

REPORT

# The art (and science) of behavior change in human risk



# Table of contents

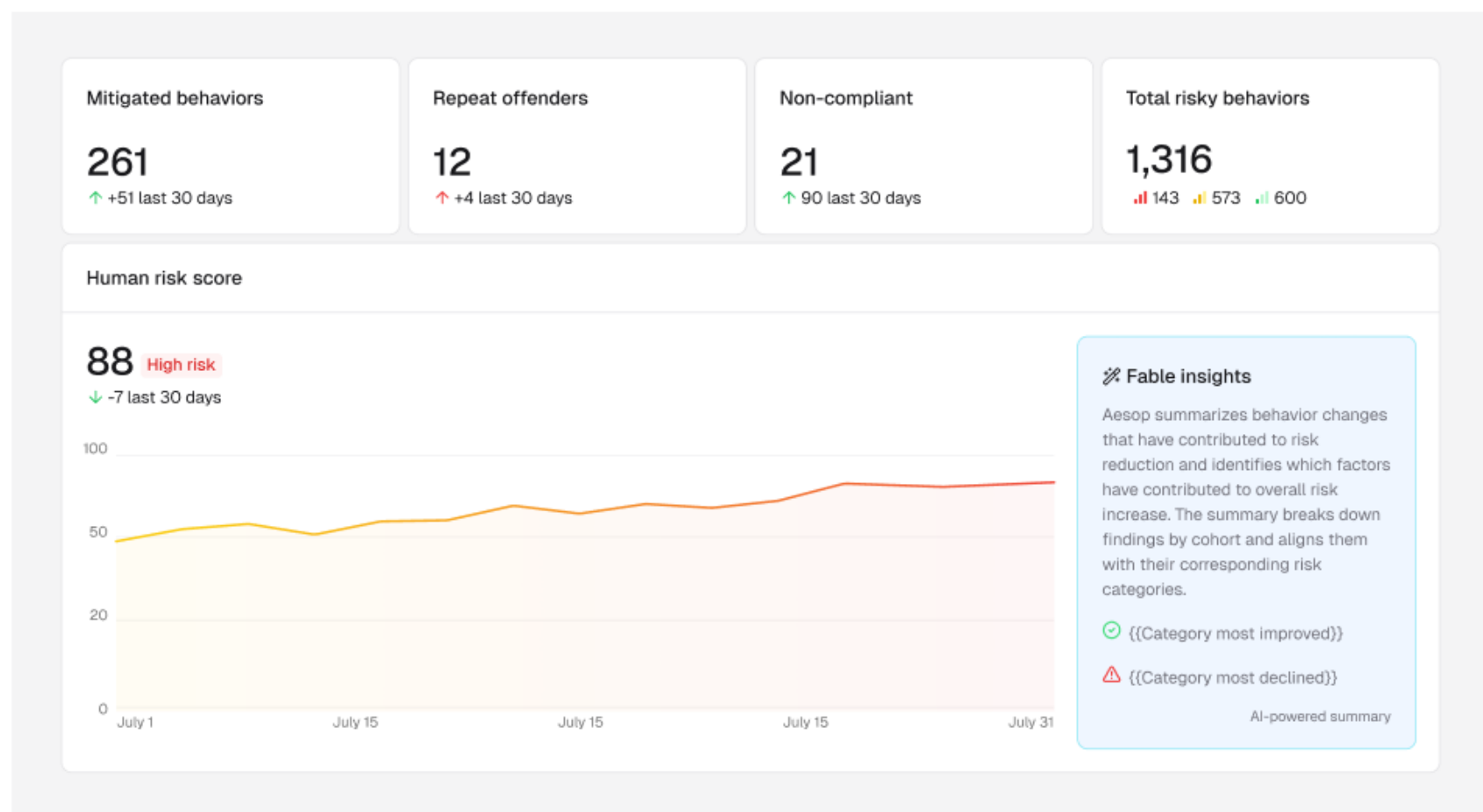
Explanation	01
Core human risk metrics	02
Ten common behavior contributors to human risk	03
Campaign maturity model	04
Targeting delta	07
Behavior change	08
Time to behavior change (TTBC)	09
Cohort comparisons	10
Behavior decay interval	11
AI swashbucklers	12
Toxic combinations	13

# Explanation

This report marks the start of an ongoing exploration into how organizations measure, understand, and reduce human risk. Here, we introduce the foundational metrics that show how employee behaviors shape an organization's security posture. Through anonymized, real-world case studies in specific customers or across customers, we bring these metrics to life. Over time, we'll evolve them into industry-wide trends and benchmarks—but for now, this report serves as the opening chapter. The data in this report reflects activity from hundreds of thousands of employees across dozens of customers through October 31, 2025. Campaign examples draw from various periods in the preceding year.

# Core human risk metrics

The core metrics security teams track and report on in Fable include a single risk score at the individual, departmental, and organizational level; top risk factors; and specific risky behaviors that comprise those factors. They report on the status of training completion and phishing simulation performance, as well as risky behaviors that have been mitigated with interventions delivered in Fable.



# Ten common behavior contributors to human risk

Security teams track dozens of risky behaviors, depending on the workspace and security signals they collect as part of their risk analysis.

Here are 10 common ones we see again and again across our customer environments.



