New defenses, new threats

What AI and Gen AI bring to cybersecurity

Capgemini RESEARCH INSTITUTE

#GetTheFutureYouWant

Table of contents

Executive summary

Cybersecurity incidents on the rise: As the number of cybersecurity incidents rises and the threats – including phishing, spear phishing, ransomware, deepfakes, and fraud schemes – grow in sophistication, organizations must enhance their cyber defenses. Our research indicates that 92% of organizations experienced a breach last year, a significant rise from 51% in 2021. The repercussions of these are frequently highly damaging, with around half of organizations reporting estimated direct and indirect losses in excess of $50 million over the past three years. It is clear that new cybersecurity risks are emerging due to AI and Gen AI. At the same time, the use of these technologies presents an opportunity to enhance an organization's cybersecurity. This represents a transformative shift in how security professionals predict, detect, and respond to threats.

Three ways in which AI and Gen AI can pose risks:

1. More sophisticated attacks and more adversaries: Threat actors are exploiting AI, including Gen AI, in various ways. Gen AI lowers barriers for these actors, enabling more sophisticated attacks. Typical uses of Gen AI by cybercriminals include phishing, social engineering, deepfakes, malware development, bypassing security controls, exploiting vulnerabilities, automated hacking, creation of malicious GPTs (Generative Pre-trained Transformers), bypassing security controls by mimicking real user behavior.

97%

of organizations reported security incidents related to Gen AI in the past year.

2. Expansion of the cyber-attack surface: With 97% of surveyed organizations reporting security incidents related to Gen AI in the past year, organizations must contend with an expanded attack surface. "Prompt injection" attacks manipulate Gen AI models and compromise the integrity of their model outputs.

Executive summary

The external attack surface is becoming increasingly complex and multifaceted with the increasing use of AI and Gen AI across various industries. In addition to the traditional attack surfaces that organizations need to protect such as networks, endpoints, data platforms and applications, new vulnerabilities are emerging from these technologies, including conversational AI agents, apps with AI integration, and multiple AI assistants, advisors and new search tools.

Additionally, these technologies can significantly expand the internal attack surface, as internal actors or employees may misuse them — such as, by uploading confidential information to external tools such as ChatGPT. Another concern is 'shadow AI', where unsanctioned applications are installed and used by employees unaware of company policies, outside of IT's control.

3. Lifecycle management of custom Gen AI solutions: The entire lifecycle of Gen AI solutions —from enterprise data collection and model customization to development and maintenance —must be secured to prevent sensitive data used in customization from being compromised and to ensure the availability and integrity of the solution.

Further, Gen AI also brings additional risks, including hallucinations and introduction of vulnerabilities, when used for code generation, which can lead to further security issues. Our research finds that organizations are aware of these threats, and about 60% see the need to boost their cybersecurity budgets consequently.

Integrating AI and Gen AI into cybersecurity and its benefits: On a positive note, three in five organizations believe AI to be essential to effective threat response and a majority rely on AI to strengthen their data security, application security, and cloud security. AI enhances threat detection and reporting by providing real-time response capabilities. It significantly reduces analyst fatigue and guides analysts to the most relevant investigation paths, thereby improving both speed and accuracy. Further, organizations also believe Gen AI will strengthen cybersecurity in the long term. The leadership at more than half of the organizations believe Gen AI can advance their security strategies.

Executive summary

Exploring AI and Gen AI applications cases for security: AI offers a wide variety of use cases for cybersecurity across IT, OT, and the Internet of Things (IoT), and many organizations are already realizing the benefits. Many are also experimenting with security use cases of Gen AI, such as generating threat intelligence and vulnerability assessments.

Enhancing cyber defenses with AI and Gen AI: Organizations must embrace a comprehensive strategy to safeguard their operations. We recommend:

- Develop a clear strategy for integrating AI and Gen AI into existing security systems. Assess the efficiencies gained and risks mitigated relative to the investment into these technologies. Maintain an incident response protocol with actionable instructions for swift, effective action.
- Continuously reassess the security landscape, enabling timely identification of new risks and deployment of adaptive defense mechanisms.
- Acquire necessary infrastructure, including advanced communication systems, data management solutions, and cloud computing resources.
- Establish a robust framework, policies, and governance to ensure data safety and integrity, fostering trust in AI models. Focus on model selection and training tailored to organizational needs.
- Invest in AI and Gen AI-based solutions to integrate with existing security operations centers (SOC) systems enhancing their effectiveness. Gradually integrate AI agents into cybersecurity operations to assist analysts in responding to incidents and mitigating threats effectively. Ensure ongoing monitoring and updates of AI systems to counter evolving threats.
- Invest in comprehensive AI cybersecurity training to ensure employees understand AI's and Gen AI's capabilities and limitations, thereby warranting responsible use.
- Finally, given the rise in cyberattacks, safeguarding business processes and fostering a culture of risk awareness among employees should be a top priority.

Who should read this report and why?

Who?

This report presents an overview of AI as a key aspect of developing and enhancing cybersecurity resilience, with Gen AI both feeding this drive and benefiting from it in terms of protection for Gen AI projects. This report is written for C-suite executives and cybersecurity leaders working across automotive, consumer products, retail, banking, insurance, telecom, energy and utilities, aerospace and defense, high-tech, industrial equipment manufacturing, pharma and healthcare, and the public sector.

Why?

In this research, we explore the role of AI and Gen AI in strengthening cybersecurity. We hope our insights will help C-suite executives and cybersecurity leaders to identify use cases of interest and provide recommendations that organizations can act upon to strengthen their defenses.

This report is based on the findings of a comprehensive survey of 1,000 industry executives and in-depth interviews with selected executives. Excluding public-sector organizations, all the organizations surveyed have annual revenue of over $1 billion and 60% have over $5 billion. All have either already begun to use AI for cybersecurity or are considering it.

See the research methodology at the end of the report for more details on the organizations surveyed.

Definitions

For the purposes of this research, we use the following definitions:

- **Artificial intelligence (AI):** Development of computer systems capable of performing tasks that historically required human intelligence, such as recognizing speech, making decisions, and identifying patterns. AI is an umbrella term that encompasses a wide variety of technologies, including machine learning, deep learning, and natural language processing (NLP).[1]
- **Machine learning (ML):** This is a subfield of AI that uses algorithms trained on data sets to create self-learning models that are capable of predicting outcomes and classifying information without human intervention. It is used today for a wide range of commercial purposes, including suggesting products to consumers based on their past purchases, predicting stock market fluctuations, translating text from one language to another, and much more.[2]
- **Generative AI (Gen AI):** It is a type of AI that has the capability to learn and reapply the properties and patterns of data for a wide range of applications, from creating text, images, and videos in different styles to generating tailored content. It enables machines to perform creative tasks previously thought exclusive to humans.[3]

01 Cybersecurity to the fore

"The [number of] attacks we've encountered has doubled over the past four years."

Julio C. Padilha
Chief Information Security Officer
at Volkswagen and Audi,
South America

An unwelcome by-product of the rapid advancement of digital technologies is the exponential increase in security incidents and breaches, causing serious concern for organizations worldwide. A security incident is defined as any event that undermines the exposure, integrity, or availability of information systems, potentially leading to significant operational and reputational damage. Corence Klop, Chief Information Security Officer at Rabobank, says: *"Over the past year, we had 90 million attacks on our bank. The trend keeps rising, which makes technologies like AI crucial. We need to be smarter in our defense strategies as the number of attacks continues to grow."* Further, Julio C. Padilha, Volkswagen and Audi's Chief Information Security Officer for South America, comments: *"The [number of] attacks we've encountered has doubled over the past four years."* A cyber security expert from a multinational aerospace and defense company elaborates: *"We are witnessing a rise both in the number and variety/type of attacks organizations face. Today, threat actors find it increasingly easier to orchestrate attacks, in terms of complexity, frequency, and accessibility."*

The advent of Gen AI has further complicated this landscape. These technologies, while transformative, have also exposed systems to threat actors. As well as protecting, AI can be weaponized to automate and enhance the sophistication of cyberattacks, making them more difficult to detect and mitigate. US intelligence officials observe that the government regulations must evolve to keep pace with the recent rapid advancements in AI.[4] At the same time, Microsoft-supported studies reveal that 87% of UK organizations are at risk of AI-powered cyberattacks.[5]

Moreover, Gen AI's ability to generate highly realistic synthetic content poses new risks, such as advanced phishing schemes, misinformation campaigns, and deepfakes. Jason Urso, Chief Technology Officer at Honeywell Connected Enterprise, says: *"Prior successful attacks on critical infrastructure involved substantial complexity beyond the capability of an average hacker. However, Gen AI enables less experienced malicious actors to generate malware and initiate sophisticated phishing attacks to gain access to systems and perform automated penetration testing."*[6]

In our research, we found that:

- On average, organizations see around 30 security incidents a day.
- As much as 61% of organizations in the banking sector recorded 10–50 incidents a day, the highest among the sectors surveyed. Aerospace and defense (60%), insurance (58%), and telecom (58%) follow closely.
- Australia, Canada, the Netherlands, and the US have the highest percentages of organizations reporting 50-100 incidents daily (22%, 21%, 20% and 19%, respectively).

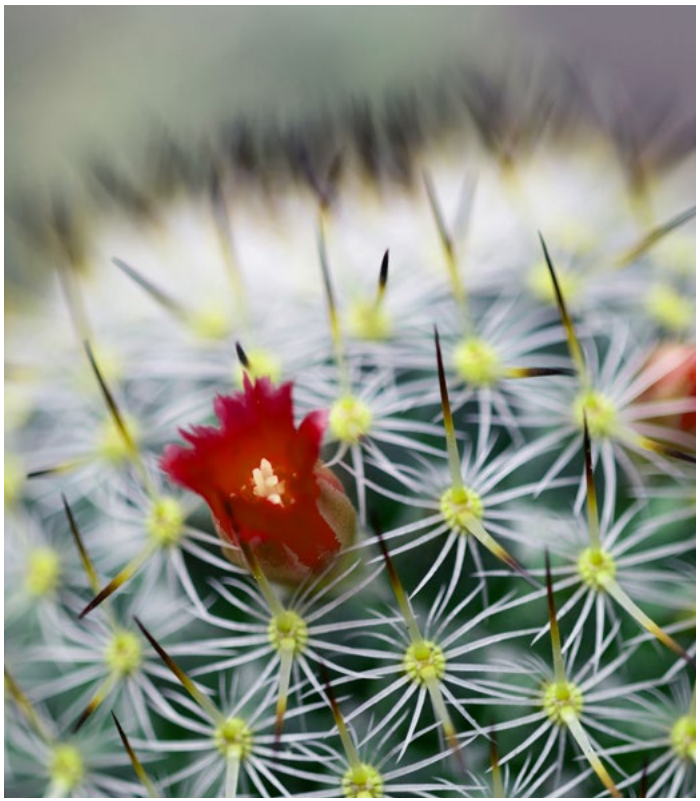
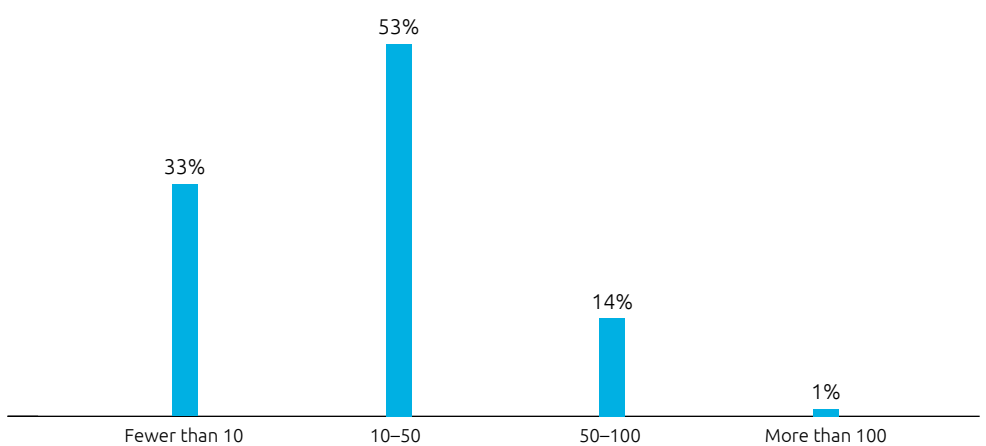


Figure 1.
Around one in six organizations sees more than 50 incidents a day



Source: Capgemini Research Institute, AI and Gen AI in Cybersecurity survey, May 2024, N=1,000 organizations.

As figure 2 shows, in the past three years, the proportion of organizations experiencing one or more breaches has grown from 51% in 2021 to 92% in 2023. The rapid digitalization of organizations and institutions in response to the COVID-19 pandemic likely contributed to a significant increase in cyberattacks, especially in 2022. Threat actors are evolving, with smaller, more agile groups forming to evade law enforcement. Additionally, hackers are increasingly targeting business collaboration tools such as Slack, Microsoft Teams, Microsoft OneDrive, and Google Drive with phishing exploits.[7] More interestingly, when we looked at all three years, almost all (99%) organizations surveyed have had a breach in one or more of the years.

In one massive data breach, a US telecom organization disclosed that malicious actors had stolen the call and text records of more than 100 million consumers from a third-party provider's cloud.[8] Similarly, customers of an American bank holding company were notified of a potential breach of their data at the beginning of March 2024, including their names, account numbers, and card details. They were urged to monitor their accounts for fraudulent activity over the following 12 to 24 months.[9] An American retail firm's crowdsourcing delivery service suffered a cyberattack, with malicious actors accessing the sensitive data of some of its drivers between early December 2023 and early February 2024, including social security numbers, driver's license numbers, and other contact information.[10]

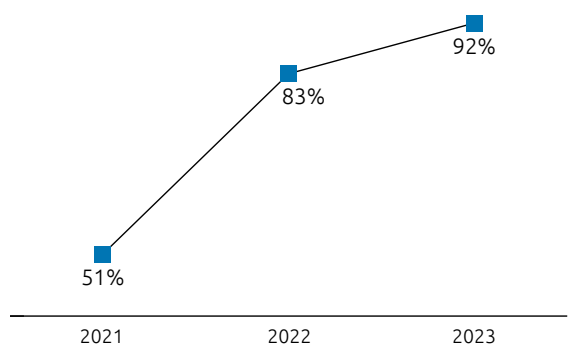
There has been a rising number of cyberattacks targeting the public sector as well. Incidents of data leaks within the public sector in Singapore rose by 10% in 2023, likely due to the increase in digital services.[11] A data breach exposed sensitive information of Canadian government employees.[12] The Australian Government is investigating a "large-scale ransomware" data breach of a health organization, impacting individuals' personal and health information.[13]

99%

of organizations surveyed have had a breach at least once in the past three years.

Figure 2.
Substantial increase in cybersecurity breaches in organizations from 2021 to 2023

% of organizations that experienced a cybersecurity breach, 2021–23



Source: Capgemini Research Institute, AI and Gen AI in cybersecurity survey, May 2024, N=1,000 organizations.

- Australia experienced a dramatic increase, from 48% in 2021 to 97% in 2023; the US also saw a notable rise, from 61% in 2021 to 95% in 2023.
- Across sectors, there is a noticeable increase from 2021 to 2023 in the percentage of organizations experiencing cybersecurity incidents. Automotive, for instance, saw an increase from 49% in 2021 to 91% in 2023. In December 2023, a cyberattack on an automotive original equipment manufacturer's (OEM's) division in APAC compromised the personal data of 100,000 employees and customers. The breach included names, contact details, and government-issued IDs.
- In financial services, 88% and 93% of banking organizations experienced security incidents in 2022 and 2023, respectively, while 89% of the insurance sector reported a breach in both these years.

