LEARNING MADE EASY

Netskope
Special Edition

Securing Generative AI

for dummies®
A Wiley Brand

Enable responsible use of generative AI

Discover the risks and debunk the myths of generative AI

Create a secure strategy for your data

Brought to you by

netskope

Carmine Clementelli
Krishna Narayanaswamy

About Netskope

Netskope, a global SASE leader, helps organizations apply Zero Trust principles and AI/ML innovations to protect data and defend against cyber threats. The Netskope One platform provides optimized access and real-time security for people, devices, and data anywhere they go. Learn how Netskope helps customers be ready for anything on their SASE journey, visit netskope.com.

We would like to thank a number of individuals who, along with the authors, made this book possible:

From Netskope: Chad Berndtson, Catie Halliday, Atul Malik, Elena Matchey, Naveen Palavalli, Stephenie Pang, Bruno Raimondo, Neil Thacker

From Evolved Media: David Penick, Karen Queen, Evan Sirof, Lauren Wagner, Dan Woods

Securing Generative AI

Netskope Special Edition

by Carmine Clementelli and Krishna Narayanaswamy

for dummies®
A Wiley Brand

Securing Generative AI For Dummies®, Netskope Special Edition

Published by
John Wiley & Sons, Inc.
111 River St.
Hoboken, NJ 07030-5774
www.wiley.com

ISBN 978-1-394-26422-3 (pbk); ISBN 978-1-394-26423-0 (ebk); 978-1-394-28354-5 (epub)

Publisher's Acknowledgments

Some of the people who helped bring this book to market include the following:

Editor: Elizabeth Kuball
Acquisitions Editor: Traci Martin
Editorial Manager: Rev Mengle
Client Account Manager: Jeremith Coward
Production Editor: Saikarthick Kumarasamy
Special Help: Nicole Sholly

Introduction

Generative artificial intelligence (GenAI) exploded into the mainstream in late 2022. Since then, it has begun to transform (or promises to transform) many aspects of how businesses operate. Tools like ChatGPT and Google Gemini are more than just novel technologies — they've instigated a seismic shift in how data is managed and secured. It's estimated that at least one in four corporate employees interacts with a GenAI tool daily, mostly unseen and undetected by employers and security personnel. That's a concern because GenAI has a voracious appetite for your data, consuming the most mundane and the most sensitive data with equally aggressive greed.

GenAI has been around for years, to say nothing of the much bigger category of AI itself. But because of its mushrooming popularity, we now face a new, complex, and unexpected set of data security challenges. As these GenAI tools offer to enhance efficiency and drive innovation, they're also blurring the traditional lines of data protection and sending sensitive information beyond the safe confines of company networks into an expansive digital realm, where they may reemerge anywhere and in anyone's hands.

Effectively managing data security amidst the proliferation of GenAI is not just about risk mitigation; it's a strategic imperative. When approached correctly, effective data security enables businesses to leverage GenAI for competitive advantage. Conversely, neglecting the need for data security related to GenAI use can lead to catastrophic consequences for your business.

About This Book

This book equips your organization with the knowledge to balance the innovative potential of GenAI tools like ChatGPT and Google Gemini with robust data security practices. We emphasize the importance of understanding the unique risks posed by GenAI and advocate for advanced data loss prevention (DLP) strategies, regular risk assessments, and the integration of modern cybersecurity tools. We also underscore the necessity of cultivating a culture of responsible AI usage among employees. This book is a vital resource for navigating the complexities of data security in a landscape transformed by GenAI. It provides actionable insights

for organizations to protect their data while embracing important advancements in technology and business workflows.

Foolish Assumptions

This book assumes a few things about you:

- You have basic experience and familiarity with public GenAI applications like ChatGPT or Google Gemini.
- You're familiar with cloud-based apps and their significance in enhancing a business's productivity.
- You understand the need to secure interactions with such tools, especially given the remote and mobile workforce environment.
- You're driven to enable your organization to leverage the potential of GenAI while ensuring that sensitive data is never compromised.

Icons Used in This Book

We use icons to call attention to important information.

TIP

Anything marked with the Tip icon is a shortcut to simplify a specific task.

REMEMBER

The Remember icon flags facts that are especially important to remember.

WARNING

Heed anything marked with the Warning icon to save yourself some headaches.

Beyond the Book

This book is full of detailed information, but GenAI and security are evolving rapidly. If you find yourself at the end of this book wondering, "Where can I learn more?," just go to `www.netskope.com`.

IN THIS CHAPTER

- Tracking the history of GenAI
- Exploring the factors behind the rise of GenAI
- Discovering the many uses of GenAI
- Weighing the risks of GenAI against the rewards
- Looking at the security consequences of shadow IT
- Debunking the myths of GenAI

Chapter 1
Understanding GenAI

Technology can often feel complex and distant, but generative artificial intelligence (GenAI) stands out as both an accessible marvel and a tantalizing mystery. The story of GenAI, the latest stage in AI, is a testament to human ingenuity and innovation. But what's behind the curtain of this technological wonder, and why did it go so fast from a pursuit of scientists and hobbyists to the talk of the technology town?

From their emergence on the public stage in late 2022 to their current prominence, GenAI applications have had a remarkable impact. Just two months after it launched, ChatGPT (arguably the most well-known GenAI app) had more than 100 million users. By late 2023, ChatGPT had recorded 14.6 billion visits over ten months. Google Gemini (formerly known as Google Bard), another GenAI app, launched in late March 2023 and recorded 241.6 million visits in its first six months. Microsoft Copilot (yes, also a GenAI app) launched in February 2023 and today reports having 100 million daily active users. More GenAI tools are arriving every week, if not every day.

According to a joint survey by *Fortune* and Deloitte in the summer of 2023, more than half of CEOs surveyed are evaluating or

experimenting with GenAI, while 37 percent are already at limited production use or pervasive adoption. And GenAI usage overall by enterprise employees topped 10 percent as of December 2023.

From Humble Beginnings to Global Phenomenon

The roots of GenAI can be traced back to the early days of AI research. AI began around 1950 with Theseus, a robotic mouse that could remember its path and find its way out of a maze. By stages, AI advanced — first, with rule-based systems, then to context awareness, followed by domain-specific uses, then to reasoning machines, followed by self-aware/GenAI, which is where we are today. Initial endeavors were modest, focusing on creating algorithms that could mimic basic human tasks. But as technology advanced and computational power grew, so did the ambitions of AI researchers.

At the core of ChatGPT and similar tools are large language models (LLMs) powered by advanced neural networks. These digital brains, developed through deep learning techniques, have been trained on extensive datasets of digital text. This training began in earnest about five years ago and enables these chatbots to comprehend and generate human-like text. LLMs bring together machine learning (ML), deep learning, and natural language processing (NLP) technologies, allowing these AI tools to interact, learn, and evolve in ways that closely mimic human thought and language patterns.

The evolution of GenAI apps can be thought of as assembling building blocks, each stage laying the foundation for the next:

1. Model development.

   At the heart of any AI application lies its model. This foundational step involves designing algorithms capable of learning patterns and generating content.

2. Model training.

   The next step is training. Feeding vast amounts of data into the model helps the model learn, adapt, and refine its capabilities.

3. Build applications on the model.

   After it's trained, the model serves as the backbone for various applications, including chatbots and content generators.

4. Prompt interface.

   A well-designed interface ensures that users can seamlessly interact with the AI application, harnessing its capabilities to the fullest.

The Rise and Shine of GenAI

So, why is GenAI taking off now? The answer lies in a confluence of factors:

- **Technological leaps:** With more powerful ML models, and the computational might to back those models, we can now generate content that's eerily close to what a human might produce.
- **Broad horizons:** From churning out movie scripts to predicting financial trends, the applications of GenAI span a broad spectrum. That versatility is a hot topic across industries.
- **The economic factor:** Businesses see GenAI as a golden ticket, promising both cost savings and avenues for new revenue. When both happen, business processes change.

The Many Hats of GenAI

GenAI is a chameleon, adapting, evolving, and fitting into roles we never imagined.

As a writer's muse, it's more than just an editing tool; it's becoming a coauthor, aiding creative writing processes. In coding, imagine having an AI assistant that not only spots errors but suggests optimizations. In the media industry, it's a maverick, composing novel melodies and video narratives that reshape the way we create and consume content.

A Risk-versus-Reward Debate Rages

As a transformative technology, GenAI offers immense potential to reshape industries. In a world driven by efficiency, enterprises are at a pivotal point — drawn by the allure of efficiency and innovation that tools like ChatGPT provide.

