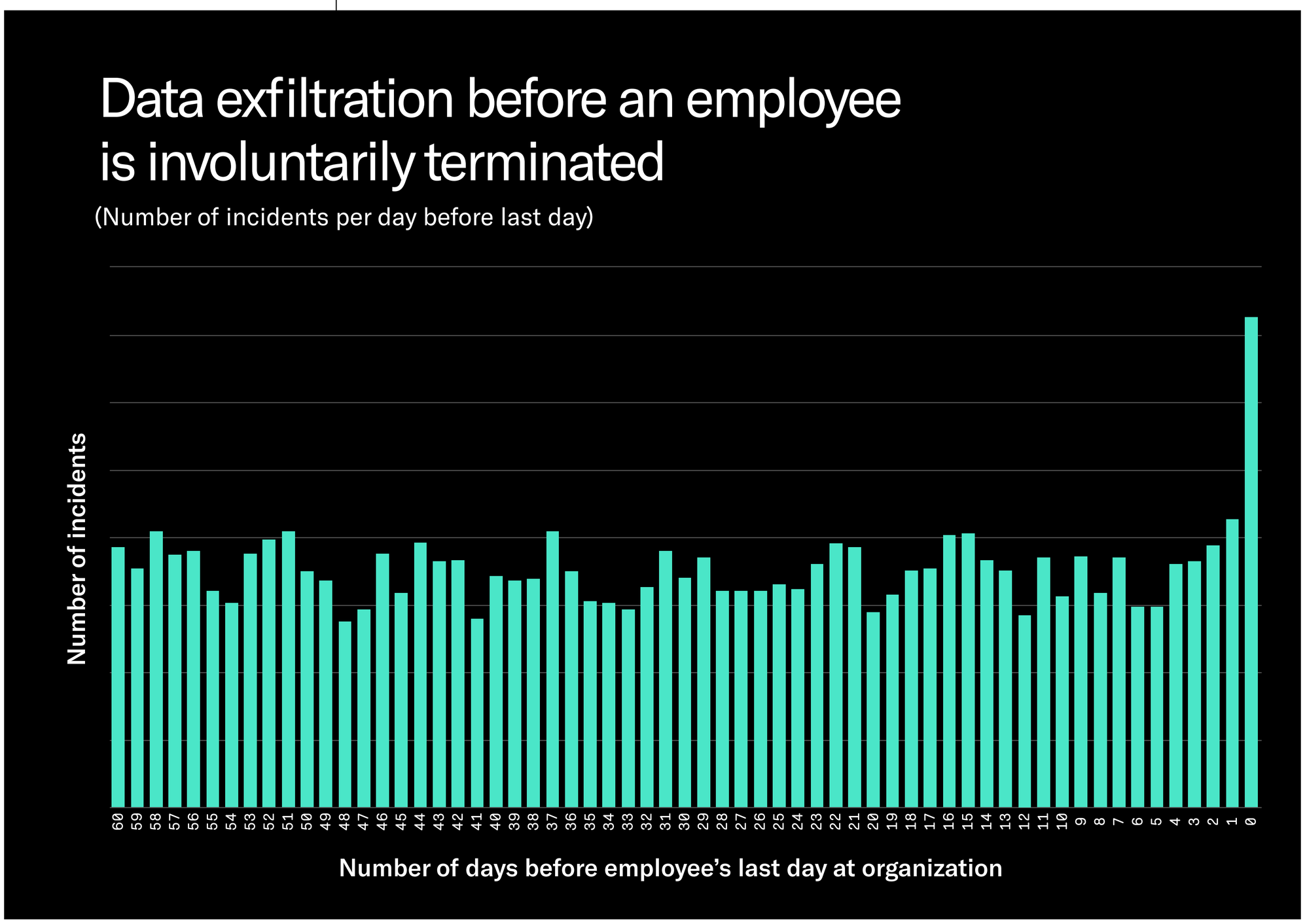
It’s not uncommon for corporate HR departments working with security teams to monitor the activity of employees who quit starting when they announce they’re quitting and lasting until their last day. One reason is to ensure employees don’t take any sensitive information with them as they depart. In the United States, employees typically provide two weeks’ notice. However, we found a noticeable increase in data exfiltration occurring before the employee gave notice, when the employee knows they intend to leave but their employer does not.



During the period between when an employee gives notice and their last day, we measured a 37.7% increase in the number of data exfiltration incidents compared with the baseline. However, during the two week period before the employee gave notice, we found an 83.1% increase in incidents. Of the increase in data exfiltration before an employee voluntarily departs, 68.7% occurs before they notify the company, when they are less likely to be monitored.

When departing employees exfiltrate data (cont.d)

You would not expect to see a similar spike in data exfiltration preceding an involuntary termination under the assumption most employees don't know they're about to be fired. But we found a 23.1% increase in data exfiltration from employees the day before they were fired and a 109.3% increase the day they were fired. It appears some employees find out or sense their impending dismissal and decide to collect sensitive company data for themselves, and others may be notified they're terminated and collect data before their access is turned off.



Get a personalized audit of your insider risk today

We'll analyze how employees handle data with Cyberhaven's Data Detection and Response platform and show you:

1 HOW REGULATED DATA AND IP MOVES THROUGH YOUR ORGANIZATION

2 WHO IS EXFILTRATING SENSITIVE DATA AND HOW THEY'RE DOING IT

3 WHAT DEPARTING EMPLOYEES DO BEFORE THEY JOIN A COMPETITOR

Request an assessment now

Cyberhaven.com/data-risk-assessment

"CYBERHAVEN HAS BEEN AN INTEGRAL PART OF BUILDING OUR INSIDER RISK PROGRAM TO GAIN VISIBILITY INTO POTENTIAL BAD BEHAVIOR AND PUT IN PLACE PREVENTATIVE CONTROLS TO STOP IP FROM LEAVING THE COMPANY,"

Gary Huber, Sr. Director of IT Security

CRESTRON

cyberhaven

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

www.cyberhaven.com