In the past three years, breaches resulted in direct or indirect damages (such as time taken to fix the breach, reputational damages, etc.) of more than $50 million for half of organizations. In 2023, another automotive OEM's IT security and data protection policies were compromised when nearly 100 gigabytes of confidential data from customers, employees, and business partners was leaked. The breach could potentially result in a $3.3 billion fine for the organization.[14] A US hospitality company suffered a data breach in 2019 that cost over $100 million and exposed the personal information of over 142 million guests.[15] One ransomware attack on a multinational in 2023 resulted in data theft that disrupted operations and cost the company over $27 million in damages.[16] Both automotive and insurance sectors have relatively high percentages (12% and 10%, respectively) of organizations experiencing financial damages in excess of $100 million.

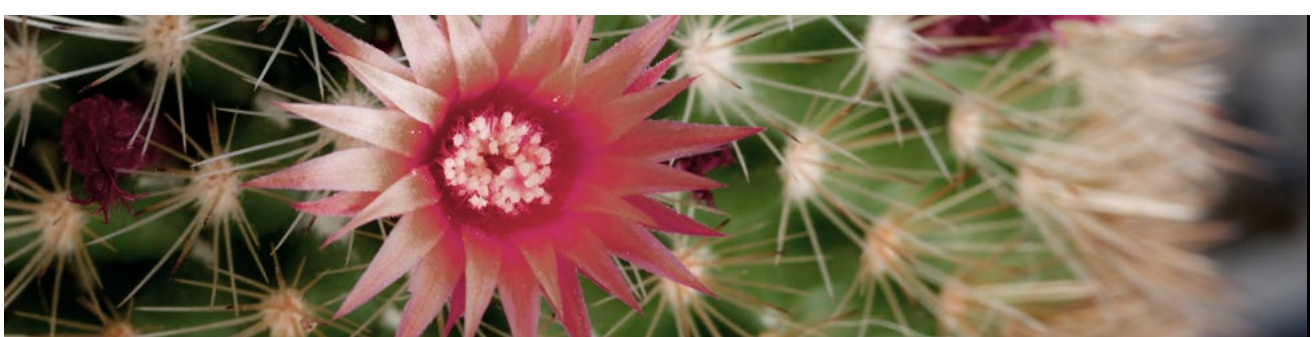
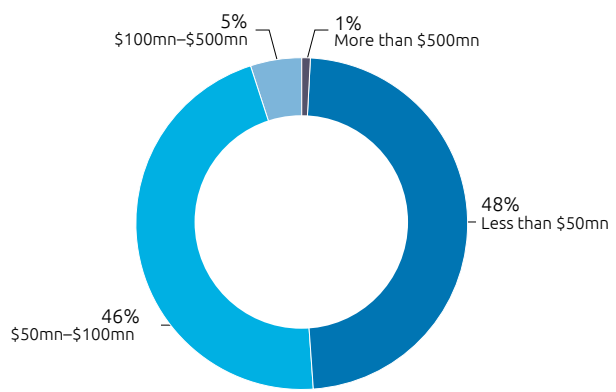


Figure 3.

Breaches resulted in average financial damages of $50 million for half of organizations

Direct and indirect financial damage to organizations resulting from a breach in the past three years



Source: Capgemini Research Institute, AI and Gen AI in cybersecurity survey, May 2024, N=991 organizations that suffered a breach in the past three years.
Indirect damages imply time taken to fix the breach, reputational damages, etc.

Organizations with revenues between $1 billion and $5 billion incur $1.60 in damages for every $100 earned, while those with revenues exceeding $20 billion experience $0.40 in damages. Organizations with revenues in between average $0.70 in damages per $100 earned.

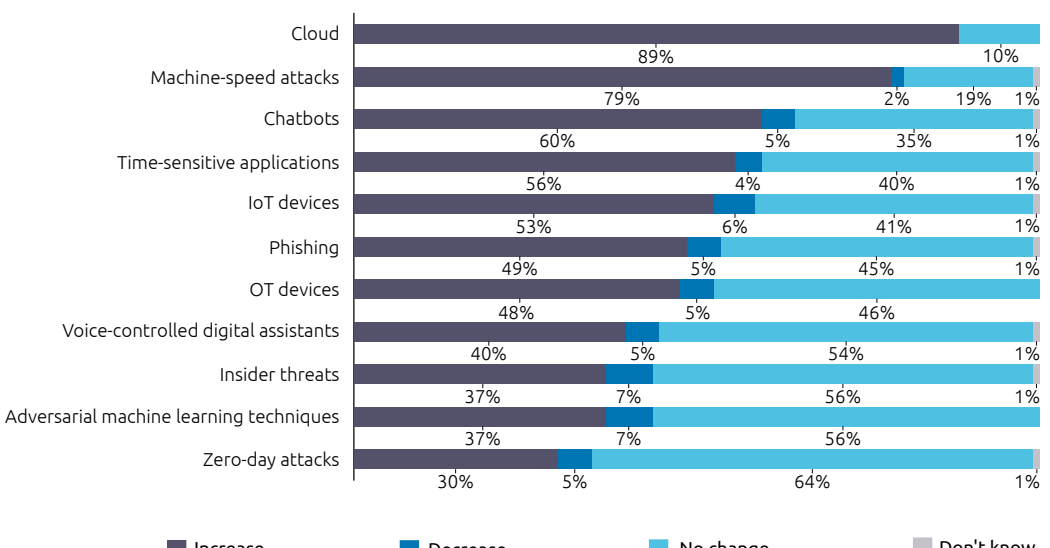
As organizations across sectors adopt digital technologies and interconnected systems, they also increase their attack surfaces, with an ensuing rise in security incidents affecting cloud services, machine-speed attacks,[17] chatbots, time-sensitive applications, and IoT devices in the past year.

89%

rise in cloud related security incidents in the past year.

Figure 4.

Cloud saw a nearly 90% rise in security incidents in the past year



Source: Capgemini Research Institute, AI and Gen AI in cybersecurity survey, May 2024, N=1,000 organizations.

- Cloud services incidents saw the highest reported increase across all sectors, with the highest percentages in aerospace and defense (95%), followed by retail and consumer products (both 92%).
- Sectors such as aerospace and defense (89%), consumer products (87%), retail and banking (both at 81%), high-tech (78%), and pharma and healthcare (84%), are particularly vulnerable to machine-speed attacks.

"The scalability of cyberattacks poses a significant challenge. However, we can use AI technologies to assist analysts in the Security Operations Center by swiftly searching multiple sources, advising on alert responses, and automating actions to handle high volumes more effectively."

Corence Klop
Chief Information Security Officer
at Rabobank

02 The AI and Gen AI risk landscape

The AI risk landscape is evolving rapidly, driven by the latest advancements in technology, increasing adoption across industries, and the emergence of more sophisticated (Gen) AI models. As AI systems become more integral to decision-making, the potential for unintended consequences, biases, and security vulnerabilities grows, necessitating a reevaluation of risk management strategies to ensure trustworthy, secure, and responsible AI deployment.

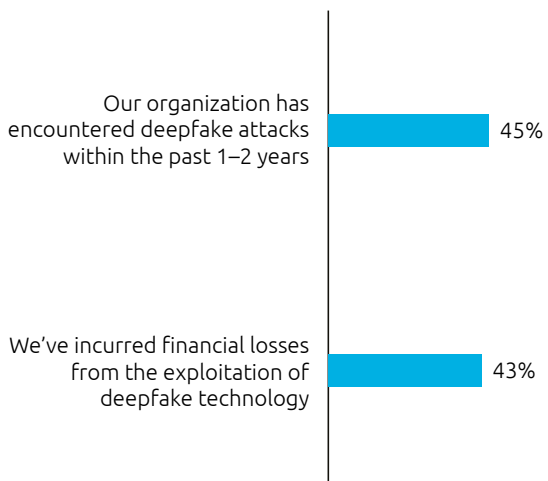
The increase in the proportion of organizations that faced breaches (seen in figure 2) and also the increase in machine-speed attacks (as shown in figure 3) could point to the fact that more and more threat actors are relying on AI and Gen AI today to cause cyberattacks. The ways in which AI and Gen AI can pose risks can be categorized into three areas.

1. More sophisticated attacks and more adversaries

Threat actors can leverage AI, including Gen AI, in a number of ways. Typical uses of Gen AI by cybercriminals include phishing, social engineering, deepfakes, malware development, bypassing security controls, exploiting vulnerabilities, automated hacking, creation of malicious GPTs (Generative Pre-trained Transformers), bypassing security controls by mimicking real user behavior. For instance, these actors can generate advanced phishing emails, scripts for file manipulation, or code to evade detection.[18] In January 2024, the UK government's National Cyber Security Center (NCSC) released an assessment that highlighted that AI will almost certainly increase the volume and heighten the impact of cyberattacks over the next two years.[19]

- **Sophisticated threats:** Gen AI can lower the barriers for threat actors, resulting in heightened cyber risks and more sophisticated attacks. Additionally, attackers might manipulate AI systems to produce incorrect predictions or deny customer service.
- **Prompt injection risks:** This involves using malicious inputs to manipulate AI and Gen AI models, compromising their integrity. Attackers can embed harmful scripts and commands in images, causing the model to comply. Multimodal prompt injection attacks can exfiltrate data, redirect queries, spread misinformation, and override safety measures, leading to risks such as fraud and operational sabotage.
- **Social engineering attacks:** AI and Gen AI-based social engineering attacks such as deepfakes or phishing emails are particularly challenging to defend due to their high personalization and realism. These attacks use custom lures in chats, videos, or audio, mimicking individuals with remarkable accuracy, targeting multiple systems or individuals with tailored messages. Recently, there have been several notable cases of deepfakes created with Gen AI.

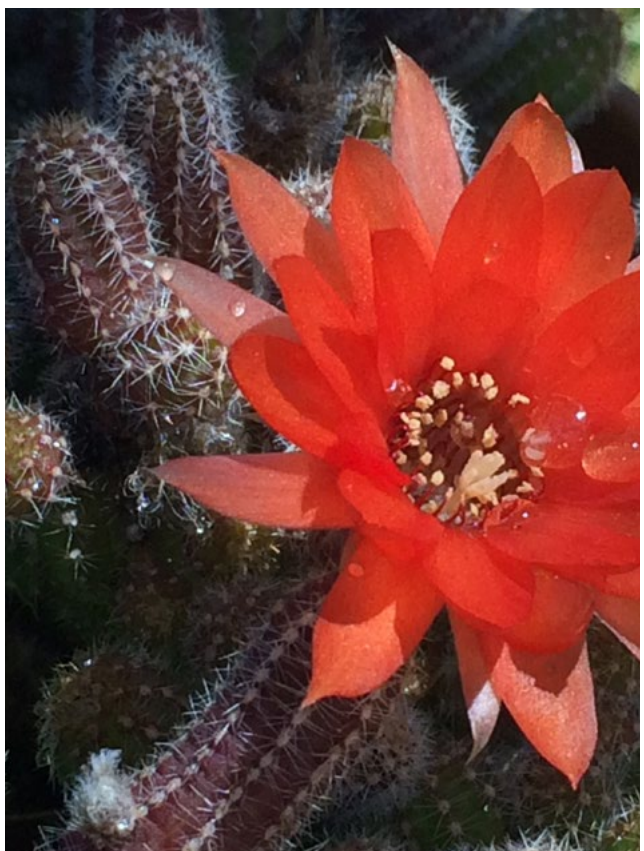
Figure 5.
More than two in five organizations have suffered financial losses arising from the use of deepfakes



Source: Capgemini Research Institute, AI and Gen AI in cybersecurity survey, May 2024, N=1,000 organizations.

An executive at a sports car manufacturer received unexpected messages that appeared to come from the CEO, asking for the signing of several agreements. Using deepfake technology, the attacker conducted a live phone conversation, delivering a voice that closely mimicked the CEO's. However, the attack was thwarted when the executive noticed inconsistencies and realized something was wrong.[20] After a design and engineering firm lost $25 mn to a deepfake scam in early 2024, the Asia branch of another multinational organization became the next victim when a digitally recreated version of its CFO deceived an employee at the Hong Kong office. This fraudulent impersonation led the employee to transfer half a million dollars, purportedly to fund a new branch of the organization.[21]

- **AI worms:** Further, researchers warn of a new shift in the cyber landscape – AI worms. These have the potential to spread from one system to another, potentially compromising data security or deploying malicious software in the process. As Gen AI systems gain autonomy, they are becoming a medium of susceptibility to exploitation by malicious actors. These AI worms can exploit this interconnectedness, spreading through Gen AI ecosystems and infecting numerous devices. These worms can potentially compromise critical infrastructure, steal sensitive data, or wreak havoc by disseminating fake news. They can outmaneuver conventional security by analyzing user behavior to craft personalized attacks.[22]



2. Expansion of the cyber-attack surface

The launch of ChatGPT in November 2022 pushed excitement around Gen AI technologies to fever pitch. From 2023, organizations worldwide began to experiment with the new tools, piloting a flurry of use cases. According to our latest research, nearly one-quarter (24%) of organizations have enabled Gen AI capabilities in some or most of their functions and locations.[23] However, as stated previously, the increased adoption of Gen AI brings heightened vulnerability as well as opportunities. In our research, we found that 97% of organizations encountered breaches or security issues related to the use of Gen AI in the past year. Organizations, today, have to deal with an expanded attack surface area that is becoming increasingly complex and multifaceted. Besides protecting the traditional attack surfaces such as networks, endpoints, data platforms and applications, they also need to secure the newer applications enabled by AI and Gen AI such as conversational AI agents, apps with AI integration, and multiple AI assistants, advisors and new search tools.

Further, AI and Gen AI can significantly expand the internal attack surface, as internal actors or employees may misuse them.

Legal risks: Without stringent governance and oversight, the use of Gen AI can also amplify legal risks, such as exposing trade secrets, proprietary information, and customer data due to inadequate data security measures.

Shadow AI: Another concern is the rise of shadow AI within the organization, where unsanctioned AI applications are installed and used inappropriately. This poses a dual security risk: one from user behavior (such as disclosing confidential information) and the other from the applications themselves (if they have security flaws or vulnerabilities). In a survey, Microsoft found that 75% of knowledge workers around the world use Gen AI at work and that 78% of AI users bring their own AI to work (tools not provided by their organization).[24] Our latest research on Gen AI revealed that unauthorized usage among employees is relatively common. Among the 39% of organizations with a ban or limitation policy, half of them say there is still unauthorized usage of Gen AI in the workplace.[25] Furthermore, our recent research on Gen AI for software engineering also reports that 63% of software professionals using Gen AI use it in an unauthorized manner, while only 37% use a licensed tool provided by their organization.[26]

A few organizations have taken the extreme step of completely banning their employees from using AI tools such as ChatGPT. For instance, a multinational organization banned ChatGPT after its engineers accidentally leaked confidential elements of the company's source code via these tools.[27] However, the rise of shadow AI calls to question whether an outright banning of these tools is indeed effective.

3. Lifecycle management of custom Gen AI solutions

Securing the entire lifecycle of Gen AI solutions is critical to ensuring the protection of sensitive data and the reliability of the system. From the initial phase of enterprise data collection, where valuable and sensitive information is gathered, to the customization of Gen AI models tailored to specific business needs, every stage must be safeguarded. During development and deployment, vulnerabilities, as identified in the insert "Mapping and mitigating Gen AI risks," can arise that could expose confidential data or compromise the system's performance. Implementing robust security measures across this lifecycle not only protects sensitive data from being compromised but also guarantees the availability, reliability, and integrity of the solution, allowing organizations to maximize the benefits of AI while minimizing risks.