Weak, reused, or shared credentials



Over-provisioned access (time or privileges)



Failure to rotate credentials exposed in a breach



Unpatched OS software



Weak MFA in critical applications



Exposure of sensitive information in generative AI or cloud applications



Secrets in code or private information in cleartext



Susceptibility to social engineering



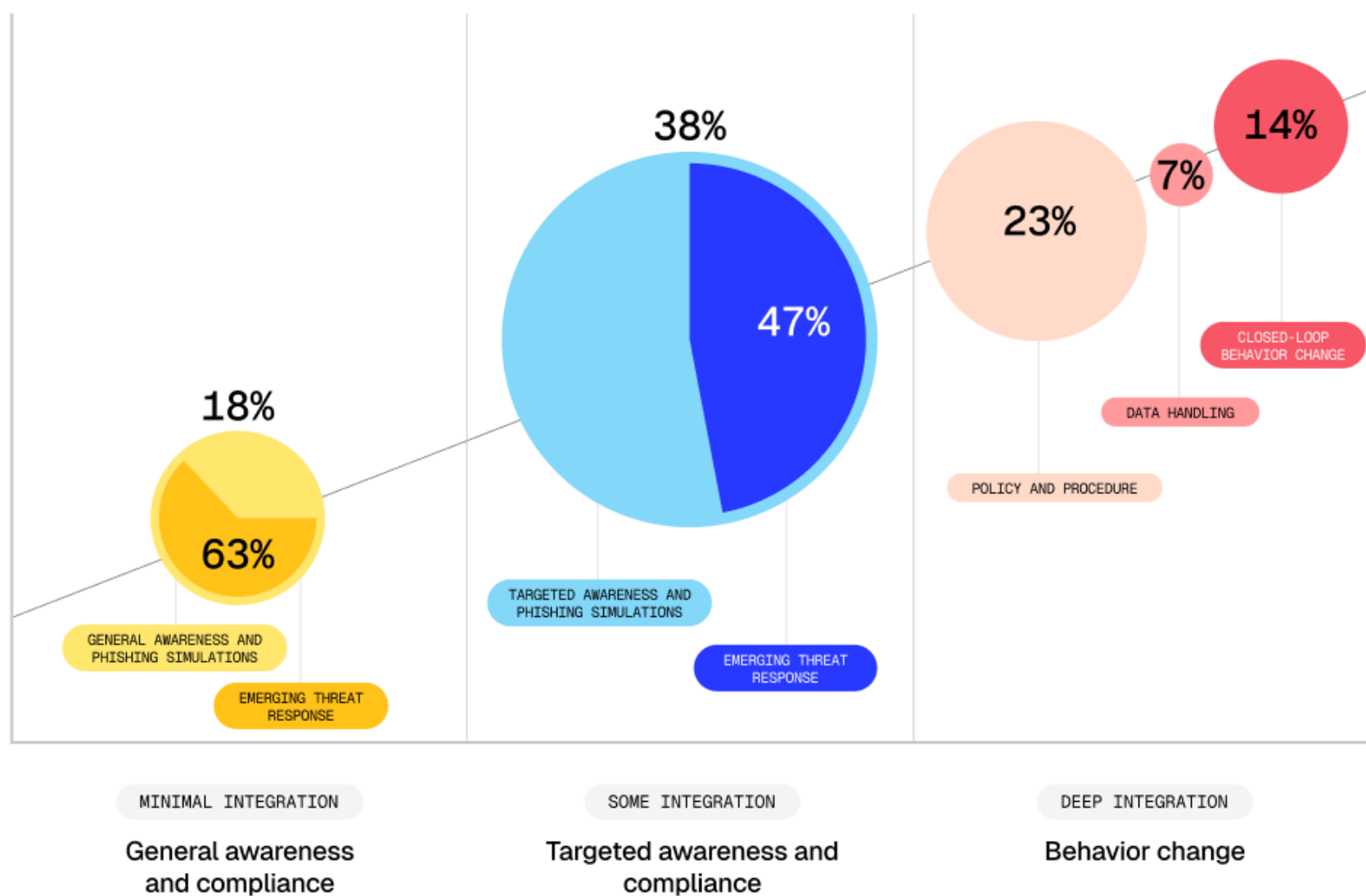
Oversharing of personal information online



Unsafe websites or browser extensions

# Campaign maturity model

Security teams use Fable to achieve objectives ranging from raising general awareness to raising awareness for targeted groups to shaping specific behaviors to reduce risk. Below, we group these campaigns by type and goal and place them on a maturity map aligned with integration depth and targeting level.



Starting on the left, 18% of our customers’ campaigns are in the general compliance category—security awareness and phishing simulations sent to everyone in the organization. Even though the campaigns are general, the topics aren’t generic: 63% are in response to an urgent threat, such as an emerging ransomware attack, impersonation scam, or social engineering tactic.


18%

OF CUSTOMER CAMPAIGNS ARE GENERAL “COMPLIANCE”

63%

OF CAMPAIGNS ARE IN RESPONSE TO AN URGENT THREAT

Fable



How to spot deepfakes

Made for you an hour ago

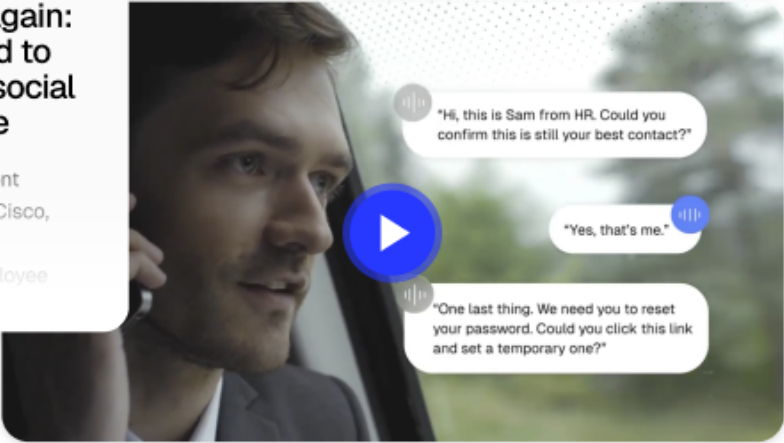
In the middle, where customers have integrated their directory or workspace platform such as Google Workspace or Microsoft 365, 38% of campaigns are somewhat targeted, focusing on cohorts based on role, access, or risk. Of those, 47% are in response to an urgent threat, such as ShinyHunters scammers targeting customer database administrators.