But on the flip side, GenAI introduces significant challenges, particularly in data security and ethics. Because GenAI systems consume vast amounts of data and require users to provide data in order to be useful (in fact, data provided to the tool could be used for future enhancements of the underlying model), they can lead people to step around the usual habits and processes that protect sensitive information. As this technology advances, it brings with it concerns like deep fakes, misinformation, and potential data exposure — the information one user provides to the model can influence the answers given to other users, which could include competitors and threat actors. The loss of intellectual property (IP) is a big concern when using these tools.

WARNING

Businesses and individuals must approach GenAI with caution, ensuring its responsible use for the greater good of society. The *unchecked* adoption of GenAI poses significant risks, including potentially exposing sensitive information.

So, how are businesses striking a balance between risk and reward?

- **Risk assessment:** Before diving headfirst into the world of GenAI, enterprises must undertake comprehensive risk assessments. This involves understanding the potential pitfalls, from data leaks and ownership rights to the propagation of misinformation.
- **Custom implementations:** While GenAI apps that are publicly available offer ready-to-use solutions, some enterprises are exploring private implementations. By training these models on enterprise-specific datasets, businesses hope to exploit the power of GenAI while maintaining tighter control over data and outputs — lowering or eliminating the risk of valuable data or code leaking into the public domain.
- **Ethical guidelines:** The introduction of GenAI in workflows necessitates the establishment of ethical guidelines.

Businesses are drafting policies that dictate the responsible use of AI, ensuring that generated content aligns with enterprise ethics, brand values, and societal norms.

- **Continuous monitoring:** Because the world of AI is ever-evolving, businesses need to maintain constant monitoring, stay abreast of advancements, and recalibrate their strategies accordingly.

REMEMBER

Public models offer broad applicability, but a burgeoning market of AI solutions is tailored to specific enterprise needs. These private implementations, trained on select datasets, offer businesses a more controlled, customized AI experience, mitigating some of the risks associated with more generic models.

THE CHATGPT PHENOMENON

Given how quickly the GenAI landscape is broadening, it already feels dated to say this, but ChatGPT's rapid ascent — credited with breaking open the floodgates of mainstream interest in GenAI — can be attributed to several factors.

At the forefront is ChatGPT's user-friendly design. Its intuitive interface was crafted so that even those not well versed in technology could easily navigate and engage with the platform. Additionally, ChatGPT's versatility sets it apart. Whether answering intricate queries or assisting in content creation, the application repeatedly proved its mettle. Plus, the free version, which drives adoption, is fully functional and versatile. (Of course, *free* often means that you or your data are the product, and that's perhaps why free GenAI tools receive more scrutiny from businesses.)

But ChatGPT's rise wasn't just about the technology. The media's unwavering spotlight, combined with the buzz generated through word of mouth, amplified its reach, drawing users in unprecedented numbers. The app originated with OpenAI, an established AI research entity, so OpenAI's credibility bolstered users' trust, enticing them to explore ChatGPT's capabilities.

A Silent Surge and Its Security Implications

Much of the adoption of GenAI remains unseen, lurking in the shadows, away from the watchful eyes of organizational security teams. This phenomenon — known as *shadow adoption, shadow IT,* or increasingly, *shadow AI* — is reshaping security in ways previously unimagined.

GenAI's ability to provide tailored solutions often tempts departments and teams to integrate these tools without formal oversight. However, the very strength of GenAI — its inclination toward specificity — can also be its Achilles' heel. When data tailored to address a specific problem is fed into a general-purpose model, hosted and run by a third party, it risks exposure. This data leakage is exacerbated because many AI product terms don't offer robust protection for submitted data.

REMEMBER

After data enters the GenAI product's realm, control over the data diminishes, posing significant security challenges.

But data leakage is just the tip of the iceberg. Shadow adoption brings with it a host of other risks. Unauthorized access, murky ownership rights, potential misuse of AI-generated content, and the inadvertent amplification of biases are all pitfalls that organizations must navigate.

Recent studies have shed light on the magnitude of this trend. Almost all such studies underscore AI's value, but they also amplify data integrity and security challenges.

A case in point is ChatGPT's training methodology. Leveraging publicly available data on the internet increases ChatGPT's accuracy and knowledge base. But it also raises a question: When should a company's data be off-limits for such training? Some news organizations, cognizant of the risks, have shielded their data from the prying eyes of AI models.

Myth Busting: Getting to the Truth about GenAI and Data Security

When it comes to GenAI, the market is saturated with buzzwords, inflated promises, and tech jargon, leading many organizations to feel confused and overwhelmed by their options. But the truth is, not all perceptions about GenAI are accurate. This section helps you distinguish between fact and fiction, debunking some common myths surrounding GenAI.

Myth: Blocking GenAI apps is the only way to protect sensitive data

Reality: With effective data protection and proper management, GenAI apps can be used securely and to great benefit. It's about understanding and managing the risks, not avoiding the technology.

Myth: All GenAI applications pose the same level of risk

Reality: The risk profile varies. Publicly available tools may have different data handling practices compared to enterprise-focused solutions. It's essential to evaluate each application's merits and understand its data handling and privacy policies.

Myth: GenAI is 100 percent accurate and unbiased

Reality: AI models can inherit biases present in their training data. It's crucial to continually monitor, refine, and ensure fairness in AI outputs.

Myth: GenAI can only be as good as the data it's trained on

Reality: Quality training data is crucial, but the algorithms, model architecture, and continual learning processes play significant roles in the effectiveness of GenAI.

IN THIS CHAPTER

- Examining an often-overlooked risk factor: the human element
- Learning from a cautionary tale about data exposure
- Considering the potential risks of third- and fourth-party apps
- Leveraging the power of GenAI while keeping data safe
- Balancing innovation with privacy and safety

Chapter 2
The Cybersecurity Implications of GenAI

Chapter 1 explores the origins and evolution of generative artificial intelligence (GenAI), considering its potential to transform industries and redefine our technological interactions. As we venture further, our attention shifts to how traditional cybersecurity defenses now face challenges from threats borne of the integration of GenAI into our daily digital interactions.

The biggest security issue with GenAI is the risk of accidentally exposing sensitive data. When employees are focused on completing a project, they may give GenAI tools important, proprietary, and/or private information without considering the implications. After the AI has this data, exposure is possible. For example, the AI may use that data for enhancing the underlying model and then, in turn, share this confidential information when answering questions from other users. Additionally, someone with bad intentions could probe the GenAI with carefully crafted queries intended to elicit sensitive data.

Additionally, shadow adoption practices, where employees circumvent security protocols to utilize GenAI tools, exacerbate the

data protection problem. The manipulation of algorithmic biases, previously a topic primarily of concern for AI fairness, has now become a weapon for adversaries aiming to corrupt GenAI system outputs. But the biggest problems GenAI creates from a security point of view are data exposure and data loss.

The Human Element in GenAI Security

We must address the most direct and easily overlooked risk factor when dealing with GenAI: the human element. In this section, we consider some of the behaviors and decisions of employees that can significantly amplify the risks of data exposure and loss in the context of GenAI.

One of the most significant risks comes from well-intentioned employees who may inadvertently feed sensitive or private data into GenAI tools in their quest for efficiency and solutions. This could range from asking the GenAI tool to refine a confidential document or fine-tune an email containing personally identifiable information, to seeking help in debugging proprietary source code, to generating text from audio/video or the meeting minutes from a videoconference. The breezy ease of interaction with GenAI tools masks the potential danger of such actions, risking the exposure of critical information (as shown in Figure 2-1) and creating a repository of sensitive data within the AI system, which could be exposed.

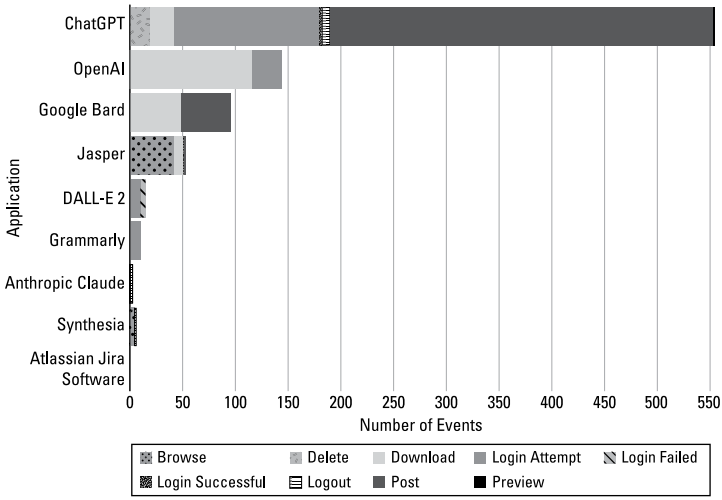

FIGURE 2-1: Users' activities in AI apps.