With AI, organizations also have the challenge of ensuring the AI models they build, or use, are free of biases. The adoption of Gen AI could further increase an organization's vulnerability to issues such as hallucination. This is when the model produces an apparently authentic and valid output that it has, in fact, partially or wholly invented. Earlier this year, a Canadian airline was ordered to pay

compensation to a customer after its bot fed inaccurate information to a customer, misleading them into buying a full-price ticket.[28]

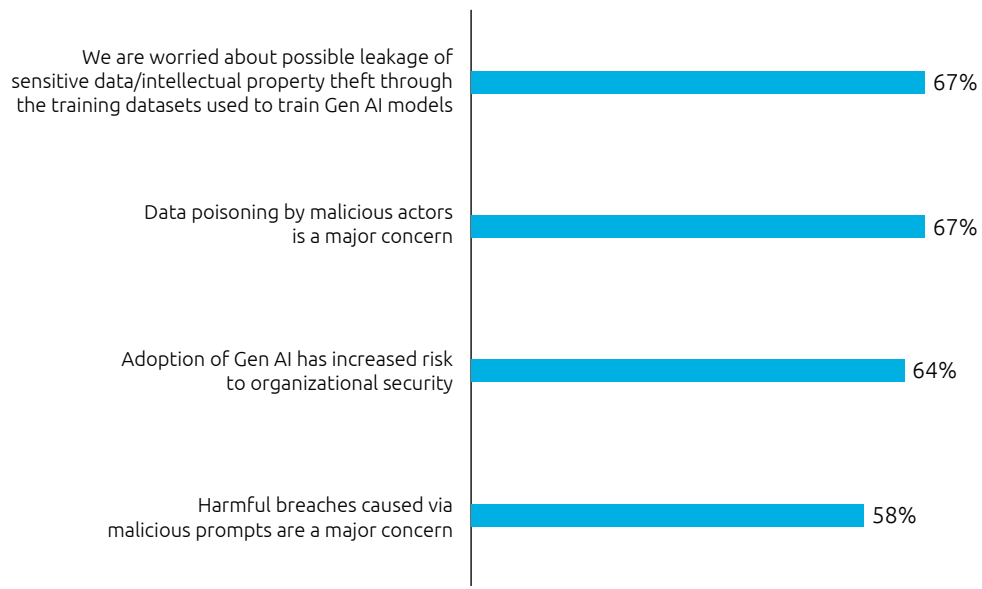
Additionally, organizations are increasingly relying on Gen AI for code generation. While this can improve proficiency, it can also introduce well-known vulnerabilities (e.g., the MITRE CWE Top 25 Most Dangerous Software Weaknesses) into the code.[29]

Two in three organizations are wary of increased exposure to threats

While organizations are excited about Gen AI's potential, they are aware of the risks that come with adoption. Most risks associated with Gen AI are not novel to application security but rather amplified versions of existing concerns. However, some risks are unique to AI, including model drift, model theft, and data poisoning. Gen AI specifically introduces additional risks and vulnerabilities, such as biased, harmful, or inappropriate content generation, hallucinations, and prompt injection attacks.

In 2023, Apple restricted the use of ChatGPT and other external AI tools, such as GitHub, for some of its employees. The organization was concerned that employees could leak confidential data while using these tools.[30] Similarly, Amazon has prohibited employees from using third-party Gen AI tools particularly for handling confidential data. This policy is intended to prevent data-ownership issues and protect sensitive company information.[31] In data poisoning, the AI model is compromised, for example, by injecting malicious data into the training dataset or manipulating the training data to create vulnerabilities. Researchers have, to date, discovered about 100 machine learning (ML) models uploaded to Hugging Face, an open-source platform for ML, that could act as an enabler of the injection of malicious code into user machines.[32]

Figure 6.
Two in three organizations are worried about data leakage and data poisoning

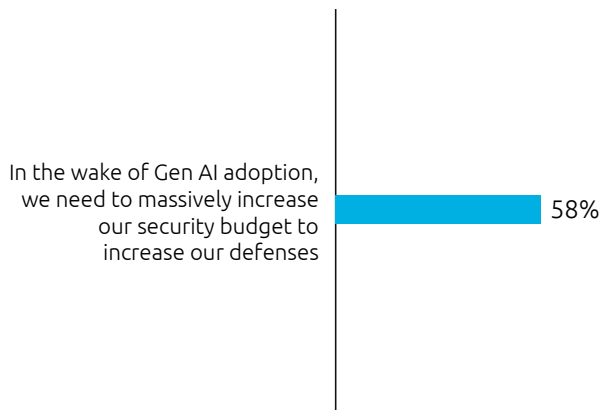


Source: Capgemini Research Institute, AI and Gen AI in cybersecurity survey, May 2024, N=1,000 organizations.

Most organizations agree they need to increase their security budgets to fortify their defenses

Our research into harnessing the potential of Gen AI shows that organizations have achieved overall productivity improvements of nearly 8% on average over the past year, and by 2026, they expect significant improvements in operational efficiency, cost reduction, and sales.[33] Our current findings also highlight that organizations are increasingly aware of the heightened security risks and the necessity to enhance their investment in cybersecurity measures. As illustrated in figure 7, most organizations acknowledge the need to increase their allocation toward Gen AI to strengthen their defenses.

Figure 7.
Nearly 6 in 10 organizations believe they need to increase their security budget to bolster their defenses



Source: Capgemini Research Institute, AI and Gen AI in cybersecurity survey, May 2024, N=1,000 organizations.

Organizations and increasingly governments are spending more than ever to protect their databases and critical defense systems from cyberattacks. Cybersecurity now constitutes 12% of overall technology budgets, up three percentage points since 2020.[34]

Mapping and mitigating Gen AI risks

The value chain of Gen AI can be complex, with multiple external parties and introduces diverse risks across its stages. A crucial element of developing a secure AI strategy is understanding the various risks the organization is responsible for mitigating.

The key to building trust in AI and Gen AI models lies in:

- Understanding the use case
- Mapping associated risks
- Evaluating these risks and
- Implementing appropriate mitigation strategies

The below architecture highlights the various risks that can be introduced at each layer. This framework serves as a guide to help organizations map the risks and create a comprehensive and secure AI strategy. This helps them address the challenges inherent in the Gen AI pipeline, ensuring a robust and trustworthy approach to AI implementation and deployment.

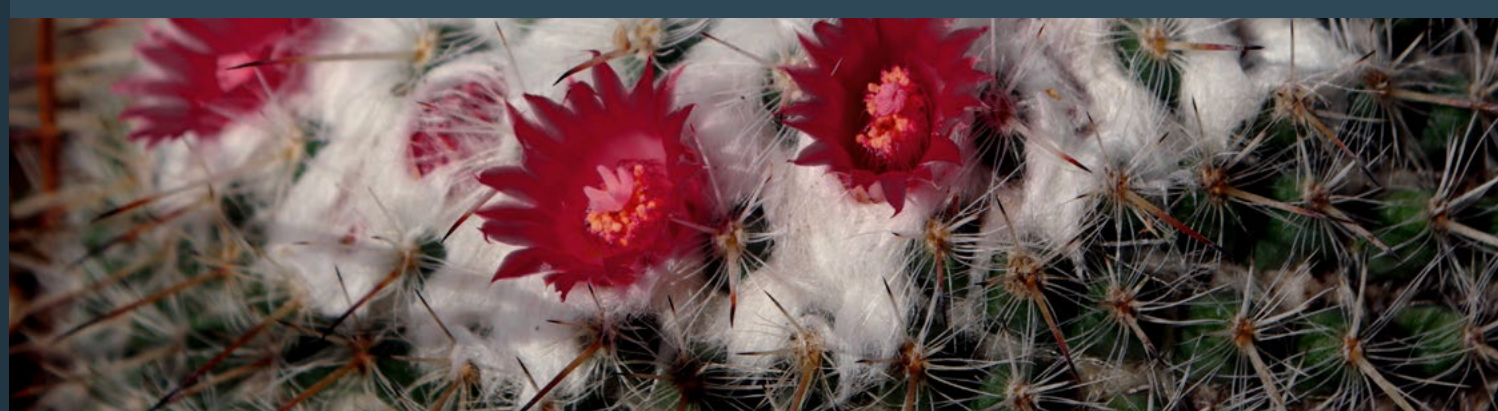
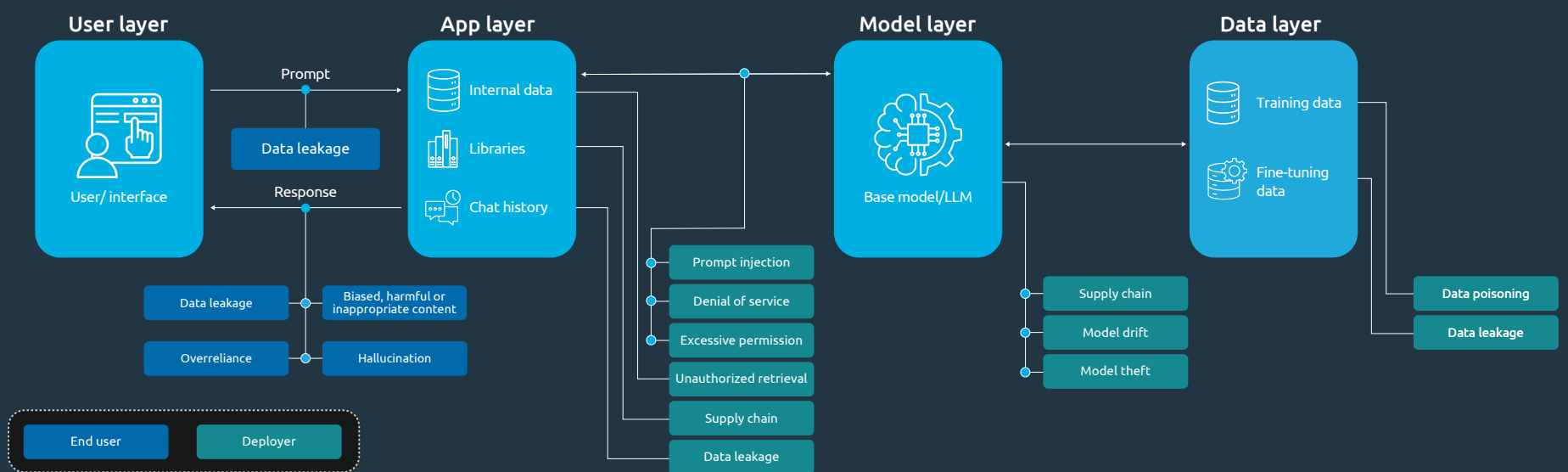


Figure 8.
Gen AI risk reference architecture



Source: Capgemini Group Cybersecurity, Trusted AI offer.

'End users' refers to those that interact with Gen AI only through the prompt input and output interface. 'Deployers' refers to anyone who develops, implements, integrates, or manages any part of the Gen AI system. As can be seen in the figure, security concerns or risks can exist at the data level or at the model layer too, in addition to the application layer and the user layers.

In order to secure the initiatives, organizations must understand and minimize each of these risks. Developers are introducing risk-mitigation features. For instance, in Q3 2024 Google is expected to preview its Model Armor protection system, which will enable customers to inspect, route, and protect foundation model prompts and responses, mitigating risks such as prompt injections, jailbreaks, toxic content, and sensitive data leakage. Model Armor will integrate with products across Google Cloud, including Vertex AI.[35]

03 We rely on AI

Three in five organizations believe AI is paramount for detecting and responding to attacks

Organizations' reliance on AI to reinforce their security infrastructure is intensifying, reflecting AI's potentially transformative effect on cybersecurity. Most cybersecurity solutions available on the market and utilized by organizations rely on traditional AI/ML technologies. AI enhances threat detection and response by rapidly analyzing vast amounts of data and identifying patterns and predicting potential breaches. This proactive approach significantly reduces response times and minimizes damage. Moreover, AI-driven automation streamlines routine security tasks, allowing human experts to focus on more complex issues. In the long term, AI's self-learning should allow the technology to adapt to evolving threats.

A cyber security expert from a multinational aerospace and defense company adds: *"Currently, we face the challenge of dealing with completely unknown threats, where traditional methods fail. This is where AI emerges as a crucial tool, as its strength lies in its ability to analyze vast amounts of data and detect specific behaviors. The integration of AI and automation is essential to close the gap between detection and response, ensuring rapid and effective cybersecurity processes."*

The relevance of AI to cybersecurity is affirmed by 66% of organizations prioritizing its use in this context. Additionally, 60% recognize AI as essential to effective responses to cyber threats, emphasizing its strategic significance.

60%

of organizations recognize AI as essential to effective responses to cyber threats, emphasizing its strategic significance.

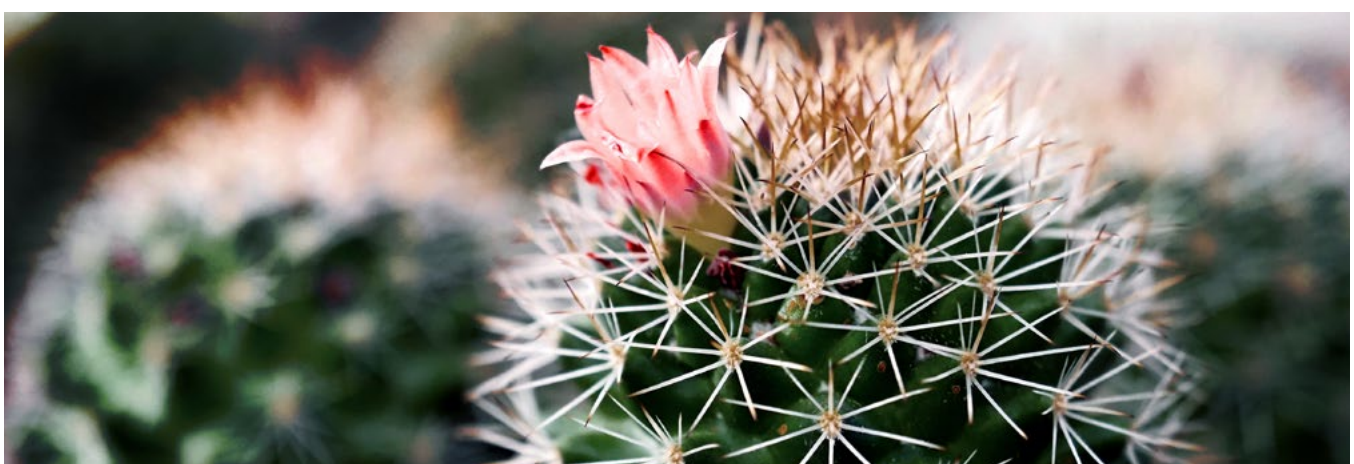
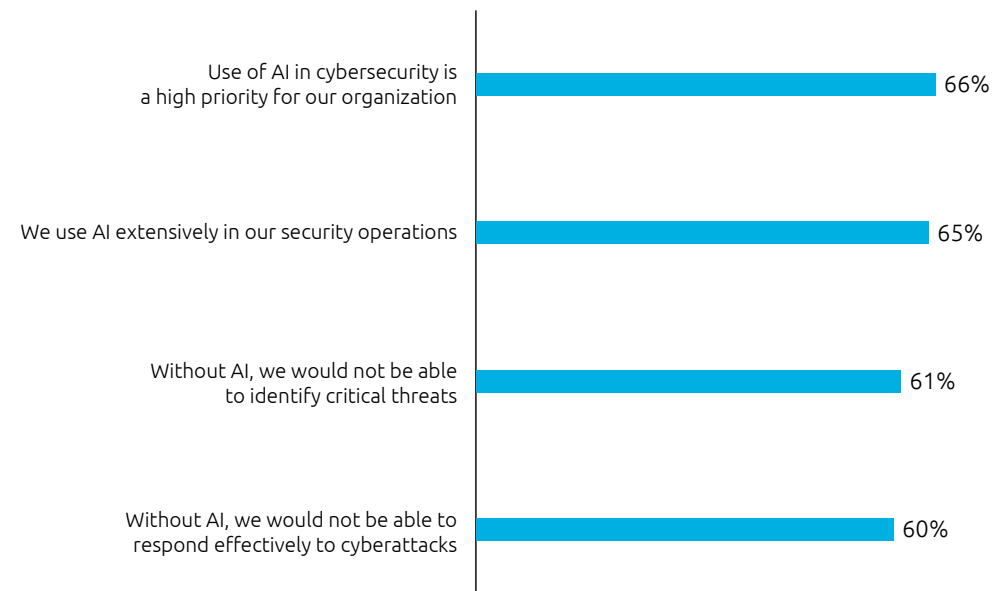