CSO

BY SHWETA SHARMA • AUGUST 19, 2025

ShinyHunters strike again:  
Workday breach tied to  
Salesforce-targeted social  
engineering wave

Experts say the attack mirrors recent breaches at Google, Pandora, and Cisco, revealing a coordinated campaign exploiting CRM platforms and employee



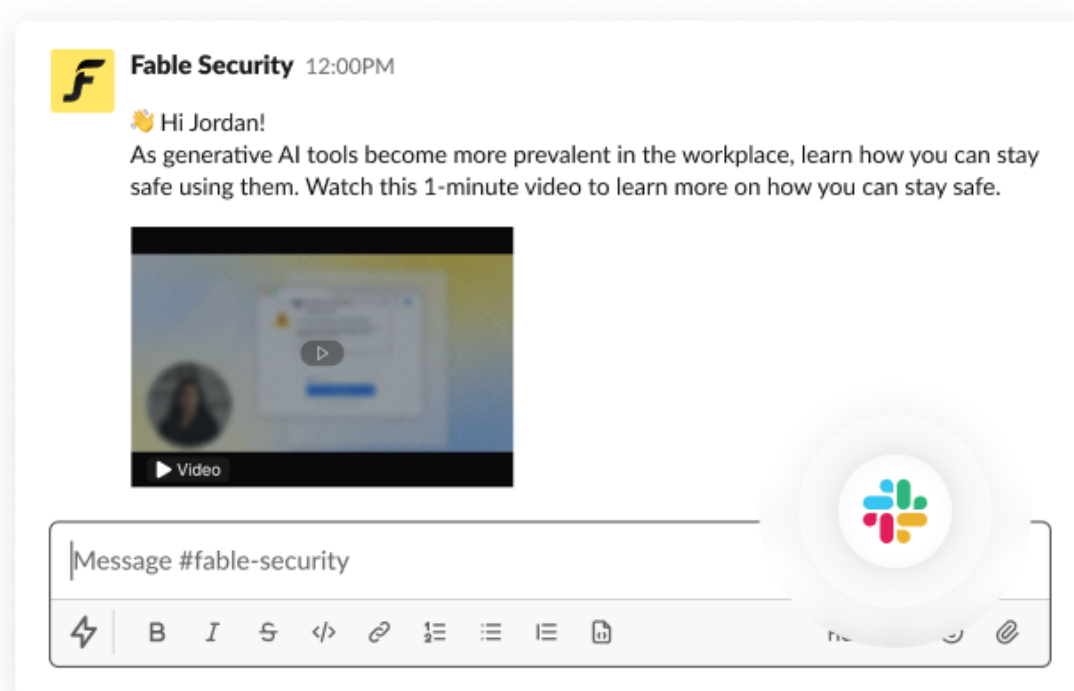
05 THE ART (AND SCIENCE) OF BEHAVIOR CHANGE IN HUMAN RISK



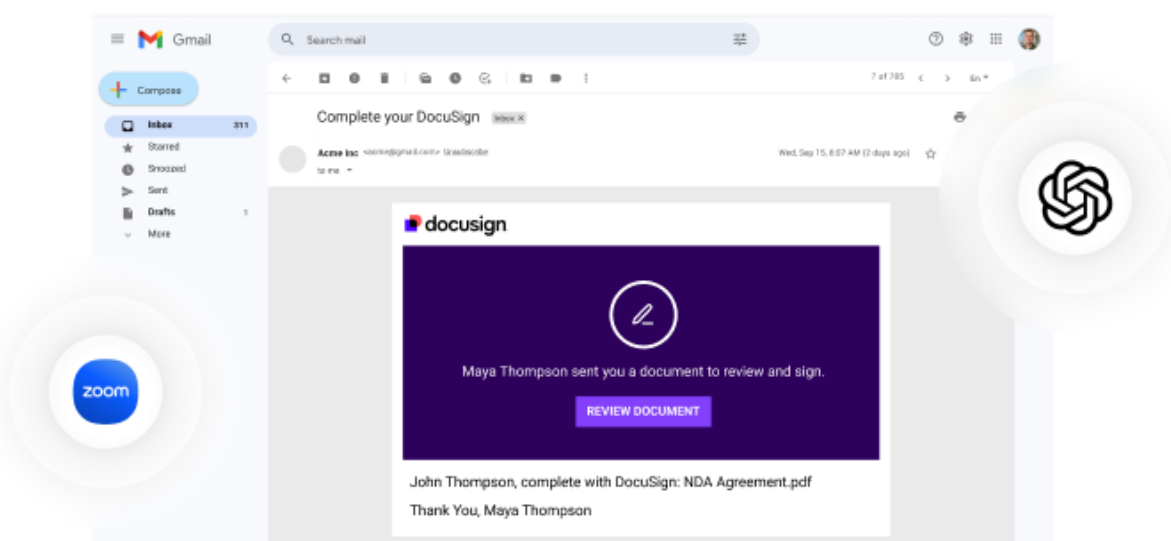
On the right, where customers have integrated Fable more deeply into their technology stacks—such as with single sign-on like Okta, enterprise browser like Google, SASE like Netskope, and endpoint detection and response like CrowdStrike—they’re running highly-targeted campaigns to shape behavior. Examples include prompting employees not to upload sensitive content to unsanctioned generative AI, to rotate credentials when they’ve been exposed in a breach, and to comply with security protocols such as enabling multi-factor authentication or adopting a password manager.