To mitigate these risks, organizations must:

- **Establish clear guidelines.** Develop and enforce clear policies regarding what types of data can be shared with GenAI tools. These guidelines should be easily accessible and regularly updated to reflect the evolving nature of these AI systems.
- **Educate and train.** Implement comprehensive training programs emphasizing the risks of sharing sensitive data with GenAI tools. Employees need to understand the potential consequences of their interactions with these systems.
- **Monitor.** Continuously monitor the organization's usage of GenAI tools. Implement systems that provide real-time reminders and alerts when employees are about to upload sensitive information. Set up guardrails that activate if sensitive data begins to be uploaded, preventing unintentional data exposure.
- **Audit.** Regularly conduct audits (quarterly or biannually) to review the organization's GenAI tool usage. Audits should assess compliance with data security policies, identify patterns of risky behavior, and highlight areas needing additional training or policy reinforcement.
- **Promote a culture of security.** Encourage a workplace environment where data security is a shared responsibility and employees feel empowered to question and report potentially risky behaviors or practices.

REMEMBER

A combination of education, policy enforcement, and a culture that prioritizes data security will help guide employees to use GenAI responsibly. However, technical controls that block sensitive types of data from being shared with unapproved GenAI services, for example, are a necessary part of data security. See Chapter 4 for more about how technology can help keep GenAI in check.

A Global Response to a Cautionary Tale about Data Exposure

The global corporate community was jolted into action following a pivotal security breach at Samsung Electronics in March 2023. Employees engaged in troubleshooting problems in internal

source code (which was sensitive in nature) uploaded that code to ChatGPT. The code's confidentiality is invaluable in the semiconductor industry. The code ended up stored on OpenAI's servers, where Samsung was unable to retrieve or delete it. If OpenAI used that code to enhance their underlying model, the code could later be provided in a response to other ChatGPT users, potentially leading to intellectual property (IP) theft and competitive disadvantages.

Reacting swiftly, Samsung fortified its household defenses, tightening GenAI protocols and bolstering security measures. The company embarked on developing an in-house GenAI tool with limited prompts for employees, alongside issuing stern warnings about the risks of leaking confidential information through GenAI platforms.

This wake-up call resonated globally, prompting a reevaluation and leading to new protective measures, not unlike parental guidelines in a family setting. These included

- **Setting stricter house rules:** Like concerned parents, many entities are establishing stricter house rules. For instance, the European Union's impending Artificial Intelligence Act aims for a safe, ethical growth environment for AI.
- **Time-outs and bans:** Some stakeholders have adopted stringent measures like time-outs or bans on specific AI activities prone to causing harm, mirroring parental decisions forcing children to sit still and/or restricting certain activities.
- **Family councils for unified guidelines:** The growing trend of forming family councils — public–private collaborations — seeks to create unified AI nurturing guidelines, ensuring secure growth and household safety.

Navigating GenAI's Extended Family Network

In the interconnected world of GenAI, the concept of family extends far beyond immediate relations. Think of immediate "family" as your users and the familiar corporate applications

that have been approved and vetted by your organization. Third- and fourth-party applications are like distant relatives and acquaintances — you don't know them well enough to blindly trust them with your data:

- **Third-party applications:** The cousins in the GenAI family — not part of the core household but closely connected. An example is if your customer relationship management (CRM) system gains new features by integrating with a third party's GenAI app.
- **Fourth-party applications:** Your cousins' cousins, entities even further removed that provide services to third-party apps. An example is if your CRM were connected to a third party's marketing automation tool, which itself was then connected to a GenAI tool.

Not all companies are transparent in disclosing third parties (your fourth parties), so the supply chain should be a key audit point with a technical control to verify (where possible).

Assessing and mitigating extended risks

Just as parents set clear expectations when their child spends time with distant relatives, technology leaders must assess risks and decide whether to trust third-party apps with data (if the risk is acceptable) or not trust them (if the risk of sensitive data exposure is too high).

WARNING

Be aware of the risks, behavior, and reputations of all applications. Scrutinize their security protocols, data handling practices, and compliance histories to ensure they're safe. Continuous monitoring is essential as an early warning system for any behavioral changes in these applications that may alter their risk profiles.

The hidden dangers of data whispers

In the cloud's vast expanse, data moves silently and swiftly between applications, like secrets shared in whispers. When GenAI is part of these exchanges, the security risks intensify. A data exposure incident in a third-party application and GenAI tool could ripple inward, reaching core corporate applications that are connected to them and leading to significant data exposure

or IP theft. Protect your company's data by understanding the terms, conditions, and handling of data by third-party apps.

WARNING

Navigating the GenAI landscape requires a deep understanding of this extended landscape. Leaders and users must recognize the subtle and complex ways risk can propagate and ensure robust security measures that safeguard AI and its entire ecosystem. Protecting sensitive data and IP requires vigilance and collaboration from every member of this extended family. The goal is to foster a secure environment where innovation thrives, but not at the expense of the organization's most valuable secrets.

The Dark Side of GenAI's Evolution

As GenAI develops greater capacity for innovation and problem-solving, this fast-evolving tool is also inadvertently opening up new avenues for risks. These cyber-threats leverage the very intelligence and capabilities that make GenAI so powerful. Here are some examples:

- **Automated data scraping and aggregation:** GenAI can be misused to automate the collection of vast amounts of sensitive data from various sources, aggregating it in ways that pose significant privacy risks.
- **Copyright violations:** GenAI tools are trained on extensive datasets of texts, images, and media, which can lead to unintentional outputs that resemble the copyrighted material on which it was trained. Users may then unknowingly use these seemingly unique results, potentially creating legal disputes and copyright infringements.
- **Data poisoning through GenAI:** One subtle yet potent threat is the potential for *data poisoning.* In this scenario, GenAI may produce compromised outputs when fed intentionally skewed, inaccurate, or malicious data. This risk is particularly concerning when AI is used for data analysis and decision-making.
- **Exploitation of predictive models:** The foresight afforded by GenAI to anticipate future outcomes from existing data patterns also harbors inherent risks. Similar to data poisoning, these predictive models become vulnerable to malicious

exploitation in the wrong hands, opening avenues for data distortion or theft.

- **Prompt injection attacks:** These are stealthy manipulations where attackers use clever inputs to make AI models reveal sensitive information or take unauthorized actions. The danger of these attacks is their disguise as normal interactions, demanding no traditional hacking, just a deep understanding of the AI's response patterns.

WARNING

The emergence of new cyber threats is an inevitable counterpart to GenAI's evolution. The aim is to harness GenAI's vast potential while fortifying the digital ecosystem against the emerging perils shadowing this technological advancement.

Setting Boundaries and Being Vigilant

You can strike a balance where innovation flourishes within a framework of safety and privacy. This involves a commitment to continuous learning, staying updated on AI trends, and engaging with the broader tech community to anticipate emerging threats. Regular health checks or security assessments help ensure GenAI remains secure and adheres to your standards.

TIP

The key is to establish clear, practical guidelines for the use of GenAI: setting operational boundaries; defining what's off-limits; closely monitoring GenAI's interactions with your data; and monitoring how users share data with AI. These practices foster a secure environment for AI to thrive without risking your sensitive information.

Delineating these boundaries reveals that certain risks are acceptable in pursuing innovation, while other risks pose too significant a threat to your digital security. But even the acceptable risks, such as the following, require oversight:

- **Guided explorations in innovation:** Leveraging GenAI for research and development holds promises of unprecedented innovation. Tread carefully, making sure your digital progeny doesn't unwittingly reveal proprietary methodologies or sensitive information, much like guiding a child on what's safe to share outside the home.

- **Controlled personalization:** You may want to enhance customer engagement by using GenAI to analyze data and preferences, enabling them to offer tailored, individualized experiences. The people, applications, and processes interacting with GenAI must be equipped to handle sensitive data responsibly, ensuring that identities and other sensitive information are not inadvertently exposed.
- **Collaborative decision-making:** In processes involving sensitive data, you may employ GenAI as a decision aid rather than the sole decision-maker. Here, the AI can propose options based on its vast reservoir of data. But humans make the final call, ensuring a safety net against unintended data exposure.

Here are some examples of unacceptable risks:

- **Unfettered access to the family safe:** Without strict access controls, there's a risk of accidental exposure of everything from personal data to the secret recipes of a company's operations.
- **Sole decision authority:** Entrusting GenAI with sole decision-making power, especially in areas dealing with sensitive data, is a gamble. Without the human touch, the AI lacks context, potentially leading to decisions that could result in data breaches or compliance violations.
- **Learning without supervision:** Allowing GenAI unchecked access to learn from sensitive data is like allowing a child unsupervised internet access. The AI may stumble upon and assimilate information it shouldn't, leading to unpredictable and potentially hazardous behavior.
- **Ignoring data regulations and compliance:** Overlooking international data protection laws (such as the European Union Artificial Intelligence Act) when GenAI is at play can lead to severe penalties and loss of trust. It's like neglecting to teach a child the importance of respecting the boundaries and rules of other people and environments.