Figure 9.
Two in three organizations prioritize the use of AI in cybersecurity



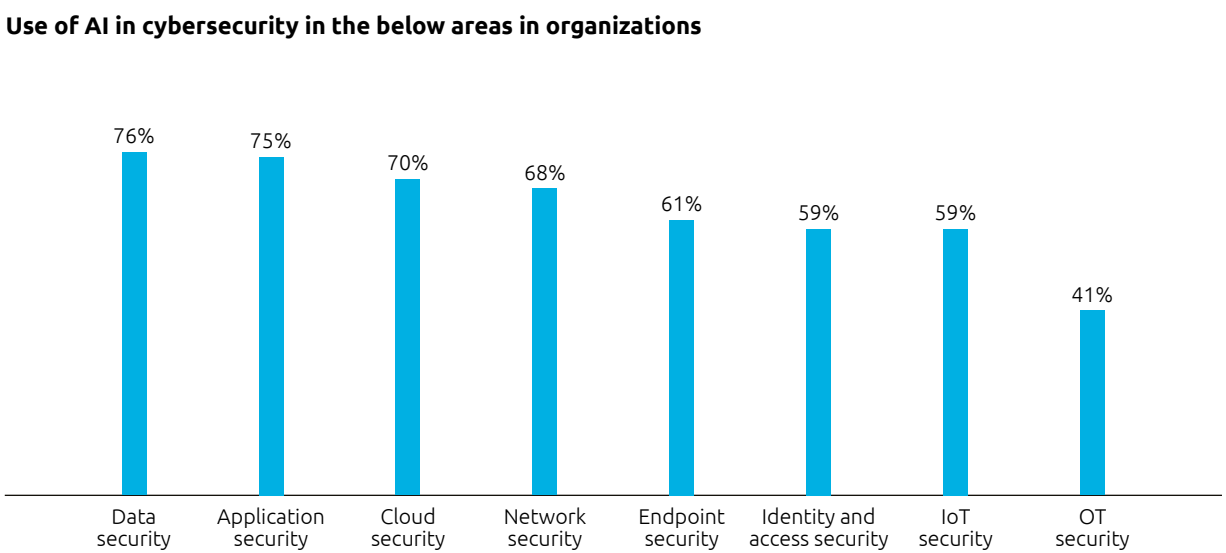
Source: Capgemini Research Institute, AI and Gen AI in cybersecurity survey, May 2024, N=1,000 organizations.

As shown in figure 10, data security (76%) and application security (75%) are the areas where organizations most commonly use AI.

66%

agree that the use of AI in cybersecurity is a high priority for their organization.

Figure 10.
More than three in five organizations use AI in cybersecurity for data security



Source: Capgemini Research Institute, AI and Gen AI in cybersecurity survey, May 2024, N=1,000 organizations.

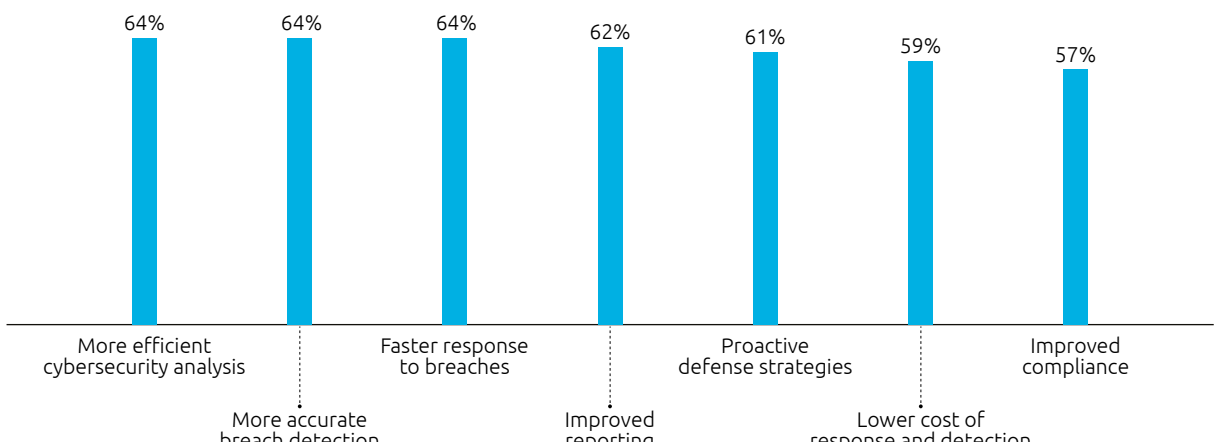
- Japan leads in the use of AI in data security (92%), application security (90%), and cloud security (77%).
- The highest adoption rates in identity and access security are in Japan, the Netherlands, Sweden and Singapore (70%, 68%, 65% and 64%, respectively).
- Banking exhibits strong AI adoption in data security (84%), application security and cloud security (both 80%).
- High-tech has high AI adoption for application security (88%) and data security (68%), whereas Industrial equipment manufacturing has high adoption in data security (80%) and application security (79%).

AI enables a faster response to breaches

By automating some elements of threat detection and response, AI minimizes manual intervention, enhances efficiency, and ensures a cost-effective, rapid and robust defense against evolving cyber threats. *"We record billions of cybersecurity events every month. Without AI, it would be impossible to analyze them all effectively,"* says Adriano Oliveira, responsible for cybersecurity at CNP Seguradora, a subsidiary of the CNP Assurances group. Mastercard's AI capabilities, for example, enable real-time detection and prevention of payment scams, cutting response times and mitigating financial and reputational losses.[36] MetLife uses AI to detect fraudulent claims swiftly, reducing investigation time and costs by analyzing data patterns, identifying anomalies, and streamlining the fraud detection and resolution process.[37]

Figure 11.
More than three in five organizations find AI provides higher efficiency and accuracy in cybersecurity

Share of organizations that have realized various benefits from the use of AI in cybersecurity

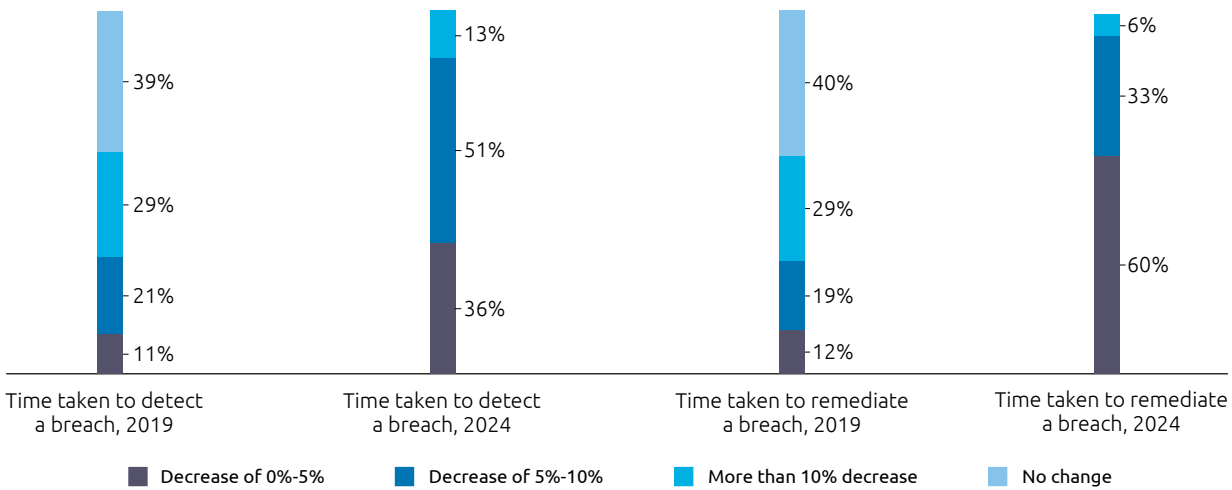


Source: Capgemini Research Institute, AI and Gen AI in cybersecurity survey, May 2024, N=1,000 organizations.

Over time, the use of AI by threat actors has rapidly increased, but organizations have not matched this pace in deploying AI for threat detection or remediation (see figure 12). This disparity highlights a substantial opportunity for organizations to advance their AI capabilities to enhance early breach detection and improve response strategies.

Figure 12.
Two in three organizations say they have cut the time taken to detect a security breach by at least 5%

Share of organizations that realized time savings after implementing AI in their security operation centers (SOCs)



Source: Capgemini Research Institute, AI in cybersecurity survey, July 2023, N=833 organizations; AI and Gen AI in cybersecurity survey, May 2024, N=1,000 organizations.

More than 60% of organizations reported that a reduction in their time-to-detect of at least 5%, and nearly 40% of organizations say remediation time fell by 5% or more after implementing AI in their SOCs.

Corence Klop from Rabobank comments: *"The scalability of cyberattacks poses a significant challenge. However, we can use AI technologies to assist analysts in the Security Operations Center by swiftly searching multiple sources, advising on alert responses, and automating actions to handle high volumes more effectively."*

"We record billions of cybersecurity events every month. Without AI, it would be impossible to analyze them all effectively."

Adriano Oliveira,
Responsible for cybersecurity at CNP Seguradora, a subsidiary of the CNP Assurances group

04 Gen AI will reinforce cybersecurity

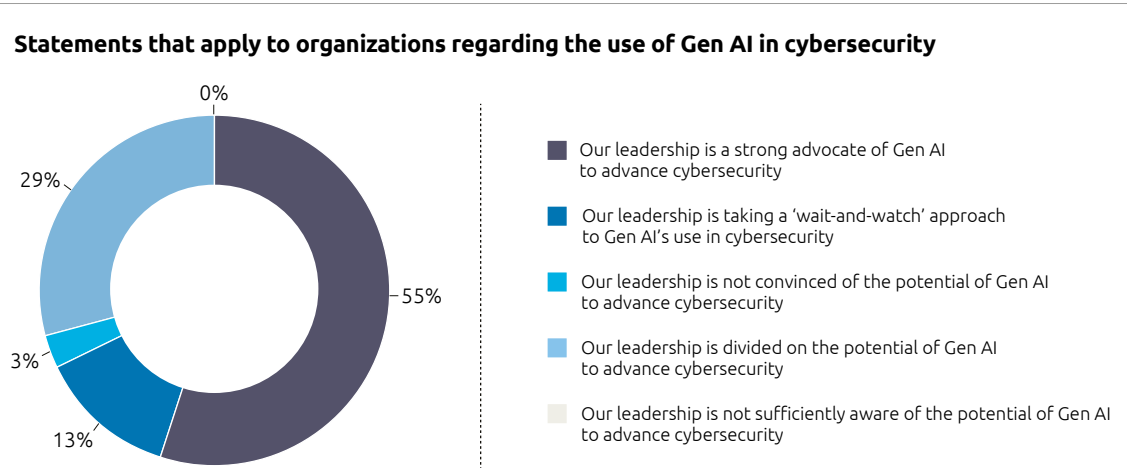
In the long term, Gen AI will strengthen cybersecurity

Gen AI will enable advanced threat simulation and proactive defense strategies. A more than solid 61% of respondents foresee Gen AI strengthening cybersecurity in the long term; a further 62% anticipate it playing a proactive role in vulnerability detection. These insights reflect growing confidence in Gen AI's pre-emptive security measures.

Hélio Cordeiro Mariano, Chief Information Officer at Cooperativa Central Ailos, states: *"The interaction and the capacity of Gen AI to learn more about what we are doing, and answer questions based on prompts, could accelerate the identification of issues in the environment."*

Gen AI's ability to anticipate and neutralize sophisticated cyber threats will boost organizational resilience against an ever-evolving digital threat landscape. Frédéric Pégaz-Fiornet, Head of Digital Health and Cybersecurity for France, Belgium and Luxembourg at Siemens Healthineers, adds: *"We have started using Gen AI for diagnosis and other applications. I am confident that in the coming months or years, AI and generative AI will advance significantly. There are numerous potential applications for generative AI, especially in enhancing proactive measures against attacks. Currently, many CIOs are reactive rather than proactive, lacking full engagement with their monitoring tools during cyberattacks."*

Figure 13.
More than half of organizational leadership believes in Gen AI for security



Source: Capgemini Research Institute, AI and Gen AI in cybersecurity survey, May 2024, N=1,000 organizations.

"We have started using Gen AI for diagnosis and other applications. I am confident that in the coming months or years, AI and generative AI will advance significantly. There are numerous potential applications for generative AI, especially in enhancing proactive measures against attacks. Currently, many CIOs are reactive rather than proactive, lacking full engagement with their monitoring tools during cyberattacks."

Frédéric Pégaz-Fiornet
Head of Digital Health and Cybersecurity for France, Belgium and Luxembourg at Siemens Healthineers

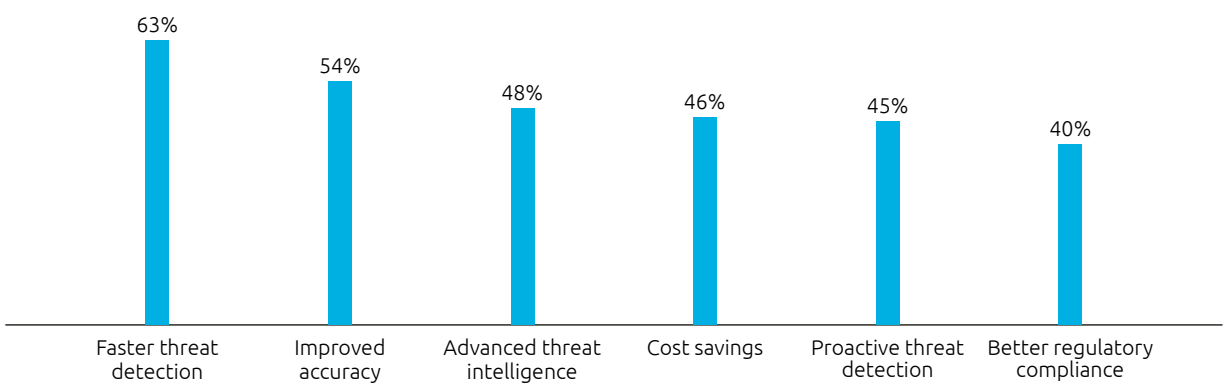
- In both Japan and Australia 62% of organizations say their leadership is a strong advocate of Gen AI to advance cybersecurity, the highest among countries surveyed.
- As many as 70% of organizations in the public sector and 65% of organizations in the high-tech and industrial equipment manufacturing sub-sectors say their leadership is a strong advocate of Gen AI to advance cybersecurity, the highest across sectors.
- Over half (54%) say Gen AI can create realistic threat scenarios to enhance cybersecurity analyst training.