24%

OF CAMPAIGNS  
RAISE AWARENESS  
ABOUT TECHNOLOGY  
AND BRAND  
IMPERSONATIONS



Of the campaigns designed to raise awareness about an issue or emerging threat, 24% are brand impersonations of commonly-used technologies. The top three faked technologies are e-signing such as DocuSign, generative AI such as ChatGPT, and video conferencing such as Zoom.





# Targeting delta

Many security teams are experimenting with targeting to understand what campaign elements drive change. In our experience, the more targeted the campaign, the better it performs in terms of engagement and action. In this simplified example, we compare two campaigns of roughly the same duration and topic—one sent to a cohort and the other to the whole company. The targeted campaign performed 33 percentage points better than the general one. We expect to see even more dramatic differences as we study the additional meaningful ways our customers target.

## Targeted campaign



74%

ENGAGEMENT RATE

## General campaign

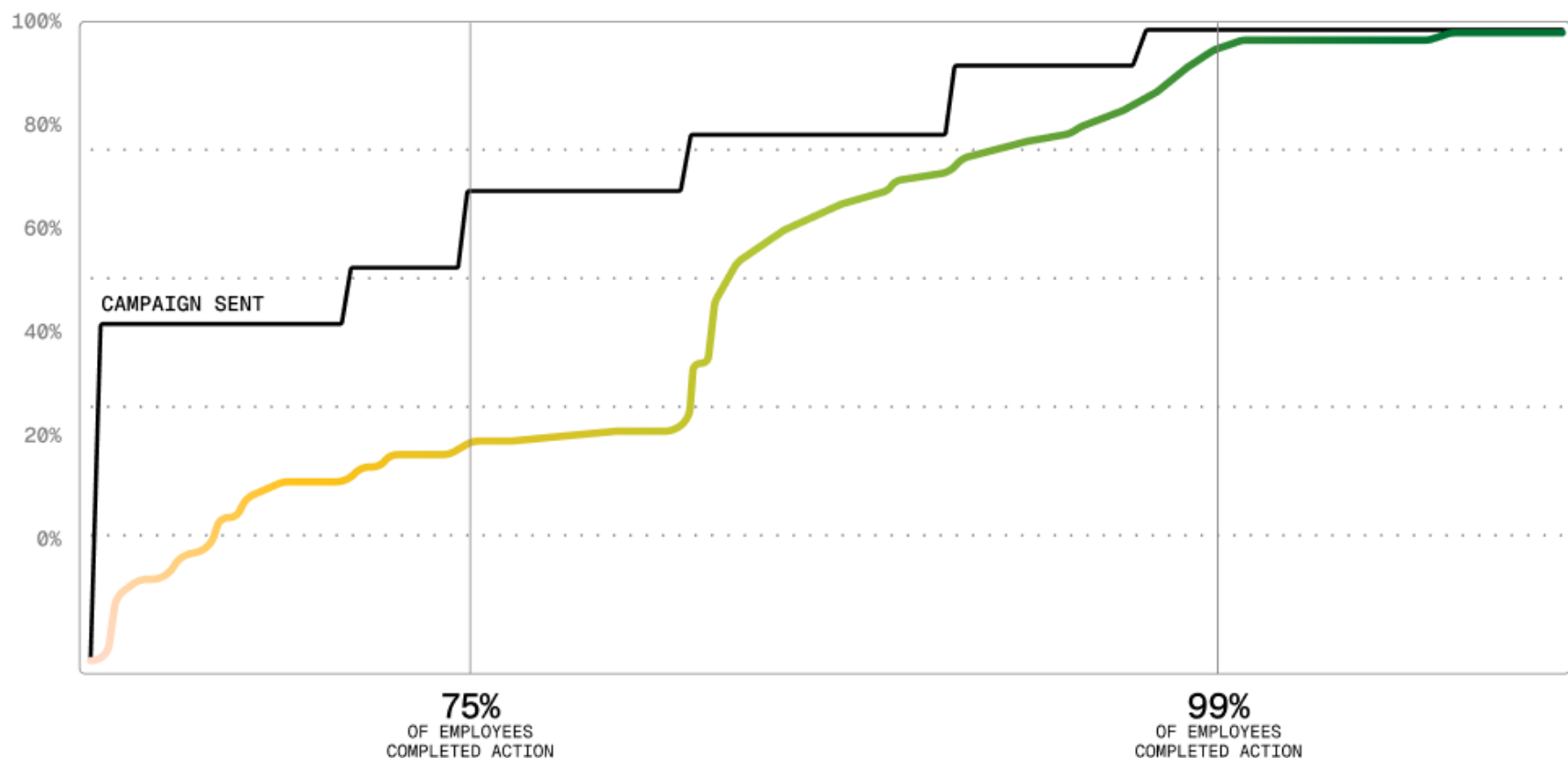


41%

ENGAGEMENT RATE

# Behavior change

While some security teams are focused on campaign engagement (such as video views in phishing campaigns), others use Fable to drive actual behavior change, such as adopting a password manager or discontinuing the upload of sensitive data into generative AI applications. Below is an example of how a security team used briefings and nudges to drive compliance of device OS updates. While most organizations stop at “video completed,” this included a critical next step: “action completed.” In other words, did the person complete the requested behavior change? In this case, the campaign reached 75% behavior change within the cohort after about two weeks, ultimately leveling off at 99% steady-state compliance in the weeks following.