IN THIS CHAPTER

- Seeing why blocking GenAI isn't always the right move
- Exploring the culpability of GenAI providers regarding security
- Combatting human error (and nefarious actions) with Zero Trust

Chapter 3
Security Precautions and Protecting Data

With generative artificial intelligence (GenAI), we're treading on both familiar and uncharted grounds. Any cloud-based Software as a Service (SaaS) app has established security protocols. For example, you control access and monitor usage and data exposure. Tools like cloud access security brokers (CASBs) remain instrumental, offering insights into app use and highlighting potential data vulnerabilities. But GenAI introduces unique challenges. The volume of data and the spontaneous nature of interactions mean even the most SaaS-savvy employees may unintentionally overshare. You can limit access, but you need robust data protection measures tailored for this landscape.

How do you safely interact with GenAI tools? What can you do to ensure your data remains secure? This chapter outlines the precautions and steps you can take.

Is Blocking GenAI the Answer?

Simply blocking GenAI tools may seem like an easy fix to potential data exposure risks. But the reality is that even if you try to block these tools, tech-savvy employees seeking to work more

efficiently will find a workaround. Such unregulated access, so-called *shadow IT* (see Chapter 1), ramps up risks instead of dialing them down.

Even if you completely block GenAI tools, will that solve the problem? After all, it's not just about slamming the door on threats. GenAI also opens the door to valuable innovation and productivity by turbocharging workflows, sparking creative ideas, and taking the grunt work off people's plates.

The smarter move is education, combined with robust advanced data protection capabilities. Teach your team to use GenAI responsibly so they know what data to keep out of AI's reach. Modern cybersecurity tools can help by sniffing out sensitive data, hitting the brakes on unauthorized use, and flashing alerts and coaching in real time.


TIP

New GenAI tools will keep popping up. Organizations must take a proactive stance. Embracing GenAI with a strategy of smart control and adaptation, instead of a flat-out no, sets you up to securely integrate future benefits.

With that in mind, let's consider some familiar security guidelines and their relevance when it comes to GenAI.

Strict access control

Robust access control is the gatekeeper for corporate resources, especially for GenAI. But currently, many GenAI tools are like the Wild West — unsanctioned third-party apps often accessed by employees through shadow IT. Even though we may lack full control today, you need to monitor who's using these tools, how often, and for how long, laying the groundwork for the stricter controls that will inevitably come.

Not all GenAI tools present the same level of risk. Companies must ensure that risky apps stay out of reach or that only authorized employees can access them. As some of these tools become more integrated into corporate ecosystems, limiting access to sanctioned tools may become critical.

Data anonymization

Let's consider a couple of nightmares:

- Sarah, an employee at a tech firm, is drafting an email to a potential client. She wants to ensure her pitch is perfect, so she runs it through a GenAI tool for suggestions. She pastes the entire email, complete with names, addresses, and other specifics. The AI may enhance her pitch, but she's unwittingly risked exposing sensitive data.
- In the same firm, the research and development (R&D) team is working on a groundbreaking piece of software. Eager to troubleshoot a tricky section of their code quickly, they upload the code to a GenAI tool. That code contains the name and details of a new unreleased R&D project, internal Internet Protocol (IP) addresses, private keys, and other highly confidential info. In their haste, the team may have just shared their hard work with the world.

These scenarios underscore the critical importance of *data anonymization*, which is like giving your data a mask, allowing it to interact without revealing its true identity. By replacing names with pseudonyms, or replacing or deleting certain confidential pieces of information (a project name, internal IP addresses, private keys, and so on) from code, you help to ensure that even if the data gets into the wrong hands, it tells no tales.

Data minimization

GenAI entices users to share more in the belief that more data leads to better results. *Data minimization* is the practice of using only the essential data required for the AI model to function effectively. It's not just about stripping out personal identifiers but also about being judicious in what we feed these models. Data minimization aligns closely with data anonymization. Although data anonymization masks identifiable information, data minimization reduces the volume of data shared. It's a proactive approach, ensuring that even if there's a breach, the exposed data is limited.

Imagine you're crafting a press release and you need help with a particular paragraph. Instead of uploading the entire document,

which may contain sensitive details, focus on that specific section. Similarly, when an R&D team working on groundbreaking software wants to troubleshoot a tricky section of the code, instead of uploading the code in its entirety to a GenAI tool, they can upload only unassociated fragments of it. By minimizing the data that's shared, you're protecting potentially confidential information and ensuring that the AI concentrates on the task at hand.

REMEMBER

With GenAI tools, it's crucial to pause and ask, "Is all this data necessary?" More often than not, you'll find that a minimalistic approach not only suffices but is also safer. Regarding data security, the mantra should be, "The less, the better."

Local deployment

An ideal scenario would be to follow long-standing security principles and keep GenAI securely within your walls. Running GenAI tools directly on your company's machines would minimize risks by helping ensure that your data stays within the company's network. However, nearly all GenAI tools, especially the leading ones, are cloud-based. Thus, the established security principle to deploy software locally doesn't fit in today's GenAI landscape.

The encryption paradox

Encryption is like a digital shield guarding your most confidential data. When data is encrypted, it's scrambled, becoming readable only with the correct decryption key. Similar to anonymization, this tactic ensures that even if someone gets their hands on the encrypted data, they can't understand it.

But to work, GenAI tools need to read the data. Handing an encrypted file to a chatbot is like giving someone a book in a language they don't understand. The tool can't process it. So, although encryption is a gold standard in digital security, its applicability to interacting with GenAI is limited.

Audit trails

Audit trails serve as an essential record of interactions, meticulously logging every activity related to data handling and GenAI interactions. These logs act as a reference for investigations, making it possible to pinpoint the origin of an issue. This forensic capability is invaluable in a world where data breaches and leaks can have monumental consequences.

Audit trails also offer insights into user behavior. By analyzing these logs, you can gauge how your employees interact with GenAI tools. Such analytics can reveal surprising trends, patterns, behaviors, and potential vulnerabilities. This understanding can help companies strike a balance between security and accessibility.

Local defenses

Regular updates and patches play a crucial role in fortifying defenses against known vulnerabilities. Cloud-based solutions are actively updated, whereas local software deployments require a more hands-on approach. If a team managing a local deployment skips an update, security could be compromised. As we look to the future, this practice will grow in importance if local deployments of GenAI tools become prevalent.

Third-party audits and certifications

When you're evaluating a GenAI service, lean toward providers that have undergone stringent third-party audits and earned certifications such as ISO 27001, SOC 2, and General Data Protection Regulation (GDPR) compliance. These certifications indicate a commitment to high standards of security and data protection.

Unfortunately, many of today's GenAI tools provide limited visibility into their assurance mechanisms. The absence of widespread third-party audits and certifications in the GenAI sector underscores the need for users to be discerning and proactive.

Recent industry shifts, such as major tech companies offering insurance against AI-related mishaps, hint at a growing awareness of the importance of accountability. These steps are commendable, but they're just the beginning. For GenAI to truly earn the trust of its users, a more robust framework of third-party validations is imperative. Until then, users must navigate cautiously, prioritizing providers that demonstrate a clear commitment to transparency and security.

Data use policies

Every organization, regardless of size, needs a comprehensive data use policy outlining how employees manage and share sensitive information. Informing every team member about these guidelines is essential in order to minimize data disasters, but their real effectiveness lies in enforcement. Because of the inevitability of

human error, integrating advanced technological tools (like those discussed in Chapter 4) is necessary to ensure consistent adherence to these policies.

A clear data use policy sets expectations and reflects an organization's dedication to data security. Ultimately, a data use policy represents an organization's commitment to protect its data assets and use GenAI tools responsibly.

Data backups

GenAI tools designed for content creation handle a lot of data. When employees use the generated output in their work, there's a risk that an original dataset will be replaced by an inaccurate one generated by AI, propagating mistaken or even malicious information.


TIP

That's why data backup is crucial. Whether data is mistakenly changed by AI, lost in a data breach, or unintentionally spread, backups give you a way to recover.

Constant reviews

The fast pace of GenAI's evolution makes it imperative for organizations to remain informed and understand the ever-changing landscape. As GenAI tools advance, the nature of their interactions keeps changing. This is especially true when integrations occur behind the scenes through application programming interfaces (APIs) with familiar corporate platforms like Salesforce, Microsoft 365, and others.