More than half of organizations anticipate faster threat detection and increased accuracy through the use of Gen AI

With the right data and the right model, Gen AI's ability to analyze and interpret vast datasets swiftly allows for early identification of potential threats, though it remains a challenge for more advanced use cases. As a result of its promising capabilities, organizations are increasingly integrating Gen AI to fortify their defenses, anticipating a marked improvement in their ability to counteract sophisticated cyber threats effectively (see figure 14). Interestingly, fewer than 50% believe Gen AI's ability to save costs. Gen AI's training costs and operating costs are certainly a cause for concern and can pose a barrier for adoption.[38]

Figure 14.
More than three in five organizations anticipate faster threat detection by using Gen AI

Benefits anticipated from use of Gen AI in cybersecurity



Source: Capgemini Research Institute, AI and Gen AI in cybersecurity survey, May 2024, N=1,000 organizations.

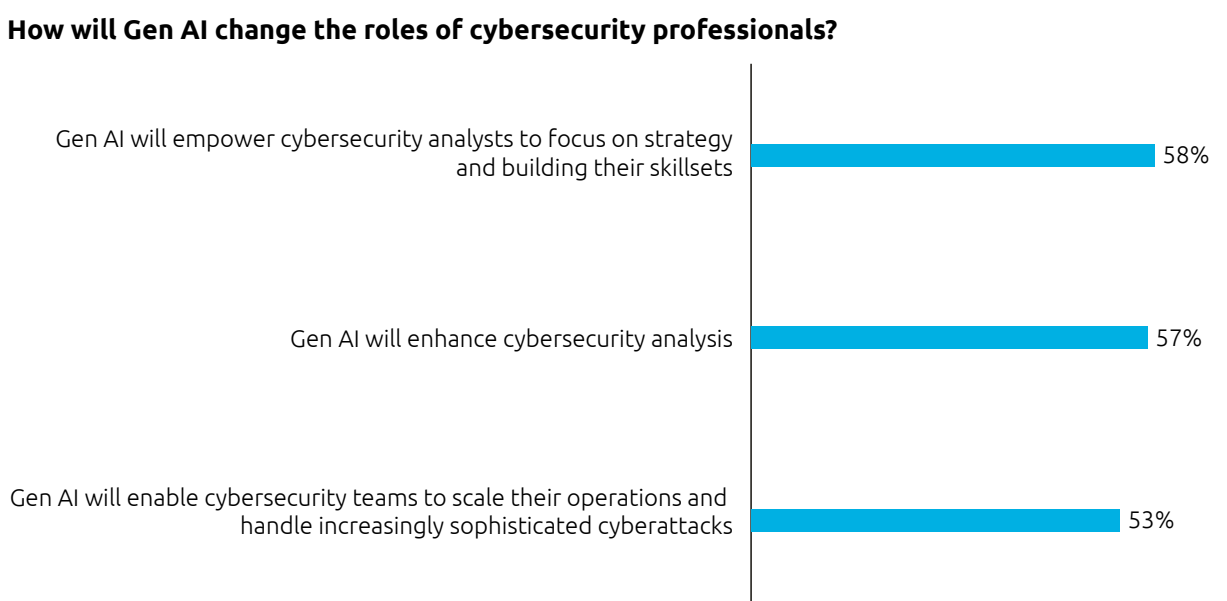
Nearly three in five organizations believe Gen AI will enhance cybersecurity analysis

Gen AI's generative capabilities and simulation tools improve overall security measures, making analysts more efficient and effective and substantially freeing them to concentrate on more complex threats. Fifty-seven percent of organizations acknowledge the importance of specialized training for using Gen AI tools in tasks such as threat detection, incident response, and vulnerability management.

58%

of organizations say Gen AI will empower cybersecurity analysts to concentrate on strategy for combating complex threats

Figure 15.
Over half (58%) of organizations say Gen AI will empower cybersecurity analysts to concentrate on strategy for combating complex threats



Source: Capgemini Research Institute, AI and Gen AI in cybersecurity survey, May 2024, N=1,000 organizations.

Frank Hamilton Moraes, IT cybersecurity superintendent, and Luciano Carolino, IT security specialist at Bradesco Bank, comments: *"Gen AI can support decision-making, enabling swift action and providing valuable support for security analysts, particularly when utilized in a supervised or semi-supervised manner. This approach is particularly important in complex and critical environments."*

Gen AI can also create sophisticated simulations to train security systems and personnel, enhancing preparedness for real-world attacks. JPMorgan Chase uses AI and Gen AI models to detect fraud by analyzing transaction patterns, identifying anomalies, and improving real-time monitoring, enhancing overall security and fraud prevention.[39]

"Gen AI can support decision-making, enabling swift action and providing valuable support for security analysts, particularly when utilized in a supervised or semi-supervised manner. This approach is particularly important in complex and critical environments."

Frank Hamilton Moraes
IT cybersecurity superintendent at Bradesco Bank

Luciano Carolino
IT security specialist at Bradesco Bank

05 Exploring AI and Gen AI use cases

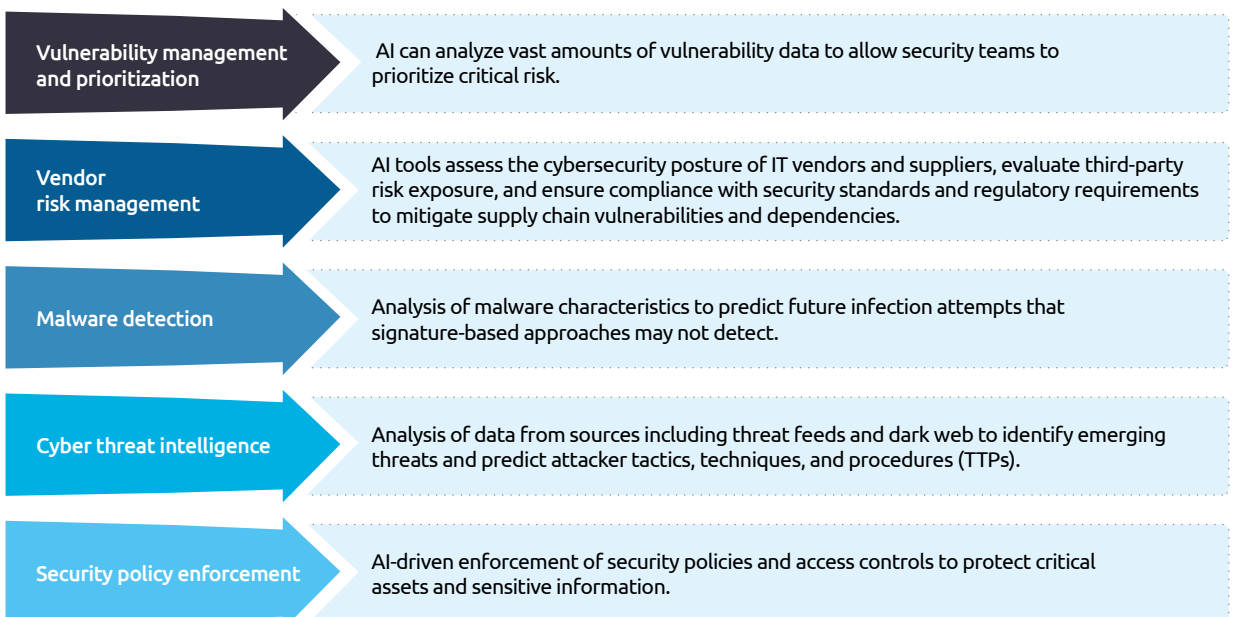
AI use cases across the organization

Exciting AI applications in cybersecurity arise across information technology (IT), operational technology (OT), and Internet of Things (IoT). Below, we look at a selection of notable use cases.

Top use cases: IT

Organizations use AI in IT security to detect threats, automate responses, and analyze vast amounts of data, enhancing cybersecurity for end-point devices, networks, identity, access management, etc. (see figure 16).

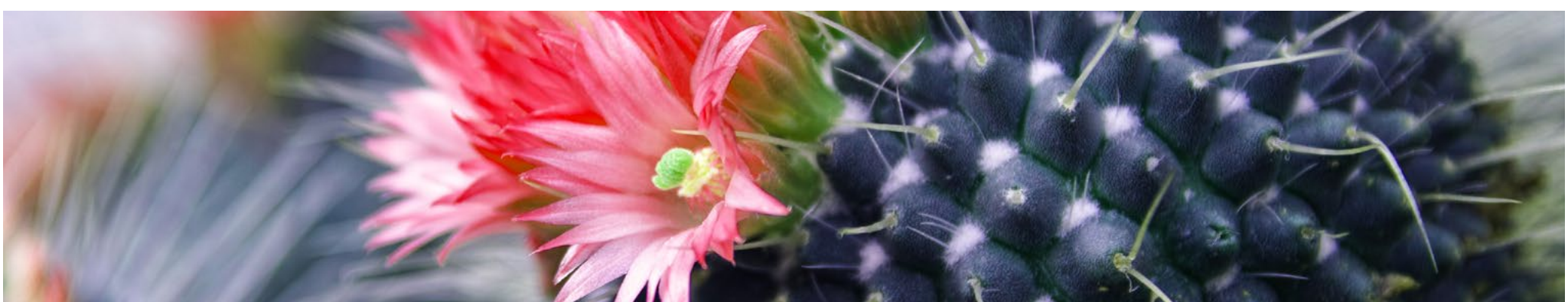
Figure 16.
Top five AI use cases in IT



Source: Capgemini Research Institute, AI and Gen AI in cybersecurity survey, May 2024, N=1,000 organizations.

Here are examples illustrating how organizations are leveraging AI in IT:

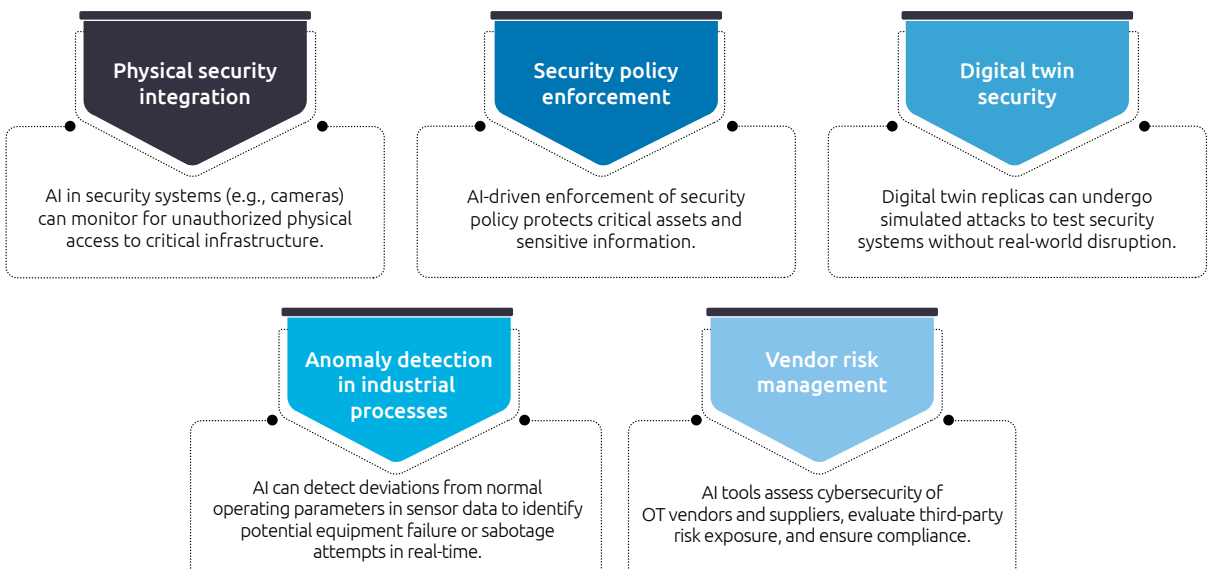
- Cybersecurity company CheckPoint and China's Qihoo 360 Netlab security firm discovered a botnet called "Reaper," which had affected over 1 million internet-connected devices, including routers and IP (Internet Protocol) cameras. AI threat intelligence can help identify and block botnet traffic, preventing attacks.[40]
- The US federal government uses an AI platform to analyze billions of events in real time, protecting against all types of attacks, from commodity malware to sophisticated state-sponsored intrusions.[41]
- American Express utilizes AI to analyze customer transactions in real time and identify suspicious activity such as unusual spending patterns, location inconsistencies, and known fraudulent activities.[42]
- JPMorgan Chase utilizes AI-powered vulnerability management solutions to monitor extensive networks for vulnerabilities and prioritize patching efforts based on severity and probability of exploitation.[43]
- Cisco uses AI-driven analytics to enhance customers' existing identity infrastructure, offering insights into their entire identity population; securing vulnerable accounts; revoking unused and risky privileges; identifying behavioral anomalies; and preventing high-risk access attempts.[44]
- PayPal uses AI to examine each transaction for red flags and identify and block malicious web content and potential cyber threats.[45]



Top use cases: OT

Organizations use AI in OT (hardware and software used to monitor and control devices, processes, and infrastructure in industrial settings) to detect anomalies, predict threats, and automate industrial control, supervisory control, and data acquisition systems (see figure 17).

Figure 17.
Top five AI use cases in OT



Source: Capgemini Research Institute, AI and Gen AI in cybersecurity survey, May 2024, N=1,000 organizations.

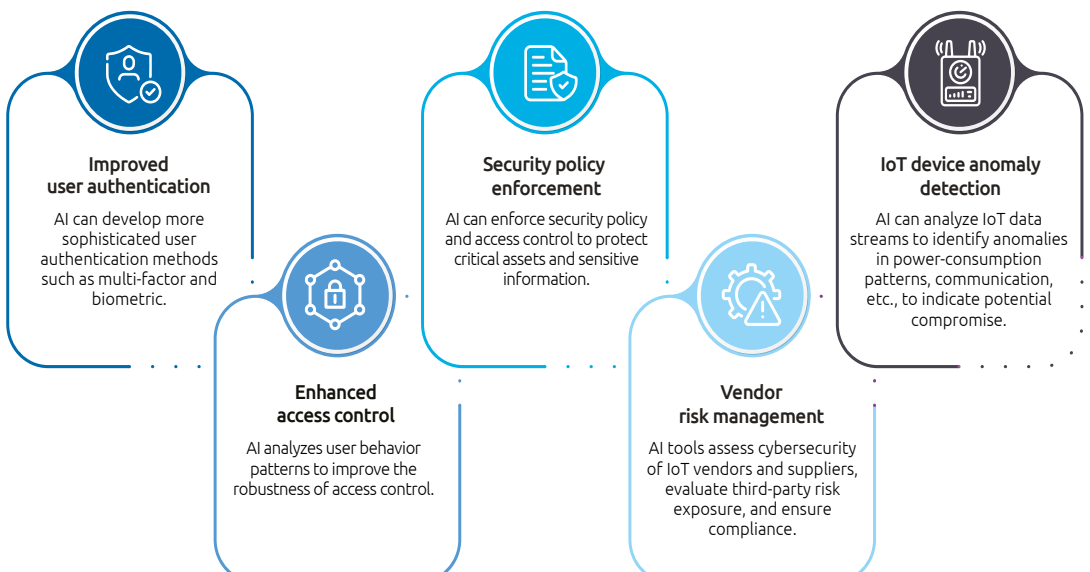
The below examples illustrate how organizations use AI in OT:

- The US Transportation Security Administration (TSA) has used facial-recognition technology at various airports nationwide since 2019. The TSA's more than 50,000 agents work across 430 local airports, assessing around 8 million passengers each week.[46]
- Honeywell uses AI to swiftly analyze vast amounts of data from industrial control systems, identifying unusual patterns or behaviors. Its AI platform continuously reviews patterns from past incidents and adapts that information to mitigate new emerging threats.[47]
- BBVA, a Spanish multinational financial services organization, has a dedicated AI-driven cybersecurity hub to provide holistic security response to each operational and business element of the bank, including anticipating threats and preparing operational tactics, offering resilience strategies, and protecting BBVA's data processing centers (DPCs).[48]

Top use cases: IoT

AI enhances IoT security by enabling real-time monitoring of device or plant sensors, threat detection, and automated response (see figure 18).