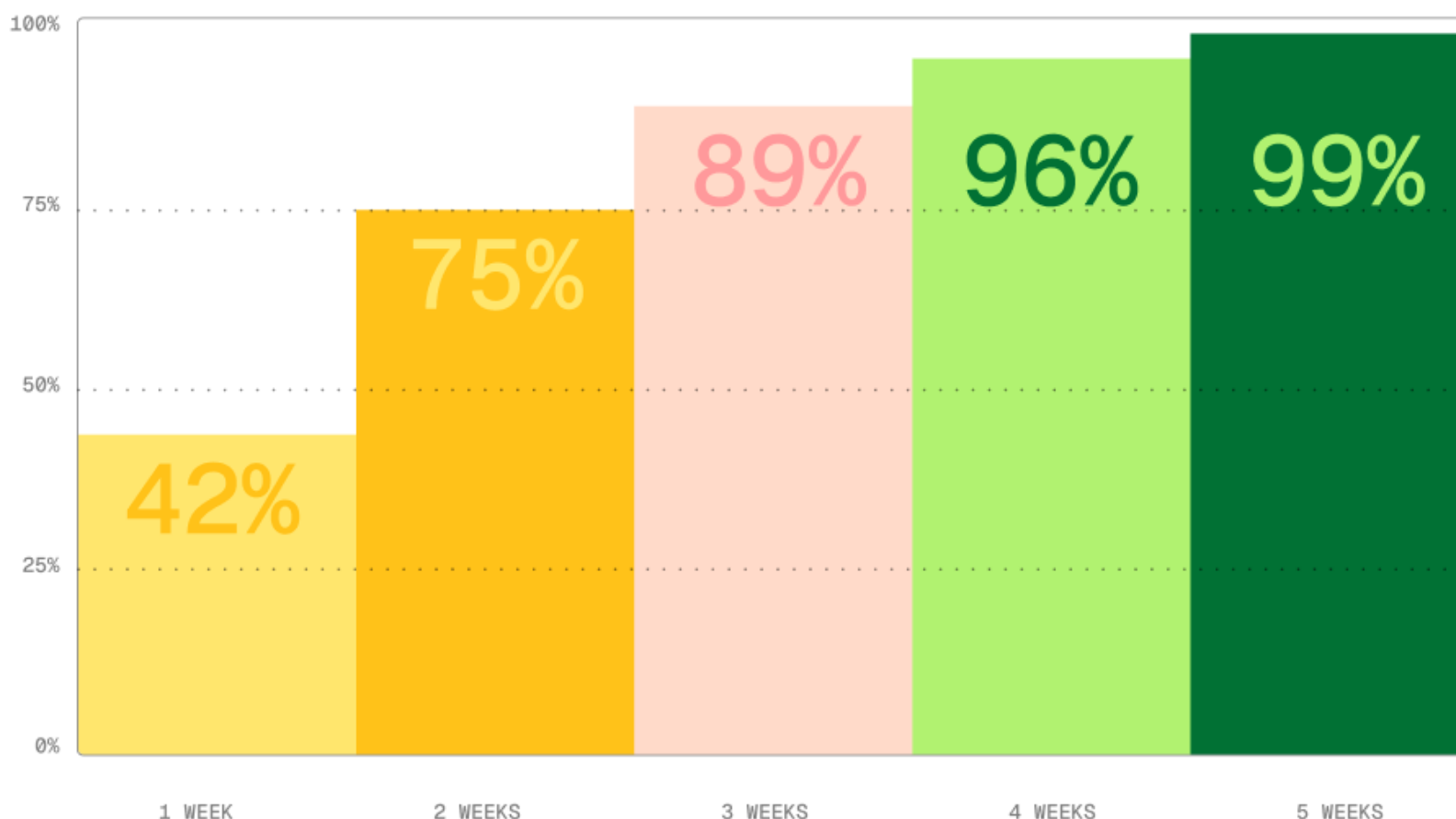




# Time to behavior change

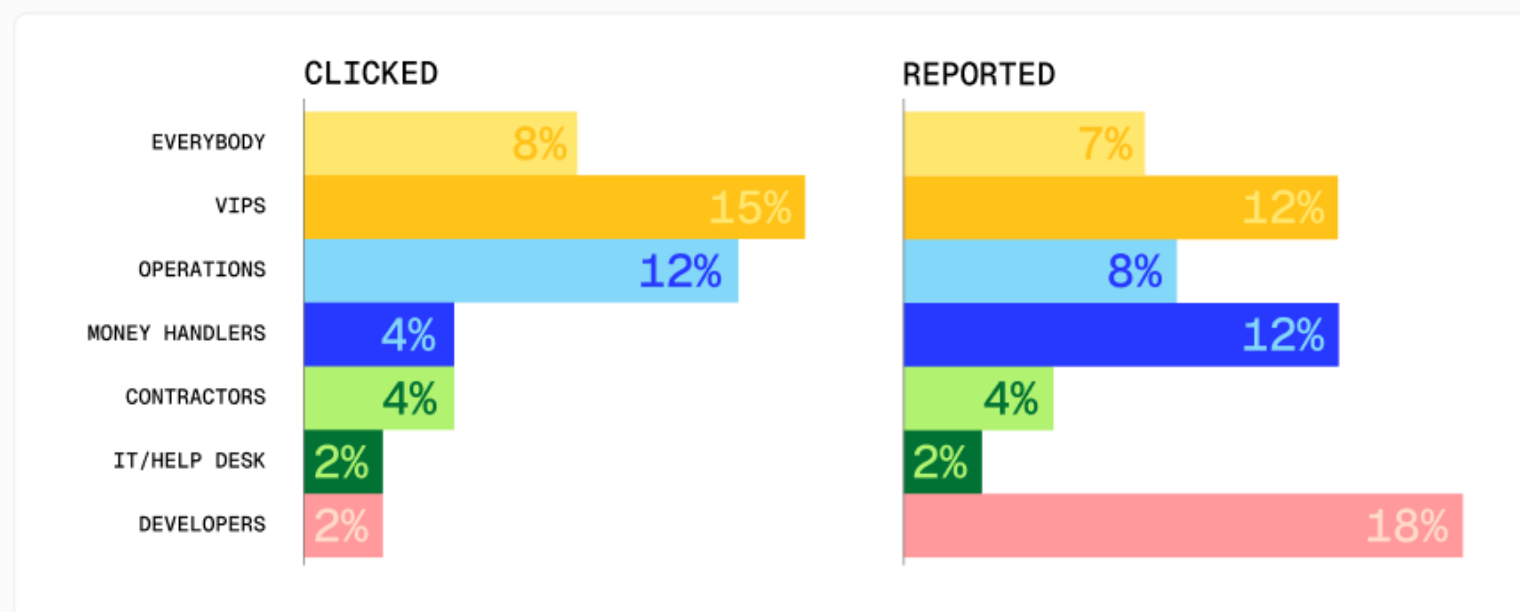
Behavior change is an important indicator of human risk reduction, but a critical companion metric is time-to-behavior change (TTBC). In other words, it measures how long it takes to reach a meaningful threshold of behavior change—say, when 75% of a cohort has taken a desired action. Similar to the well-known mean time to remediation (MTTR) in security operations, TTBC can be expressed either as an absolute duration (e.g., 8 days) or, when benchmarked against a control group, as a relative percentage (e.g., 20% of the control group's time).

It's an important metric because it demonstrates how quickly we can close the exposure window on human risk. For example, in the prior campaign, the security team asked employees to update their operating system software so their devices would have less of a chance of being hacked because unpatched vulnerabilities were left unchecked. They launched their campaign with a