This dynamic nature underscores the importance of a vigilant, proactive approach to constantly reviewing and updating usage policies and terms of service, as well as keeping an updated inventory of approved GenAI apps (which is supported by the National Institute of Standards and Technology [NIST] AI Risk Management Framework [RMF]). Without this inventory, very few controls are effective. You must also understand what type of data is flowing to what app or service, so you can put controls in place to protect sensitive data. Finally, constant reviews must be supported with advanced analytics to automatically produce a visual map of data flows.

What Is the Responsibility of GenAI Providers?

Unlike traditional SaaS applications, which can function without deeply analyzing user data, GenAI thrives on data. Data is the lifeblood that powers its capabilities. This distinction raises a critical question: What should GenAI providers offer to ensure the security of their applications?

In an ideal scenario, AI providers would offer a suite of security measures to ensure the safety of user data, including the following:

- **Transparent data protection policies:** Providers clearly outline how they handle, store, and use data. This transparency gives users confidence in the platform's commitment to data security.
- **Data retention limits:** After using the data, the system deletes it promptly. This minimizes the risk of data exposure over time.
- **Data sovereignty guarantees:** Ensure that data is not exposed to other countries or external entities. The data provided is used solely for the intended interaction, not for any other purpose.
- **Anonymization and minimization:** Beyond what users do, the AI system further anonymizes and minimizes data during processing.
- **User data deletion:** Users can delete their data upon request.
- **Model security:** The model uses security controls to safeguard against adversarial attack, data breach, and unauthorized access.
- **Infrastructure security:** The underlying cloud infrastructure uses a security system.
- **Incident response:** The provider responds to incidents quickly and effectively.

- **Ethics:** The provider's moral principles and techniques ultimately benefit society.
- **Environmental impact:** The provider's physical infrastructure has a minimal negative impact on the environment.

Asking these models to forget, however, is like asking someone to unlearn everything they've experienced. It's not gonna happen. So, although collective intelligence is what sets GenAI apart from other technologies, it also means that these models are unlikely ever to provide the full breadth of data security measures we want.

WARNING

We can outline the ideal security measures, but you have to understand that the inherent nature of GenAI may make some of these measures unattainable. Always approach GenAI with a clear understanding of its capabilities and limitations.

REMEMBER

Ultimately, the burden falls on both AI providers and users. Providers should strive to offer as many security measures as possible. But providers are not following the ideal scenarios we're laying out here. So, whether providers do their job or not, you must remain informed and cautious, understanding that GenAI operates differently from other tools you may use.

Human Error Makes Zero Trust an Imperative

A staggering 82 percent of data breaches occur due to human mistakes. Everyone from entry-level employees to top-tier executives is susceptible to making errors when interacting with GenAI. The pressure of modern work environments, coupled with the immediacy of digital communication, often leads to shortcuts.

You may expect your teams to anonymize data, for example, but in their rush to get answers, they may forget. In their haste, they may also inadvertently expose sensitive data through emails or other communication channels. The quest for the easy button often trumps security precautions. This is what makes the security model known as Zero Trust so important. Zero Trust addresses many of the challenges we've described by imposing the rule "Never trust, always verify."

Zero Trust recognizes that the context of interactions is ever-changing (see Figure 3-1), making it necessary for verification to be constant and ongoing. Zero Trust is not merely a technology but a set of principles to ensure that every interaction with sensitive resources, be it data or applications, is verified and controlled.

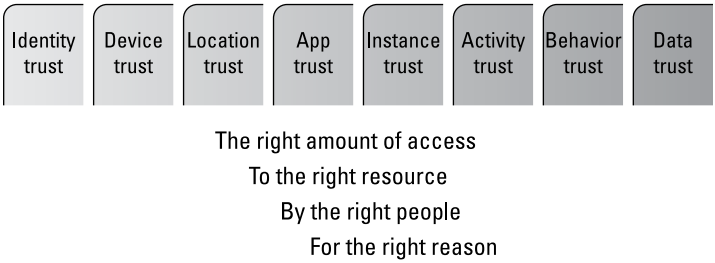

FIGURE 3-1: The Zero Trust mantra is "Never trust, always verify." Verification must be constant and ongoing.

Here are the key tenets of the Zero Trust model:

- **Identity verification:** Know who is interacting with the resources.
- **Action monitoring:** Understand the actions being performed.
- **Data sensitivity:** Control access to sensitive data and its use.
- **Destination tracking:** Identify which application or data is being accessed and by whom.
- **Behavioral analysis:** Monitor user behavior throughout the interaction.

The Zero Trust model becomes even more crucial in GenAI, which we discuss further in Chapter 4. Written policies and training sessions are a start, but more is needed. The confidence users place in chatbots and AI models, combined with the constant data interactions, amplifies the risk.

To consistently mitigate these risks, organizations must implement Zero Trust principles with technology that provides visibility and a clear understanding of the following:

- Application risks, granularly categorized
- What data is sensitive

- Where data flows occur
- User identities
- User interactions with AI apps
- Overall user behavior
- How the context of each of those things changes constantly
- Which instance (free, personal, corporate) of a GenAI service is being used
- Third-party apps

Behavioral analytics tools, in particular, can shed light on negligent behaviors, which, in the case of GenAI, are more often unintentional than malicious.

REMEMBER

As GenAI becomes an integral part of daily operations, the challenges it presents will evolve, making the call for advanced cybersecurity measures, such as the Zero Trust model, louder than ever.

IN THIS CHAPTER

- Securing AI through the power of AI
- Exploring how to apply Zero Trust principles
- Increasing visibility
- Leveraging AI/ML to gain advanced insights into app risks and more
- Enabling seamless communication for a richer real-time risk assessment
- Taking a broader view of the app landscape to adapt your business

Chapter 4
How Modern Technology Helps Secure GenAI

As we navigate the security implications of businesses using generative artificial intelligence (GenAI), we're quickly learning that most security technologies won't keep pace. In this chapter, we delve into how modern technology, especially security platforms that themselves *use* AI, will be pivotal to protecting data.

Harnessing AI to Secure AI

When it comes to GenAI, the massive volume of data involved, the dynamic nature of interactions, and the diversity of applications mean that conventional security methods are no longer sufficient. Without the right tools, many risks associated with GenAI could remain concealed, emerging only when it's too late to prevent them. Luckily, AI *itself* is going to help you safeguard AI. But we also need to consider the broader environment in which AI will operate.

Historically, technology has empowered security teams to automate tasks that would otherwise be daunting, if not impossible, to handle manually. For instance, data protection tools have granted organizations invaluable visibility into their sensitive digital assets, illuminating potential risks and vulnerabilities. And, although traditional approaches to training employees about security frequently lag behind the rapidly evolving digital landscape, automated tools can provide security guardrails and real-time feedback that ensures users are always informed about how to adhere to best practices.

Addressing the many data protection challenges GenAI presents starts with a holistic approach to data protection. Security isn't just about having the right tools — it's about ensuring those tools are integrated to work in harmony, sharing data intelligence and building a rich context for analytics that can inform the entire security apparatus.

A Platform for Implementing Zero Trust Principles

The concept of Zero Trust — "Never trust, always verify" — is the cornerstone of modern, robust cybersecurity. In the context of GenAI, this doesn't mean restricting the use of these tools. Instead, it's about ensuring that every interaction with them is monitored and considered, especially when it comes to sensitive data. We must first continuously verify that sensitive information isn't inadvertently shared with GenAI platforms. It's a data-centric approach, emphasizing the importance of safeguarding sensitive data at all times, regardless of the tool or platform in use.

REMEMBER

You can and should leverage the capabilities of GenAI, but applying Zero Trust principles ensures that you do so without compromising the security and confidentiality of sensitive data.

Zero Trust principles are what underpin a properly architected Security Service Edge (SSE) platform. Traditional security portfolios consist of numerous discrete tools that fit together loosely and certainly don't make each other smarter. In contrast, SSE is a modern cloud-based approach that emphasizes consolidating security tools into a unified platform. It encompasses data loss

prevention (DLP), web and cloud application security, and private access tools, all working together. This platform facilitates data intelligence sharing, allowing for richer context, more robust analytics, and a comprehensive way to apply Zero Trust principles to every interaction with data and applications. The goal is to create an environment where security tools are not isolated but are part of a cohesive system, exchanging data and insights to offer a more comprehensive security solution.