Figure 18.
Top five AI use cases in IoT



Source: Capgemini Research Institute, AI and Gen AI in cybersecurity survey, May 2024, N=1,000 organizations.

Ring and Nest both incorporate AI features and use computer vision algorithms to detect and track motion, send alerts to connected devices, and offer real-time video streaming via mobile apps.[49]

At least two in five organizations have conducted pilot programs using Gen AI for security

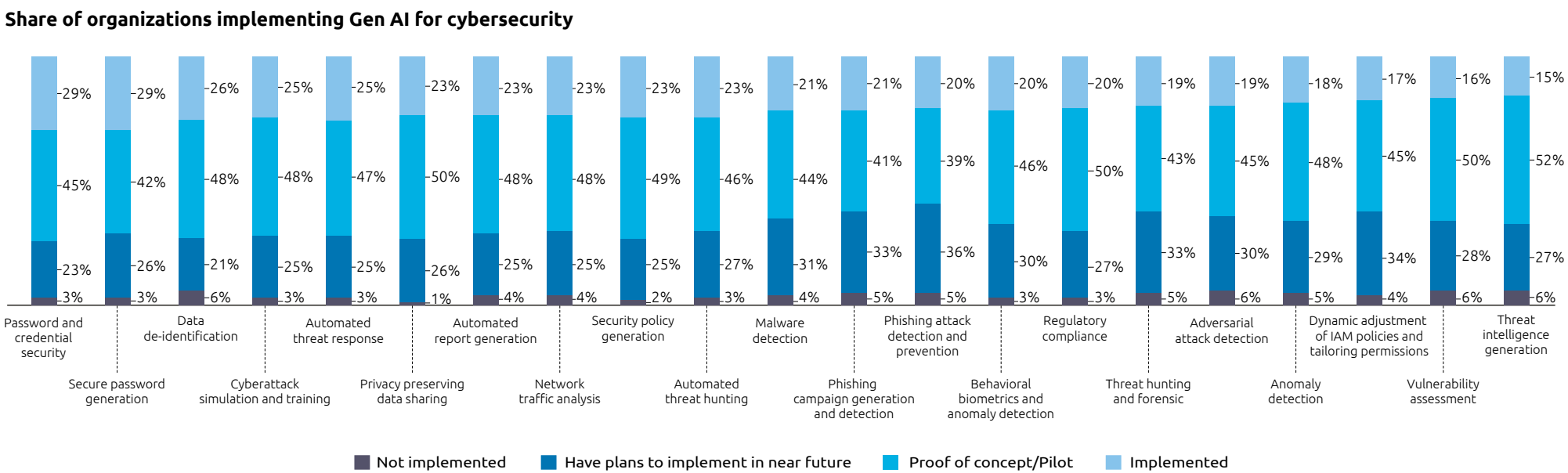
Gen AI can advance an organization's security operations. For instance, LLMs trained and fine-tuned for security use cases, such as the Google Cloud Security AI Workbench, can help analysts identify potential threats.[50] Through prompts, analysts can classify, synthesize, and summarize these insights in an intuitive way and preferred formats (e.g., translate attack graphs to human-readable explanations). Further, these tools provide assistance that allows the development of generalist talent to a security analyst role.

Organizations have begun integrating Gen AI into their cybersecurity operations. As figure 19 shows, around 40%-50% have initiated proofs of concept (PoC) or pilots. Additionally, nearly three in ten (28%) organizations plan to implement Gen AI in cybersecurity in the near future.

28%

of organizations plan to implement Gen AI in cybersecurity in the near future.

Figure 19.
At least two in five organizations are currently piloting Gen AI for security



Source: Capgemini Research Institute, AI and Gen AI in cybersecurity survey, May 2024, N=1,000 organizations, N varies per use case, ranging from 337 to 663.
"Implemented" means the use case has been deployed in one or more locations.

Luciano Valdomiro Dos Santos, Head of Cybersecurity Risk in a Brazilian retail bank, says: *"Gen AI can create synthetic datasets and insights that help us test and improve our security measures and controls without exposing real user data."*

Below are examples of how organizations use Gen AI in their cybersecurity operations:

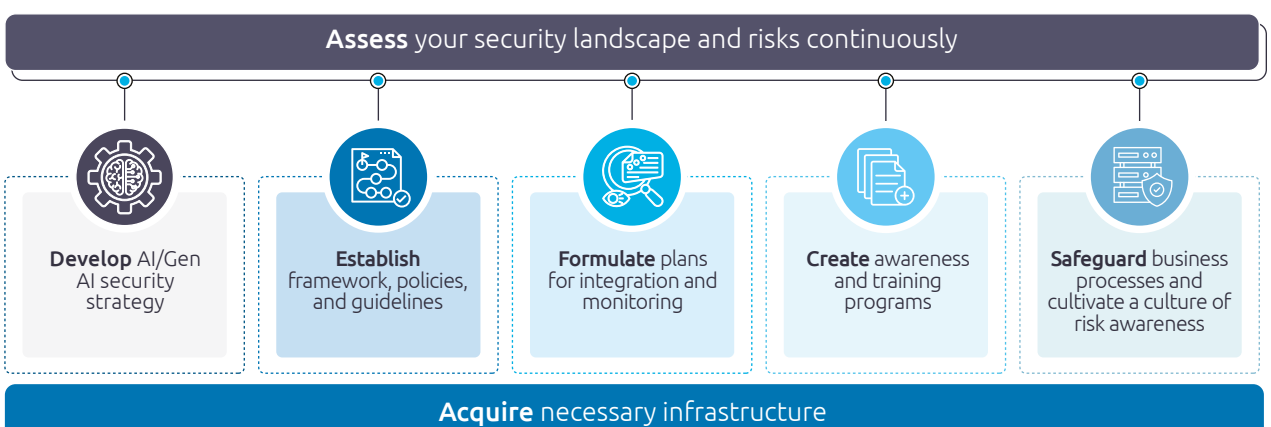
- Symantec, a division of Broadcom Inc., a US designer, developer, manufacturer, and global supplier of software products, is embedding Gen AI into its security platform in a phased rollout. The Gen AI will detect, understand, and remediate sophisticated cyberattacks.[51]
- Mastercard uses Gen AI-based predictive technology to protect future transactions against emerging threats by doubling the detection rate of compromised cards, reducing false positives during card fraud detection by up to 200%, and increasing the speed of identification of at-risk/compromised merchants by 300%.[52]
- Brazilian beauty retail and cosmetics, Grupo Boticário employs real-time security models to detect, prevent, and respond to potential fraud.[53]

Per our research, around 20% of organizations have scaled their deployments for their cybersecurity operations, less than the extent of adoption of AI in cybersecurity. The extensive data usage and interactions with models can drive up costs quickly and be a major barrier for large-scale implementation. In addition to the licensing costs of the models (or training costs), significant expenses include maintenance and updates for models and data pipelines. Critical models needing immediate responses, such as those used in live customer service, add further costs for scaling. Managing multiple tools and complex infrastructure also raises operational expenses, complicating large-scale implementations. Secondly, the quality of the Gen AI models' output depends on the size and quality of the data on which the models are trained. Without this, organizations may find it difficult to obtain the right results from their Gen AI models. These cost- and data-related concerns could hinder the large-scale adoption of Gen AI for strengthening defenses.

06 Recommendations: Using AI and Gen AI to strengthen your cyber defenses

As AI and Gen AI-based risks continue to rise, organizations' security strategies should rely heavily on using the same technologies to defend their assets and operations. Based on our survey, interviews, and experience, we propose recommendations for strengthening an organization's defenses below.

Figure 20.
Strengthening the defenses of your organizations using AI and Gen AI



Source: Capgemini Research Institute analysis.

Develop an AI/Gen AI security strategy

Roadmap and selected use cases

In the era of Gen AI, developing a robust security strategy is crucial to safeguarding organizational assets and data integrity. Organizations should implement the following:

- Formulate a clear strategy to integrate AI and Gen AI into existing security systems. Adopt a phased approach that evolves current systems to address nuanced risks and threats targeting AI technologies. Both AI and Gen AI have their own applications. For instance, while AI can help with threat or anomaly detection, automating responses to common incidents, and analyzing large datasets, Gen AI can help create realistic phishing simulations, and develop sophisticated scenarios for testing defenses.
- Evaluate the efficiencies gained or risks mitigated relative to the investment in the Gen AI tools. Developing a clear strategy for measuring these factors is essential. Beyond the initial investment, it is important to evaluate ongoing operating costs and determine whether the long-term return justifies these expenses. Investing in a tool just because it incorporates Gen AI, without looking into the overall costs, could result in inefficiencies over the long term.

- Focus on foundational elements (such as data) and gradually scale up, evaluating associated security risks, as identified in the insert "Mapping and mitigating Gen AI risks." Identify low-risk, high-value use cases to assess performance and security implications, such as vulnerability patching and enhancing incident-response capabilities. Establish a risk reference architecture to standardize and benchmark the identification and mitigation of risks, providing a consistent approach for managing them effectively. Volkswagen's Julio C. Padilha comments: *"The approach to AI and Generative AI should be grounded in a risk-based framework, which entails identifying and cataloging all potential risks associated with the technology or project prior to its initiation."*
- Conduct regular reviews of your cybersecurity strategy and AI policies to adapt to evolving threats and technological advancements. Collaboration among IT teams, legal experts, business leaders, and other stakeholders within and outside the organization, pooling knowledge, experience and resources, is crucial. Luciano Valdomiro Dos Santos from a Brazilian retail bank comments: *"Cybersecurity exercises that simulate attacks and prepare for security resilience underscore the significance of this cooperation across diverse internal departments and sectors, encompassing not only banks but also various private industries, to ensure continuity in the face of an ever-changing cyber threat landscape."*

Incident response protocols

To effectively manage serious security incidents like data breaches, leaks, ransomware attacks, or loss of sensitive information, organizations should establish a global incident response and management team available 24/7, with strict protocols in place for response and mitigation. The ISO/IEC 27035-1:2016 framework, for instance, provides protocols for analyzing, assessing, responding to, and containing cybersecurity threats, ensuring alignment with international standards.

- Outline initial actions to contain and mitigate the incident, such as isolating affected systems, shutting down compromised accounts, and blocking unauthorized access.
- Leverage Gen AI to provide insights and recommendations to first responders. However, given its relative novelty, it is crucial to keep humans involved in the decision-making process to ensure that actions are taken thoughtfully and appropriately rather than automatically initiating actions from the outset.
- Ensure a secure chain of custody for investigation evidence and maintain detailed records of all security incidents. This documentation is essential for future threat analysis, response planning, and proactive vulnerability mitigation.
- Conduct thorough post-incident analysis to identify gaps in security controls and response procedures to enhance overall cybersecurity resilience.

Alexandra Foster, former Managing Director at BT and now an independent consultant, comments: *"It's crucial to strengthen your data backup and recovery tools. Many industries collaborate on incident response and prevention, but it's essential to actively use AI to optimize the backup process and ensure swift recovery in case of data loss."*

Integrating AI, especially Gen AI, into incident response planning empowers organizations not only to react swiftly to security incidents but also to anticipate and mitigate future threats. It can simulate a wide range of cyberattack scenarios, enabling incident response teams to fine-tune their response.

"The approach to AI and Generative AI should be grounded in a risk-based framework, which entails identifying and cataloging all potential risks associated with the technology or project prior to its initiation."

Julio C. Padilha
Chief Information Security Officer at Volkswagen and Audi, South America

Continuously re-assess security landscape and risks

This pre-emptive approach enables the timely identification of new risks and the deployment of adaptive defense mechanisms.

- Identify critical assets such as sensitive data, IP, and key infrastructure, prioritizing resources and tailoring measures to reduce breach risks and minimize damage. Targeted protection ensures critical operations and information remain secure.
- Re-evaluating security posture encourages enhanced threat detection and response, helping to maintain system integrity and safeguard sensitive information.

This approach also supports compliance with regulatory requirements and fosters a proactive security culture. In our research, 62% of organizations believed Gen AI will allow them to identify vulnerabilities proactively. Gen AI can further help in interpreting complex regulations and producing detailed reports required for compliance.

62%

of organizations believe Gen AI will allow them to identify vulnerabilities proactively.

Acquire necessary infrastructure

AI and Gen AI adoption demands more sophisticated communications, data management, and cloud computing infrastructures, specialized AI processors, and extensive data storage. Exploring the synergy between Gen AI and advanced hardware is crucial. Organizations can either develop the necessary infrastructure upgrades in-house or purchase them from specialized providers.

Organizations should prioritize hardware security upgrades. High-performance components such as powerful graphics processing units (GPUs) and specialized AI accelerators enhance Gen AI models, enabling rapid processing and analysis of large volumes of data. This enables AI to swiftly identify anomalies and potential threats in complex networks.

Hardware security modules (HSMs) and Trusted Platform Modules (TPMs) provide robust encryption and secure-key management. Organizations should also adopt features such as hardware root of trust (RoT, a systemically foundational software component) and fingerprinting to enhance defense layers.[54]

"The convergence of Gen AI and hardware innovations advances cybersecurity by enabling faster, more accurate threat detection, enhancing data protection through secure computing environments, and improving user-authentication processes," affirms Mohit Sagar, Chief Executive Officer and Editor-in-Chief at OpenGov Asia, a content platform, initiating dialogue across public-sector CIOs and technology experts.[55]

As AI expands in data centers and at the edge, integrating AI-based security mechanisms into data centers and network infrastructure will be crucial. However, faster hardware and processors also consume more energy and contribute to carbon footprints. Instead of buying GPU farms and HSMs, organizations should consider leveraging the cloud to expand the required capabilities. Utilizing cloud resources can improve sustainability by optimizing the use of shared infrastructure, which leads to more efficient resource utilization and a reduced carbon footprint.

Establish framework, policies, and guidelines

Data pipelines and readiness

A robust, well-integrated, and scalable data platform ensures data safety and integrity, which are vital to nurturing trust in AI models. Further, the effectiveness of AI, specifically Gen AI in cybersecurity depends on the size and quality of the data, and the algorithms used to analyze it. Currently, neither the data volume and quality nor the algorithms are sufficiently advanced for the widespread use of Gen AI in cybersecurity – aside from a few targeted applications where data and algorithms are reliable.

To address this, organizations should:

- Identify and classify data sources, files, and unstructured data, especially confidential data (e.g., customer information, business transactions data, etc.).
- Track and evaluate data sources for accuracy, completeness, consistency, timeliness, and reliability.
- Record frequency and purpose of data access to identify dependencies and potential bottlenecks and prevent breaches.