one-minute briefing video and followed up with Slack nudges to those who hadn't yet taken action. By the end of the week two, they had reached three-quarters of the cohort, and by the end of week five, participation had leveled off at 99%.



# Cohort comparisons

How do the groups within your company stack up on human risk campaign performance? In Fable, we automatically group employees by function, department, access, geography, tenure, affinity groups like “VIP,” and behaviors, such as “weak MFA.” Slicing and dicing employee data by cohort is useful for understanding where you’re having the most impact—and where you need to take further action. In the example below, we evaluate the performance of a recent phishing campaign across functional cohorts. Notice that, in this particular example, VIPs clicked on phishing at least double the rate of the others.



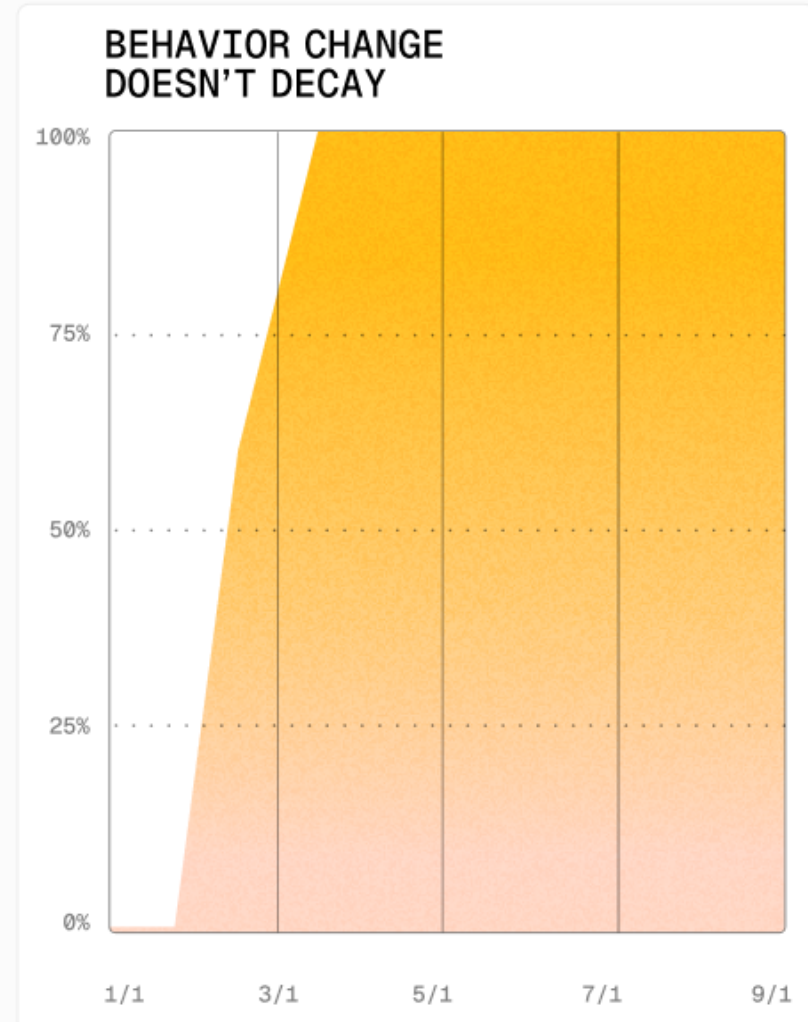
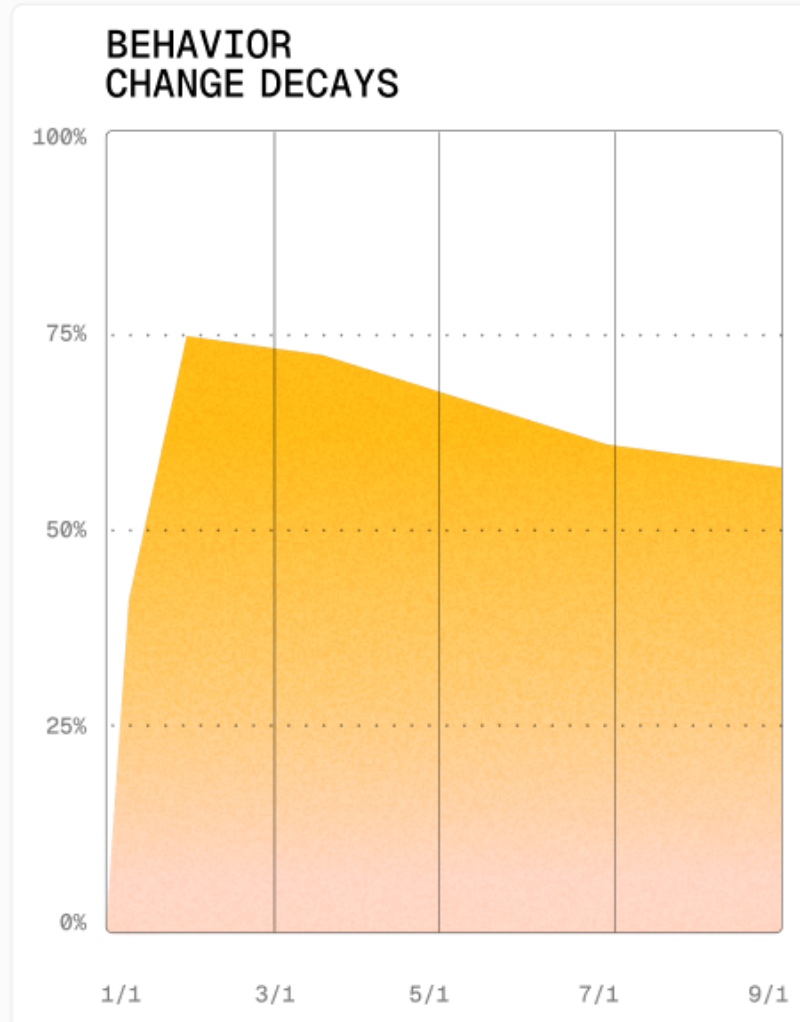
**Click rate:** the percentage of recipients who clicked the phishing simulation link