But what does all that mean in practical terms? Let's look at some of the core capabilities of SSE that become benefits when the architecture is correctly designed:

- **App usage visibility and access control:** Visibility into app usage is crucial for monitoring how GenAI tools are accessed and used within an organization. By leveraging machine learning (ML) and large language models (LLMs) for app usage discovery and app risk categorization, SSE can identify potentially risky interactions with GenAI applications, ensuring that only safe and compliant usage is allowed.
- **Automatic detection and protection of sensitive data:** SSE's automatic data protection capabilities, integrated into the platform, are vital for detecting and preventing the unauthorized sharing of sensitive data with GenAI tools. This includes curbing risky uploads, posts, or data sharing that could lead to data exposure or compliance violations.
- **Promoting user responsibility and awareness via coaching:** Given the ease with which employees can interact with GenAI tools, educating them on responsible use is essential. Real-time alerts and customizable coaching workflows inform users about the risks of sharing sensitive data with GenAI tools. This proactive approach reduces the likelihood of accidental data exposure and reinforces a culture of security awareness.
- **Data protection beyond uploads:** As GenAI tools often integrate with other cloud applications, protecting data across these interactions is crucial. The ability to monitor and secure data against unauthorized cross-application access in the cloud directly addresses the complex data flow challenges posed by GenAI integrations. This includes assessing and mitigating risks in cloud-to-cloud data transfers.

- **Addressing legal risks from AI-generated content:** GenAI tools can produce content that may infringe on copyrights or other legal rights. SSE's real-time analysis of data from these AI applications ensures that some legal risks, such as copyright infringement or noncompliance with data protection laws, are promptly identified and addressed.

As organizations continue to embrace GenAI and expand their use of SaaS applications, the importance of a robust security framework can't be overstated. SSE offers a comprehensive solution, ensuring that as technology evolves, security remains a top priority, paving the way for sustainable growth.

The Importance of Visibility: You Can't Protect What You Can't See

Cloud access security brokers (CASBs), which are a critical part of any SSE platform, illuminate hidden aspects of the digital infrastructure, including the potential vulnerabilities of new GenAI tools. A CASB provides real-time visibility into SaaS applications, ensuring that organizations are always a step ahead of the rapid proliferation and potential risks of these apps. With that confidence of security, organizations can confidently harness the power of GenAI.

However, not all CASBs are created equal. Many commercially available CASB products offer basic insights into a limited number of more popular apps, but Netskope provides a panoramic view of all modern SaaS applications. Netskope's vast and dynamic coverage is crucial given the rapid proliferation of GenAI apps.

But you need to do more than just identify new AI apps. Organizations must understand the associated risks, user behaviors, data protection concerns, and compliance implications. This demands advanced app insights and constant vigilance. Deep risk categorization and continuous monitoring of application access ensure that organizations can detect and address security concerns in real time, fortifying their defenses against potential breaches.

REMEMBER

Taken together, visibility and continuous monitoring of the application landscape help organizations make informed decisions, tailoring their security protocols to address specific threats.

Going Faster with AI/ML

GenAI applications are emerging at an unprecedented rate, making manual visibility and monitoring of the landscape increasingly challenging. The sheer volume and speed of this growth demand automated solutions.

WARNING

To safeguard your organization in this rapidly evolving landscape, harnessing the power of AI/ML for comprehensive app discovery, data security, and behavior analysis is critical.

Netskope leverages AI/ML both for rapid app discovery and categorization (see Figure 4-1) and to gain advanced insights into app risks, security postures, and behavior. This comprehensive approach ensures consistent monitoring of these apps as they evolve, ensuring that organizations remain in sync with the dynamic app landscape. Rapid app discovery is crucial, but it only scratches the surface of what's needed for risk management. As businesses grapple with the decision of whether to allow access to GenAI, they must recognize that not all GenAI apps are created equal. Some are inherently riskier than others. The challenge lies in discerning which ones to embrace, which ones to tolerate and monitor, and which to restrict.

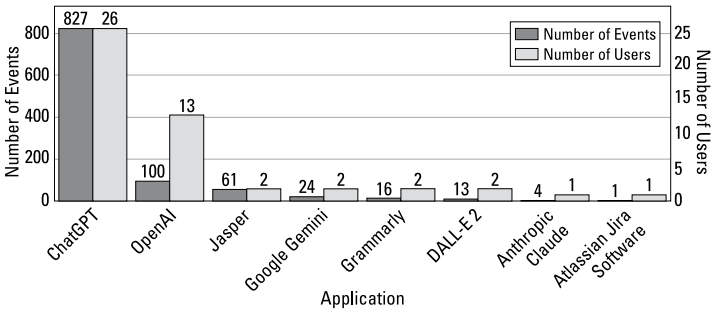

FIGURE 4-1: Example of the top ten AI apps in use in an organization by events and user count.

With its unique Cloud Confidence Index (CCI), Netskope enables organizations to dive deep into each app's potential risks, considering everything from compliance issues to data protection concerns. With the ability to identify more than 50 distinct risk

categories, CCI gives organizations a deep and comprehensive view of their app risks. This detailed insight ensures that companies can make well-informed decisions, balancing the innovation of GenAI tools with the security they need.

Defining access control policies in real-time

The modern workforce operates in a hybrid environment, seamlessly transitioning between office spaces and remote locations such as homes and coffee shops. This shift necessitates a real-time, cloud-delivered access control service. Such a service transcends the limitations of conventional firewalls and security gateways. Instead, a cloud-based SaaS security architecture becomes indispensable, offering the agility to monitor user access to apps and data from any location.

However, sifting through each app and evaluating potential risks and benefits is an arduous task for IT teams. This is where tools such as Netskope CASB excel, granularly categorizing apps and their risk profiles and enabling organizations to set access control policies to monitor, allow, or deny access to GenAI apps as needed. For instance, accessing a verified GenAI app to help draft a promotional email may be safer than using a riskier, unknown, unverified GenAI app with higher threat implications.

Policies, informed by detailed insights from the CASB and its CCI, can automatically determine which apps align with the company's predefined risk thresholds.

Real-time digital coaching

Employees, the primary users of these apps, must be well-informed about the potential pitfalls and best practices. Real-time coaching mechanisms can bridge knowledge gaps. Imagine that an employee trying to access a new app receives an immediate alert about its associated risks. Such cues serve as gentle reminders to adopt precautions, such as data anonymization. These cues also foster a culture of caution and empower employees to make informed decisions, ensuring that the organization's digital activities remain innovative and secure.

Netskope provides a unique mechanism to deliver real-time alerts and automated coaching workflows when users access GenAI and other SaaS applications. This includes customizable warning pop-ups (see Figure 4-2) that offer guidance about the responsible use of these tools, which involves the user in the access decisions after acknowledging the risk.

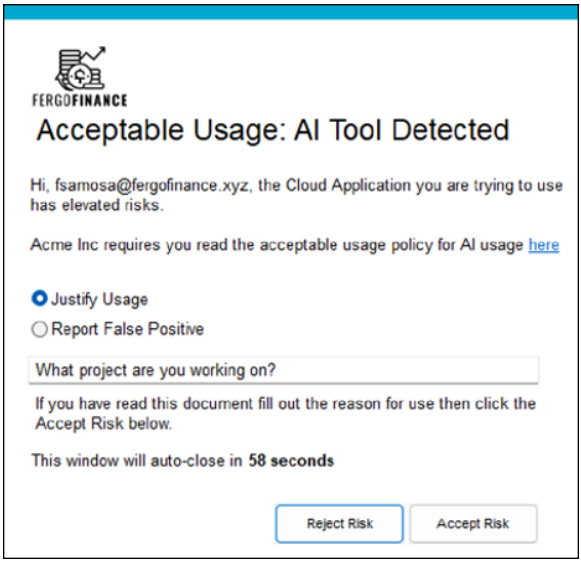

FIGURE 4-2: Real-time alerts can educate users on the potential risks involved with accessing and using various apps.

Understanding data risk through analytics

To understand the usage patterns and potential risks of GenAI applications, you need to look beyond identifying apps in use and their generic risks. You also need to look at *how* these apps are being used within a specific organization over time. This is the power you get from analytics. CASBs offer a broad perspective on app risks based on global data. Augmented with analytics, like Netskope Advanced Analytics, they can provide an even more tailored view, discerning specific trends, user behaviors, and data flow in any given organization (see Figure 4-3).