Moreover, such data platforms must scale seamlessly as organizations grow their workforces, build complex data infrastructures, and manage larger volumes of data. Corence Klop from Rabobank affirms: *"I would prioritize organizing your data, ensuring you have a centralized repository to work from. This begins with establishing a single, comprehensive source of data. Rather than starting from scratch, use existing resources and conduct experiments to see what works for you."*

Governance policies

AI and Gen AI raise important questions regarding data governance, intellectual property (IP), bias mitigation, and responsible utilization of AI-generated content. Establishing a dedicated team or department to oversee AI governance at the organizational level is crucial. Highlighting the significance of this governance mechanism, Hélio Cordeiro Mariano from Cooperativa Central Ailos explains: *"We're establishing an Innovation team to engage all departments effectively. Our focus spans beyond security leads, encompassing how the company navigates experiments, tests, and simulations. We're crafting a process to prioritize impactful actions and demonstrate their value. There are numerous market alternatives, not all are tailored to our specific needs. Our approach involves starting small, testing for value generation, and scaling promising initiatives swiftly. Governance, planning, and aligning expectations are crucial discussions, particularly when integrating AI to benefit our business comprehensively."* Further, such teams within organizations need to:

- Clearly articulate and document policies governing the use, storage, and transmission of data in AI and Gen AI systems.
- Establish policies and guidelines for the development and deployment of AI and Gen AI tools to ensure ethical practices and effective governance.
- Define clear guidelines for employees on the use of AI and Gen AI tools. Emphasize adherence to data privacy regulations and highlight the repercussions of misuse. This empowers employees to make informed decisions and reduces organizational risk. The CISO at a payment solutions company in Brazil comments: *"Anyone can potentially compromise the LLM, which is the primary security risk we are concerned about. Additionally, privacy breaches are also a significant concern; sensitive information should not be accessible to just anyone. Even within a closed environment, access to all company information should be restricted. These are among the challenges we currently face."*
- Control and restrict data access to necessary stakeholders through role-based access control and techniques such as blocking, hashing, and limiting platform connectivity to external networks.

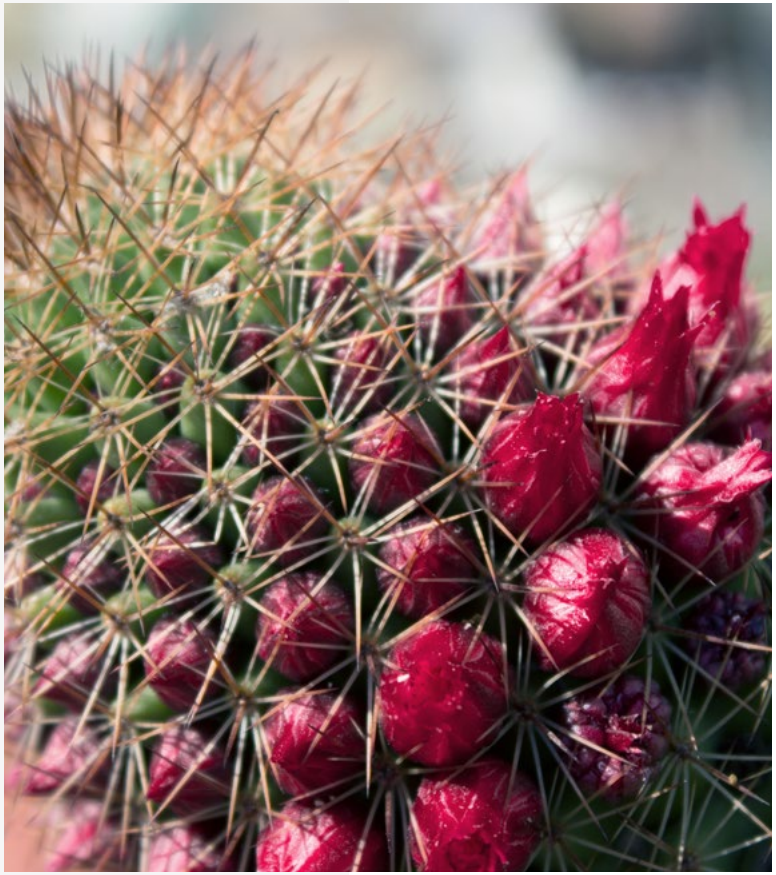
- Facilitate regular reassessment of all models in use and their generated outputs to refine and modify frameworks. The CTO at a multinational clothing company comments: *"Organizations must ensure that their models adhere to privacy principles and regulatory requirements. Another critical issue is ethical considerations and bias. Biases present in training data can lead to discriminatory outcomes. Additionally, AI models may produce artifacts that need to be understood and managed. Therefore, ensuring the quality and diversity of outputs related to gender and other factors is essential for the practical application of generative AI."*

- Be transparent about the privacy policy and give the customer easy access to resources explaining the underlying logic of AI algorithms, and offering clarity on your methodology to identify, eliminate, and prevent bias. Consider establishing specialized roles within the security team to oversee these efforts. *"Certain organizations are appointing a Chief trust officer within their security functions, which exemplifies the convergence where evolving regulations are blurring traditional roles and responsibilities. Now, cybersecurity frameworks are interlinking with considerations of AI trustworthiness, encompassing both risks and opportunities,"* said Alexandra Foster, former Managing Director at BT.

- CDOs should work closely with the CISOs and CIOs in ensuring the data integrity and data quality. By designating roles such as these, organizations can effectively leverage technology and data in fighting cyberattacks.
- Review your vendors' policies regarding data handling, storage, deletion timelines, and model training. Look for details on traceability, log history, anonymization, and other essential features.
- Solicit feedback from diverse stakeholders, including technology experts, business professionals and users, to evaluate the potential impacts and implications of AI applications. There is a clear need for increased collaboration between government and the private sector to manage complex technology platforms.
- Comply with security and notification requirements under latest regulations such as European Union's AI Act, Network and Information Systems (NIS) directive, and GDPR.

“Certain organizations are appointing a Chief trust officer within their security functions, which exemplifies the convergence where evolving regulations are blurring traditional roles and responsibilities. Now, cybersecurity frameworks are interlinking with considerations of AI trustworthiness, encompassing both risks and opportunities”

Alexandra Foster
Former Managing Director at BT and now an independent consultant



Model selection and training

To reduce the carbon footprint of LLMs, organizations should limit training. Using reputed and reliable LLMs specifically on raw cybersecurity logs minimizes false positives, reducing unnecessary alerts and clarifying genuine threats. They provide robust attack simulations and explore what-if scenarios, which are crucial to testing existing alerts and defenses.[56] To gain more flexibility with LLMs, organizations can explore open-source models such as Meta's Llama and fine-tune them to meet their specific needs.

Instead of solely focusing on LLMs, organizations can also consider using small language models (SLMs) that have fewer parameters, require less data and training time, are targeted for specific use cases, and have smaller carbon footprints. SLMs also have smaller attack surfaces, making them less susceptible to adversarial attacks.[57] The Brazilian payment company's CISO comments: *"One of the main challenges we face with AI in security is the high rate of false positives and negatives. This makes it difficult to trust AI in production. We need to ensure accuracy before deploying AI solutions on a large scale."* In the future, organizations will likely use a combination of LLMs and SLMs to meet their cybersecurity needs. Therefore, they should also focus on developing a custom Gen AI deployment pipeline to manage these models effectively.

Organizations can also explore custom GPT models, which are similar to tailored cyber tools. These models learn the organization's specific language, adapt to its nuances, maintain constant vigilance to detect anomalies, and learn continuously. Organizations must train such models on relevant data such as logs, incident reports, and threat intelligence. Organizations should synchronize them with the existing tech stack, enabling seamless communication with firewalls, intrusion detection systems and other security protocols. It should be noted that these custom GPTs are also susceptible to vulnerabilities. Setting up guardrails against misuse and, further, segregation of development and production environments will help in mitigating risks. Maintaining a clear boundary between testing and production systems and closely monitoring and validating model performance and integrity is essential.

When training models, organizations do need to note that models trained on sensitive data should also be considered as sensitive. Hence, it is also crucial to safeguard these models with the same level of security and confidentiality as that of the data they are trained on, to prevent unauthorized access. To safeguard these models and the model weights from unauthorized access and theft, organizations should:

- Establish a security plan that includes centralizing weights on a limited number of access-controlled systems and restricting access.
- Ensure interfaces for accessing the model are hardened against data-exfiltration attempts.
- Integrate confidential computing techniques to secure model weights during processing and minimize attack surfaces.[58]

Further, organizations can follow the federated learning (FL) approach which is a decentralized ML approach where training occurs across multiple devices, sending only model updates instead of raw data. This method enhances privacy by keeping personal information on local devices, improving models collaboratively without centralizing sensitive data. This approach, therefore, supports threat detection, anomaly identification, malware detection, and predictive analysis without sacrificing confidentiality of data.[59]

Additionally, as the softwarization of chips gain prominence, organizations' attack surface increases. This expanded attack surface can lead to serious security issues, such as unauthorized access to sensitive data, manipulation of chip functions, or even full system control by attackers. Therefore, introducing robust security measures at the chip level is equally crucial.

Formulate plans for integration and monitoring

Integrate with existing SOC solutions

Gen AI revamps SecOps by enhancing the current SOC capabilities, aiding in automation, data interpretation, suggesting best practices. As threats grow more sophisticated, organizations must pivot towards AI-driven solutions for enhanced detection and faster response.

Organizations should invest in AI-based solutions that can autonomously identify threats and block them. By crafting specific response playbooks, and optimizing workflows, AI also empowers security teams to efficiently prioritize, detect, and remediate issues.

Today, organizations could leverage a number of solutions available in the market to detect threats such as deepfakes and avoid bias in AI system outputs. Similarly, setting up guardrails against common attacks will help in ensuring that the AI systems do not turn rogue. It is essential to dismantle silos and foster cross-platform collaboration and promote a cohesive and unified security strategy. This integration enables comprehensive monitoring and protection of all facets of an organization's digital infrastructure.[60]

Furthermore, to safeguard against prompt injection risks and the autonomous nature of AI systems, it's crucial to implement an additional security layer that continuously monitors and intercepts potentially rogue commands. This secondary system should be designed to wrap around the primary AI system, acting as a gatekeeper that can scrutinize and filter commands before they reach the core AI functions. Adopting a zero-trust approach –where every input is treated as potentially compromised and scrutinized – can help mitigate these risks and enhance overall security.

Deploy AI agents

Organizations should strategically integrate AI agents into their cybersecurity operations. They are designed to function independently, plan, reflect, pursue higher-level goals, and execute complex workflows with minimal or limited direct human oversight.[61] In cybersecurity, such agents operate autonomously and monitor network traffic, detect anomalies, respond to threats in real time, and actively search for threats without human intervention. Additionally, AI agents simulate attacks, identify vulnerabilities, and develop defense strategies. This iterative process involves a collaborative ongoing effort among these agents.

As well as ecosystemic collaboration, organizations must also train employees to collaborate with these advanced systems. As the volume of the attacks increases, AI agents that operate within certain thresholds are critical in defending an organization's operations. However, these agents do require safeguards. Our latest research on Gen AI shows that 57% acknowledge the need for robust control mechanisms before integrating AI agents into their operations, and 73% insist that humans must verify and, if needed, intervene in AI decisions.[62] A careful balance is required between utilizing such autonomous agents and maintaining oversight due to the risks they present.

Continuous monitoring

Organizations must maintain continuous monitoring and updates of AI and Gen AI systems to defend against evolving threats. They should:

- Monitor, measure, audit, and log metrics to ensure the responsible, ethical, and secure deployment of AI models. Implementing scoring mechanisms provides real-time insights into the risk level associated with each input and output, helping users make informed decisions and maintain robust oversight. Monitoring for model drift over time and recalibrating the models will ensure the output is reliable.
- Focus testing on AI-specific risks, including jailbreaks, prompt injection, and issues related to coherence, readability, and toxicity of the generated content. Additionally, assess for biases and conduct red teaming exercises to strengthen the model's security posture.
- Invest in real-time behavioral pattern matching. As social engineering attacks increase, organizations must be able to detect and prevent any behavior that is out of the norm for an employee.
- Invest in existing solutions and tools to detect Gen AI signatures and patterns such as deepfakes, and proactively prevent potential attacks and mitigate threats effectively.

Create awareness and training programs

In our research, 58% of organizations mentioned a shortage of talented cybersecurity professionals. Additionally, 63% acknowledge the difficulty in integrating Gen AI into their existing security solutions due to talent limitations. Consequently, over half (51%) of organizations today are investing in comprehensive AI cybersecurity training programs.

These programs foster a deeper understanding of AI capabilities, limitations, and ethical considerations, ensuring responsible usage. Emphasizing the significance of diverse training programs, Alexandra Foster, former Managing Director at BT, states: *"In many organizations today, there is a strong emphasis on basic security awareness and training, particularly in areas like phishing. I believe there's great potential to expand these efforts to include comprehensive programs on social engineering and malware. This could involve incorporating simulations and leveraging gamification for effective training. Moreover, these initiatives should extend beyond just the cybersecurity team to encompass all departments, fostering a security-first culture that ensures everyone understands and values their role in maintaining security."*

58%

of organizations mentioned a shortage of talented cybersecurity professionals.

63%

of organizations acknowledge the difficulty in integrating Gen AI into their existing security solutions due to talent limitations.

51%

of organizations today are investing in comprehensive AI cybersecurity training programs.

Enhanced awareness and training lead to improved threat detection, response strategies, and overall cybersecurity posture. Frédéric Pégaz-Fiornet from Siemens Healthineers says: *"We have hackers within our company who continuously test our systems. We conduct internal testing and also engage the local Computer Emergency Response Team with hacking expertise to ensure our system hardening."*

- User awareness and education in AI cybersecurity ensure that individuals recognize potential threats and understand how to respond. This reduces human error, strengthens overall security, and promotes a culture of vigilance, enhancing the effectiveness of AI-driven defenses.
- Upskilling programs in AI and Gen AI for cybersecurity bridge the gap between cybersecurity and AI. Highlighting the importance of Gen AI training, a CISO from an automotive company adds: *"Before gaining access to Gen AI, employees must complete mandatory training to understand how to responsibly handle and utilize the platform. This ensures that sensitive information isn't inadvertently shared. It's crucial to educate employees on what type of information is appropriate for input into the system or platform before they begin using it."* Further, highlighting the importance of educating employees on Gen AI, Frederic Jesupret, Group Information Security Officer at Allianz Partners, states: *"It's essential to have checks in place to ensure that generative AI doesn't enter a negative or erroneous loop, as has been observed in some cases. This underscores the need to elevate the capabilities of my employees. While generative AI can handle a greater volume of events than manual processes, human oversight remains critical to verify conclusions periodically and ensure ethical considerations are upheld."* Over half of the organizations (56%) in our research believe that Gen AI will significantly redefine the roles and responsibilities of cybersecurity professionals within the next 2–3 years.

- Awareness programs should extend beyond end-users and security teams to include data scientists and engineers as well, who play a crucial role in the security of the models they develop and customize. Additionally, it's important to evaluate the practices of all relevant personas—such as Gen AI users, developers, data scientists, and cyber and infrastructure teams—to ensure they align with security policies and best practices.

56%

of organizations believe that Gen AI will significantly redefine the roles and responsibilities of cybersecurity professionals within the next 2–3 years.

“We're establishing an Innovation team to engage all departments effectively. Our focus spans beyond security leads, encompassing how the company navigates experiments, tests, and simulations. We're crafting a process to prioritize impactful actions and demonstrate their value. There are numerous market alternatives, not all are tailored to our specific needs. Our approach involves starting small, testing for value generation, and scaling promising initiatives swiftly. Governance, planning, and aligning expectations are crucial discussions, particularly when integrating AI to benefit our business comprehensively.”