**Report rate:** the percentage of recipients who reported the phishing simulation through the approved reporting channel



# Behavior decay interval

How long does behavior change stick? The behavior decay interval measures the staying power of a campaign—how quickly people revert to old habits. In the example on the left, behavior tapers off; on the right, a targeted campaign led developers who inadvertently logged PII to an observability tool to correct course. Clear guidance on the error and remediation steps drove 60% compliance in the first two months and 100% thereafter. Because it's not always obvious what makes a campaign take hold, you need to monitor behavior and intervene whenever performance falls below the acceptable threshold.

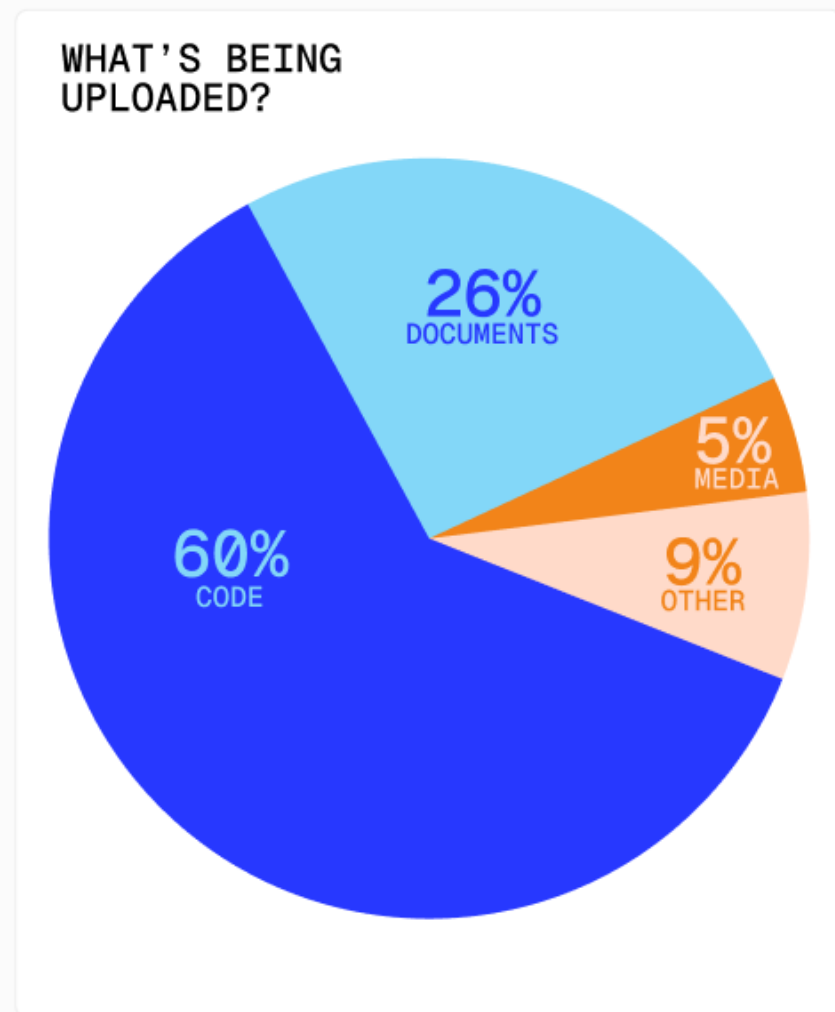
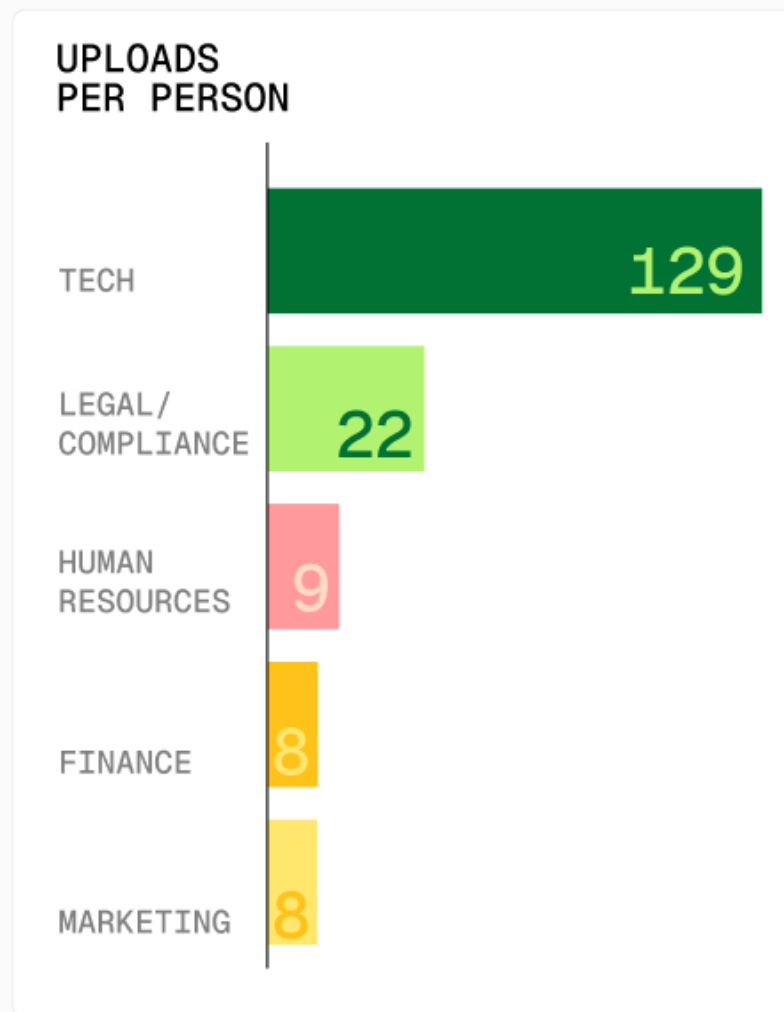


# AI swashbucklers

Who's sailing closest to the edge with generative AI?

In this section, we look at which cohorts in one organization uploaded the most content to generative AI tools over a six-month period, and what content they uploaded. As customers ingest more data from security tools, they'll be able to go beyond simple data uploads to ascertain uploads to sanctioned vs. unsanctioned AI applications, as well as data violations in the same.

**In one organization, these cohorts uploaded the most content per person.**





# Toxic combinations

Some risks are dangerous on their own, but become toxic when combined. The risk lift of toxic combinations measures how much more often two risks co-occur than you'd expect by chance. The formula we use is  $P(A \cap B) / P(A) \times P(B)$ , where the numerator is the proportion of people who exhibit both risks and the denominator is the expected overlap if the two risks were independent. A lift greater than 1.0 means the pairing amplifies overall exposure.