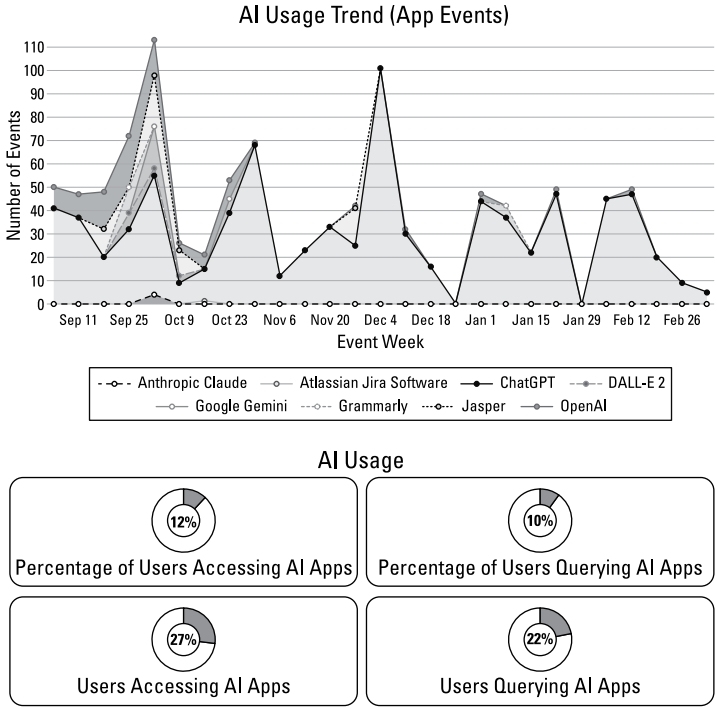

FIGURE 4-3: The analytics dashboard provides insights into usage and risks.

The granularity offered via analytics allows organizations to identify specific risk behaviors, recognize potential data bottlenecks that are meaningful to them, and detect patterns that may identify both security vulnerabilities and manageable risks. For instance, analytics may show that a particular GenAI app is being widely used in meaningful ways across the organization. This revelation, in turn, should prompt a more considered approach to risk management. Moreover, analytics can spotlight emerging trends, such as the rising popularity of certain apps. Proactive insight allows for timely interventions, be it providing additional training or implementing safeguards.

Analytics help us to understand what's happening with data after access is allowed, the volume of data being transferred, the frequency of uploads and downloads, and even the times of day when activity is most intense, pinpointing specific users and

their activities. If users aren't taking steps to minimize their data exposure, for example, organizations can deploy additional tools that automatically identify and proactively protect sensitive information.

Netskope CASB with Advanced Analytics forms a powerful (dare we say dynamic?) duo, offering insights tailored to each organization, ensuring that organizations can secure the use of GenAI applications with confidence and clarity.

Limiting sensitive data exposure

Imagine a world where every piece of data — from a generic file to a confidential document to source code or even a textual conversation — flows seamlessly across platforms. How do we ensure our most sensitive information doesn't slip through the cracks? This is the domain of modern, cloud-based DLP systems. Traditional DLP uses a limited set of data identifiers to recognize certain data types. But like the story of the boy who cried wolf, they sometimes raise the alarm for a harmless piece of data, imagining a threat where none exists. This not only disrupts people's work but also diminishes trust in the system.

With AI and ML, our sentries can be more intelligent. Netskope DLP, for example, uses AI and deep learning, advanced ML Classifiers, convolutional neural networks, natural language processing (NLP), and a pioneering "Train Your Own Classifiers" capability that helps reliably identify and protect sensitive data. These technologies dive deep into the data, sifting through layers, understanding context, and discerning between a harmless number and a sensitive piece of information (like a passport number or a bank account detail). And there's more to it than just text analysis. In a world where images speak louder than words, AI ensures that a photo of confidential information like a passport, credit card, source code, or even a whiteboard doesn't go unnoticed.

After we identify what data is sensitive, we then need to guard that sensitive data. If someone tries to send such data to a GenAI tool, the system should spring into action, blocking the data flow and immediately alerting the user (see Figure 4-4). It's a dance between protection and education, ensuring security while fostering a culture of awareness.

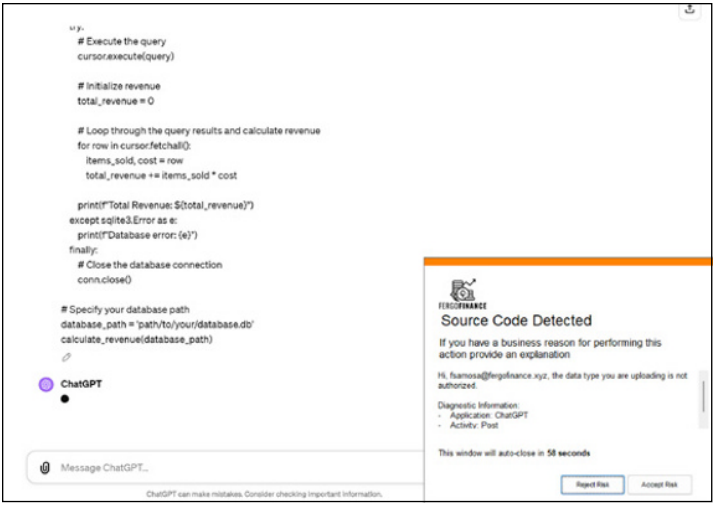

FIGURE 4-4: Preventing proprietary source code from being exposed through GenAI.

The difference between knowing and enforcing

If you think of data protection as a large theater production, the enforcement function is the job of the director, who ensures that every actor plays their part perfectly. But how do you ensure the script is followed to the letter, especially when the stage is as vast and dynamic as GenAI operations?

The real-time coaching messages we mention earlier are gentle reminders, cues that guide the actors. If a user, in their zeal to interact with a GenAI app, inadvertently tries to share sensitive information, these messages are a prompt that makes them pause and reconsider. It's like a whisper in their ear, saying, "Are you sure about this?" But sometimes a whisper isn't enough. There are moments when the director must stop a scene before things go off the rails. This is where a security platform with real-time blocking mechanisms comes into play, ensuring that no sensitive data, no matter how inconspicuous, makes its way where it shouldn't, such as a GenAI app.

Real-time coaching messages (see Figure 4-5) are fundamental to inform the user of the violation that has just occurred and of the action taken by the data protection solution. It's like telling the actor, "This time, I suggested your line. I hope next time you remember it."

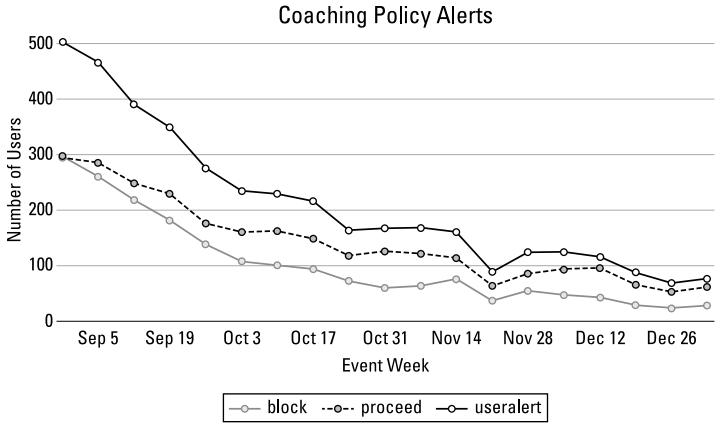

FIGURE 4-5: Reduction of risk over the time through user coaching alerts.

WARNING

An app previously deemed safe may suddenly become a potential risk due to its new associations. As GenAI evolves, enforcement strategies must evolve with it. New apps, plug-ins, and connections are continually being introduced. By monitoring these connections and understanding the changing risk profiles, you can ensure that your data remains secure, no matter where the tides of innovation take you.

Managing multiple instances of GenAI applications

As the use of GenAI applications becomes more prevalent in both our professional and personal lives, new challenges will emerge. We've seen this happen before. For example, an individual may have a corporate Gmail account for official communications and a personal one for private interactions. Although sending sensitive corporate data through the official account is permissible, doing so through a personal account poses significant risks.

The challenge lies in discerning which instance is being accessed and ensuring that sensitive data flows only through the appropriate channels. *Instance awareness* refers to the ability to distinguish between different occurrences of the same application. A common method employed by many security tool vendors is Uniform Resource Locator (URL) monitoring. Sometimes this is enough — for example, when the URL for a corporate instance differs from its personal counterpart. But this isn't always the case. Often, both instances share the same URL, making it nearly impossible to differentiate based solely on this parameter.

WARNING

Distinguishing between instances in real time isn't an easy job. Even knowing the login account may not be enough to know the true instance. Netskope provides a unique patented mechanism that identifies true instances of an application. This is based on an instance resolution mechanism that is independent of seeing the login transaction (not always visible on the wire in all transactions), but based on requesting the app itself to reveal its instance to Netskope for all sessions.

As tempting as it may be to block personal instances entirely, preventing corporate data from ever moving to personal instances may be more practical. Our focus should be on controlling the data flow. It's not about restricting access to personal instances but about ensuring that sensitive corporate data remains confined to its designated channels.

As GenAI tools gain traction and become integral to organizational operations, instance awareness has become even more crucial. It's not just about preventing shadow IT/unauthorized tools but about recognizing that even sanctioned tools can have multiple instances. The future will see a proliferation of these instances, making it imperative for organizations to have granular control, visibility, and access levels.