In this analysis, money handlers who failed phishing simulations registered a lift of 1.98—98% higher than random chance would suggest. Employees with sensitive data access and no MFA showed a lift of 1.17—17% higher. And IT admins who reused passwords had a lift of 1.13—13% higher. These examples reveal where overlapping weaknesses compound into disproportionate risk—and where targeted remediation delivers the biggest payoff.

Here are three toxic combination examples from one of our customers.

IF  $P(A \cap B) / P(A) \times P(B) > 1$ , TOXIC COMBINATION

PHISHING FAILURES 0.82%	$\cap$ 0.37%	MONEY HANDLERS 22.80%	RISK LIFT=1.98
NO MFA 30.02%	$\cap$ 8.46%	SENSITIVE DATA ACCESS 24.02%	RISK LIFT=1.17
PASSWORD REUSE 8.35%	$\cap$ 2.24%	IT ADMINS 23.65%	RISK LIFT=1.13

# Summary

This report begins Fable's ongoing exploration of human risk and behavior change. It introduces a core set of behavioral metrics that show how everyday employee actions shape an organization's security posture—and how targeted interventions can meaningfully reduce exposure. Over time, these measures will form the foundation of industry-wide benchmarks for human risk reduction—and a clearer path to sustained security behavior change.

Here are three takeaways from this data review

**01**

## Measure what matters: risk

Move beyond vanity metrics and measure real security behaviors—what people actually do, how quickly they improve, and whether those improvements last.

**02**

## Target with precision

Campaigns that target cohorts based on role, access, or risky behavior drive better engagement and faster risk reduction. Use cohort insights to tailor interventions where they'll have the highest impact.

**03**

## Fix the highest-leverage risks first

Identify where risks overlap and amplify each other. Remediating toxic combinations delivers disproportionate returns and can shrink your exposure window dramatically.

Get modern human risk management from Fable.



[FABLESECURITY.COM](https://fablesecurity.com)