A New Approach to CASB

The risks associated with third- and fourth-party interactions can no longer be ignored. But what does this mean in practical terms and how can organizations navigate this intricate web? Most CASB systems can tell if someone is accessing an application

like ChatGPT or Salesforce. But that's where their insight ends. They remain oblivious to the activities *within* these apps and certainly can't react in real time, leaving a gaping hole in security oversight. In Salesforce, for example, each new plug-in from a third party introduces potential vulnerabilities. Without real-time awareness, organizations are left exposed.

Netskope has taken a unique approach by integrating all the CASB elements that have been traditionally disjointed, such as inline control and application programming interface (API)–based app introspection, into one cohesive, user-friendly solution within SSE. At its core, the CASB solution has two main components:

- **The inline component:** This component determines who gets access to what, distinguishing between corporate-sanctioned and personal apps and applying both data security and threat prevention to real-time traffic. But its capabilities don't end there. It can also detect personal instances of corporate apps, adding another layer of security.
- **The API component:** This component provides a deep dive into the activities taking place within a corporate-sanctioned app environment, revealing the security posture and integrations with third-party plug-ins. From scanning data to monitoring user behavior, detecting malware and ensuring the best security postures, it offers a 360-degree view of what's happening within the app.

REMEMBER

What sets Netskope apart is the seamless communication between the inline and API components, which enables a much richer and more informative real-time risk assessment and reaction, a feature sorely missing in other solutions.

A Shifting Landscape Requires a Broader Perspective

As the digital landscape evolves, apps that were once deemed safe can quickly become vulnerable due to misconfigurations, new features or plug-ins, or changes in their overall security postures. Additional tools within SSE, such as SaaS Security Posture

Management (SSPM), continuously monitor these changes, ensuring that organizations are always a step ahead. But monitoring alone isn't enough; real-time action is crucial.

With a comprehensive security approach, you can avoid overlooking hidden threats. For example, although an app like Salesforce may be perceived as secure, data transfers to channels in other apps that connect to Salesforce can inadvertently expose sensitive information to external users, leading to potential breaches.

The complexity only deepens as the app ecosystem grows. GenAI apps are increasingly integrating with corporate applications, amplifying the challenges. You must meticulously monitor data transfers, especially sensitive data, not just between users and the cloud but also among cloud apps themselves. In this intricate web of digital interactions, tools like CASB are vital — but only an integrated platform like SSE that shares information with all its components and is backed by AI/ML performance and insights can provide a comprehensive approach to enabling the safe use of GenAI.

IN THIS CHAPTER

- Asking the right questions to create a secure environment for your data
- Safeguarding your assets with SkopeAI

Chapter 5
Ten Ways to Ensure Strong Data Security for GenAI

As your organization considers how to get the most out of generative artificial intelligence (GenAI), remember that interacting with GenAI is like nurturing a bright, curious child brimming with boundless potential. It's a job that requires a steady hand, ensuring that every step is safe and informed as you encourage growth. Following are ten pivotal questions to help you, the digital parent, care for that curious child. These critical checkpoints can help you foster a secure environment where innovation and security walk hand in hand.

- **Are your employees data security savvy?** A one-off training session isn't enough. Regular training on the do's and don'ts — like anonymizing personal details, safeguarding intellectual property (IP), and not leaking source code — is essential. Data security must be part of the daily conversation. Well-informed employees are a good defense.

- **Are your security measures at least as agile as GenAI?** The GenAI landscape changes daily, with new apps and hurdles popping up constantly. Keep your ear to the ground. Read up, tune in, and keep an eye on what other organizations are up to.
- **Do you know who uses what GenAI tools and how?** Shadow IT poses a significant risk, with employees using unsanctioned apps and potentially exposing sensitive data. To manage this situation effectively, organizations need advanced analytics tools that identify a wide range of GenAI applications and provide deep insights into their usage patterns, revealing the who, what, and why behind each interaction for comprehensive oversight.
- **Can you accurately measure the risk each tool presents?** With each new GenAI application comes a unique and evolving set of risks. What was safe yesterday may be risky tomorrow. You need to know what tools are in use and the threat level each app poses. This requires a dynamic and scalable approach to risk assessment that can keep pace with the rapid development of these technologies. DIY is not the best approach; you need a trusted partner. A robust security platform should offer insights into threats, compliance issues, and potential data exposures, providing a clear risk profile for each application.
- **What's your strategy for blocking high-risk apps?** The right tools allow you to limit access to GenAI applications that pose a significant risk to your organization's data security. You need a solution to enforce your access policies across all users and devices in real time, ensuring that only safe applications are used.
- **Are you coaching users in real time?** Real-time interaction goes beyond surveillance to provide mechanisms for instant feedback and guidance. When an employee is about to use a GenAI tool, immediate prompts or warnings can direct them toward secure practices.
- **Can you recognize your sensitive data and track its journey?** In today's hybrid work environments, understanding data flows is crucial for comprehending how data is uploaded, processed, and potentially exposed across various networks and devices. A comprehensive solution is needed

to accurately identify and categorize sensitive data, enabling employees to safely use GenAI tools while ensuring that sensitive information is reliably flagged and protected.

- **Will sensitive data stay secure during GenAI use?** Allowing your workforce to leverage the productivity benefits of GenAI tools comes with the responsibility of safeguarding sensitive data. Your security systems must accurately identify and protect sensitive information, such as IP or source code, ensuring that it never gets into the wrong hands.
- **What happens to your data in third- and fourth-party apps?** Data doesn't just stay put in the interconnected web of cloud services. It flows from one app to another, often through third- and fourth-party plug-ins. Monitoring these data flows is crucial to understanding the full scope of your risk exposure. Look at direct user interactions with GenAI isn't enough. You need visibility into the entire chain of custody when it comes to data.
- **Are you taking full advantage of AI to fortify your security?** AI is a transformative force that redefines how we protect our digital environments. From data protection to threat detection and beyond, AI's ability to learn and adapt is a critical component in staying ahead of sophisticated threats and managing complex data ecosystems. The question isn't *whether* to use AI but how comprehensively you can deploy it to safeguard your assets.

Netskope's SkopeAI, a set of AI/ML capabilities native to Netskope's Secure Access Service Edge (SASE), is at the forefront of this AI-driven security revolution, adopting a holistic approach where AI permeates the entire platform. SkopeAI includes the following:

- **SkopeAI data protection:** Netskope's AI/machine learning (ML) classifiers and "Train Your Own Classifier" technology shield unstructured data, adapting to protect against novel data risks with speed and accuracy.
- **AI threat protection:** Netskope provides rapid detection and response to emerging threats, from polymorphic malware to zero-day attacks, ensuring that new malicious domains and Uniform Resource Locators (URLs) are categorized and neutralized swiftly.

- **GenAI and Software as a Service (SaaS) governance:** Netskope provides unparalleled insights and governance for GenAI apps, ensuring that sensitive data remains secure while coaching users on best practices in real time.
- **AI-based user and entity behavior analytics:** By analyzing user behavior, Netskope identifies anomalies that could indicate compromised accounts or insider threats, providing a crucial layer of protection against data exfiltration.
- **Software-defined wide-area network (SD-WAN) optimization:** Leveraging AI, Netskope optimizes network access and performance, proactively identifying anomalies and offering predictive insights that ensure smooth operation across the enterprise.
- **Device access intelligence:** Netskope offers real-time intelligence on newly connected devices, monitoring for behavioral anomalies and potential vulnerabilities.

Generative AI is changing how we do business

The development and use of generative artificial intelligence (GenAI) require a balanced approach, and this book is your guide to exactly that. It explains why you can't just halt the adoption of GenAI technologies (hint: they offer significant benefits). It also explores the potential benefits and pitfalls of GenAI and advocates for responsible management, ethnical guidelines, and robust security practices. *Securing Generative AI For Dummies* emphasizes the importance of aligning AI advancements with human values and societal goals. Finally, it pushes for proactive governance and a strategic framework to mitigate risks while capitalizing on the transformative impacts of GenAI on business and society.

Inside...

- Weigh the risks of GenAI against the rewards
- Cultivate a culture of responsible AI usage
- Consider advanced data loss prevention strategies
- Create a secure environment where innovation thrives
- Explore the potential risks of third- and fourth-party apps
- Learn how to apply Zero Trust principles

netskope

Carmine Clementelli is a cybersecurity expert and technology leader for data protection, cloud security, AI, and Zero Trust at Netskope. He has two decades of experience in networking and security solutions. **Krishna Narayanaswamy** is CTO and co-founder of Netskope, and an acknowledged industry leader and subject matter expert in the use of AI/ML techniques in data security to accelerate and safeguard innovation.

ISBN: 978-1-394-26422-3
Not For Resale

for dummies®
A Wiley Brand

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley’s ebook EULA.