Hélio Cordeiro Mariano
Chief Information Officer at
Cooperative Central Ailos

Safeguard business processes and cultivate a culture of risk awareness

With the rise of social engineering attacks, it is essential for organizations to embed security awareness into the organizational mindset. Employees must be trained to recognize and report potential threats swiftly. By cultivating a culture of risk awareness, organizations can ensure that employees remain vigilant and proactive in safeguarding against security breaches. A heightened sense of awareness and critical thinking skills (such as analyzing the context and corroborating from trusted sources) from employees can help organizations counterbalance the threats posed by malicious actors.

Additionally, cyberattacks often highlight the issues in business processes. Organizational hierarchies rely on trust in the individuals, with instructions often conveyed via email or workflow tools. With the increased use of digital communication and the ability of threat actors to intercept these communications, it's crucial to verify the authenticity of these interactions. AI can address this by incorporating real-time risk assessments into business processes. For instance, when executing high-risk transactions—like transferring money—AI can score the activity, assess risk, and enable informed decision-making. Leveraging AI to evaluate risky transactions ensures greater security and protects people and processes besides assets. Ultimately, human oversight over autonomous AI systems and a clear demarcation of duties can help ensure efficiencies while remaining secure. In the trifecta of people, process, and technology, it is essential that all these three elements seamlessly integrate to create a resilient and robust defense.

“It's essential to have checks in place to ensure that generative AI doesn't enter a negative or erroneous loop, as has been observed in some cases. This underscores the need to elevate the capabilities of my employees. While generative AI can handle a greater volume of events than manual processes, human oversight remains critical to verify conclusions periodically and ensure ethical considerations are upheld."

Frederic Jesupret
Group Information Security Officer
at Allianz Partners

Conclusion

AI and Gen AI are increasingly being exploited by malicious actors, who are leveraging these technologies to enhance their attacks. The threat landscape is continuing to become more sophisticated, with threat actors leveraging AI and Gen AI to run attacks at scale, posing significant risks to cybersecurity. With the volume of threats increasing more than ever, organizations must turn to these same technologies to mount a strong defense against cyberattacks.

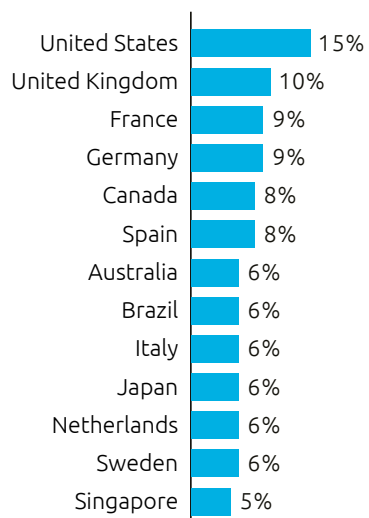
Organizations will need an innovative edge in the fight to remain secure against malicious actors. They are increasingly relying on AI for fast and accurate detection and reporting of real-time threats and, nuanced anomalies, enabling proactive defense strategies. While the adoption of Gen AI exposes organizations to some new threats – both from outside actors and from employees, it also provides opportunities to enhance security. At the same time, they must also understand the risks that are associated with Gen AI adoption and take necessary mitigative actions.

To harness AI and Gen AI effectively, organizations should embrace a culture of continuous re-assessment of the security landscape, building the infrastructure and establishing adequate framework and guidelines, as well as establishing robust employee training and awareness programs. They should develop new monitoring and control mechanisms and integrate them into their existing response processes. This will allow organizations to tap into the full potential of these technologies, creating a resilient cybersecurity posture that will be pivotal for safeguarding their most valuable assets and nurturing trust along the value chain.

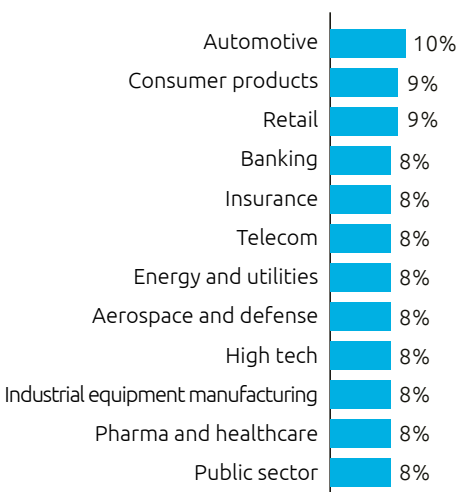
Research methodology

We conducted a targeted survey of 1,000 organizations that have either considered AI for cybersecurity or are already using it, across 12 sectors and 13 countries in Asia–Pacific, Europe, and North America. They have annual revenues of $1 billion and over. We carried out the global survey in May 2024. We provide the distribution of these respondents and their organizations below.

Organizations by headquarter location

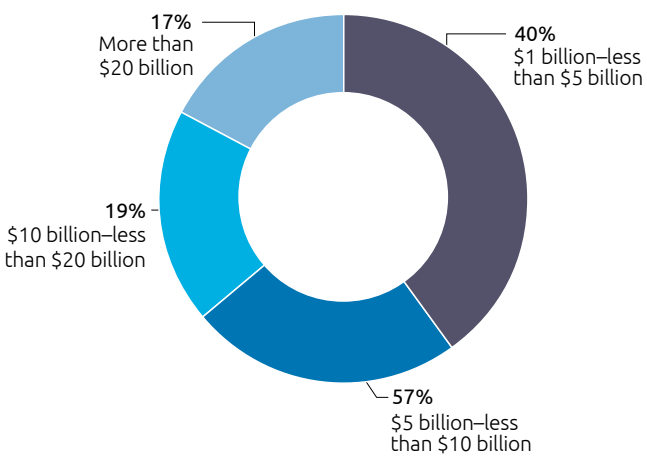


Organizations by sector

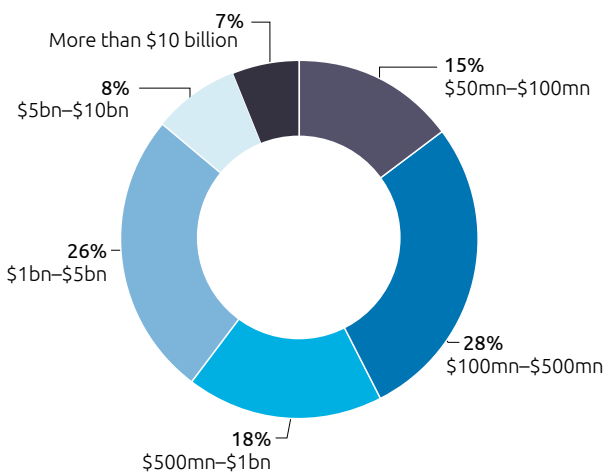


Source: Capgemini Research Institute, AI and Gen AI in cybersecurity survey, May 2024, N=1,000 organizations.

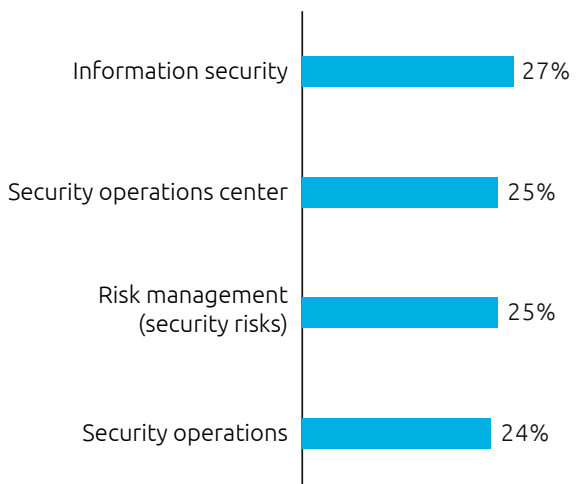
Organizations by annual revenue



Public sector organizations by annual budgets



Respondents by function



Source: Capgemini Research Institute, AI and Gen AI in cybersecurity survey, May 2024, N=1,000 organizations.

To supplement the survey findings, we also conducted in-depth discussions with 18 executives from organizations using AI and/or Gen AI in their cybersecurity defenses.

The study findings reflect the views of the respondents to our online questionnaire for this research and are intended to provide directional guidance. Please contact one of the Capgemini experts listed at the end of the report to discuss specific implications.

References

1. Coursera, accessed in August 2024.
2. Ibid.
3. Capgemini Research Institute, “Harnessing the value of generative AI: Top use cases across industries”, July 2023.
4. Forbes, “Gen AI and its malicious impact on the cyber-physical threat landscape,” April 2024.
5. bid.
6. VentureBeat, "Gen AI is the power surge cybersecurity vendors need to reduce the risks of losing the AI war," December 2023.
7. Check Point, “Check Point research reports a 38% increase in 2022 global cyberattacks,” January 2023.
8. Verdict, “AT&T in the crosshairs after a massive breach of customer data,” July 22, 2024.
9. Electric, “High-profile company data breaches,” June 2024.
10. Cybernews, “Attackers penetrate Walmart’s Spark driver portal,” February 2024.
11. The Straits Times, “Incidents of data leaks in S’pore public sector up 10%, with 201 cases recorded in 2023,” July 2024.
12. CS Hub, “IOTW: Data breach exposes sensitive information of Canadian Government employees,” November 2023.
13. ABC News, “Australian government investigating 'large-scale ransomware' data breach of script provider MediSecure,” May 2024.
14. Ekran, “7 examples of real-life data breaches caused by insider threats,” February 2024.
15. The Record, "MGM Resorts says cyberattack cost $100 million," October 2023.
16. Security Week, "Johnson Controls ransomware attack: data theft confirmed, cost exceeds $27 million," February 2024.
17. Machine-speed attacks are cyberattacks that occur at a pace that surpasses human capabilities, often leveraging automation, AI and ML to carry out the attack. The goal is to overwhelm defenses and exploit systems before human operators have a chance to react.
18. Microsoft, “Staying ahead of threat actors in the age of AI,“ February 14, 2024.
19. National Cyber Security Centre, “The near-term impact of AI on the cyber threat,” January 24, 2024.
20. Fortune, “Ferrari exec foils deepfake attempt by asking the scammer a question only CEO Benedetto Vigna could answer,” July 2024.
21. SCMP, “Hong Kong employee tricked into paying out HK$4 million after video call with deepfake ‘CFO’ of UK multinational firm,” May 2024.
22. Cybersecurity Asean, “AI Worms are crawling up as new AI parasites invade your devices,” May 2024.
23. Capgemini Research Institute, “Harnessing the value of generative AI 2nd edition: Top use cases across sectors,” July 2024.
24. Microsoft, “2024 Work Trend Index Annual Report,” May 8, 2024.
25. Capgemini Research Institute, “Harnessing the value of generative AI 2nd edition: Top use cases across sectors,” July 2024.
26. Capgemini Research Institute, “Turbocharging software with gen AI,” July 2024.
27. Forbes, "Samsung bans ChatGPT among employees after sensitive code leak," May 2023.
28. The Guardian, “Air Canada ordered to pay customer who was misled by airline’s chatbot,” February 2024.
29. Negri-Ribalta C, Geraud-Stewart R, Sergeeva A, Lenzini G. A systematic literature review on the impact of AI models on the security of code generation. Front Big Data. 2024 May 13;7:1386720. doi: 10.3389/fdata.2024.1386720. PMID: 38803522; PMCID: PMC11128619.

30. The Wall Street Journal, “Apple restricts employee use of ChatGPT, joining other companies wary of leaks,” May 2023.
31. Times of India, “Amazon has a ‘warning’ for employees using AI at work,” February 2024.
32. Dark Reading, “Hugging Face AI platform riddled with 100 malicious code-execution models,” February 2024.
33. Capgemini Research Institute, “Harnessing the value of generative AI 2nd edition: Top use cases across sectors,” July 2024.
34. IANS research security budget benchmark report, data as of October 3, 2023.
35. Google Cloud, “Advancing the art of AI-driven security with Google Cloud,” May 6, 2024.
36. Mastercard, "Mastercard leverages its AI capabilities to fight real-time payment scams," July 2023.
37. Emerj, "Artificial Intelligence at MetLife," May 2024.
38. Harvard Business Review, “What CEOs Need to Know About the Costs of Adopting GenAI,” November 15, 2023.
39. American Banker, "JPMorgan Chase using advanced AI to detect fraud" July 2023.
40. SOC Radar, "Threat intelligence can be life-saving in identifying and responding to botnet attacks," June 2023.
41. Relevant Software, "AI-driven threat intelligence systems represent a significant advancement in the continuous effort to protect data infrastructures," April 2024.
42. Relevant Software, "AI cybersecurity software can help with fraud detection," April 2024.
43. Relevant Software, "AI cybersecurity."
44. UCtoday, “Cisco unveils ‘first-of-its-kind’ AI-powered identity intelligence,” April 2024.
45. TechMagic, “AI in cybersecurity: Exploring the top 6 use cases,” April 2024.
46. The Globe And Mail, "How AI is revolutionizing the physical security industry," June 2024.
47. TechMagic, “AI in cybersecurity: Exploring the top 6 use cases,” April 2024.
48. BBVA, “BBVA signs an agreement with Telefónica Tech to boost cybersecurity on a global scale,” July 2024.
49. Forbes, "How AI Is disrupting the business of physical security," September 2023.
50. Google Cloud, “Supercharging security with generative AI,” April 2024.
51. Broadcom, “Broadcom partners with Google cloud to strengthen gen AI-powered cybersecurity,” September 2023.
52. Mastercard, “Mastercard accelerates card fraud detection with generative AI technology,” May 2024.
53. Google, “Real-world gen AI use cases from the world's leading organizations,” April 2024.
54. Opengovasia, “Exclusive! Generative AI unleashed: Bridging hardware and cybersecurity,” July 2024.
55. Ibid.
56. Nvidia, “Building cyber language models to unlock new cybersecurity capabilities,” July 2024.
57. Synergy-technical, “The significance of small language models for the future of AI & computing,” February 2024.
58. RAND, “Securing AI model weights,” accessed in July 2024.
59. Tripwire, “Federated Learning for cybersecurity: Collaborative intelligence for threat detection,” March 2024.
60. Sumologic, “How AI will impact cybersecurity: The beginning of fifth-gen SIEM,” April 2024.
61. Capgemini Research Institute, “Harnessing the value of generative AI 2nd edition: Top use cases across sectors,” July 2024.
62. Ibid.

Authors

Meet the experts

Karine Brune
CEO, Cloud and Infrastructure Services and member of Group Executive Committee, Capgemini
karine.brunet@capgemini.com

Karine was previously Capgemini's COO of Cloud Infrastructure Services since 2019. Before Capgemini, Karine was Vodafone's Technology Services Director, and she also held senior roles at Steria, Alcatel, and NCR. Karine has a strong track record of managing large transformation and infrastructure businesses. She holds three master's degrees in marketing and European Management, Economic Sciences, and European Economics.

Marco Pereira
Executive Vice President, Global Head Cybersecurity, Cloud and Infrastructure Services, Capgemini
marco.pereira@capgemini.com

Marco previously served as Chief Product and Strategy Officer at Trustwave, and Global Head of Strategy, Operations, and Product Management for Cybersecurity at DXC. Earlier in his career, he also held senior roles at HP, Wipro, and NTT Data. Marco holds a bachelor's and a master's degree in Information Systems and Computer Engineering, an MBA, and several leading industry cybersecurity certifications, including CISSP, CCSP, CISM, CISA, and ISO 27001 Lead Auditor.

Marjorie Bordes
Group Chief Information Security Officer, Group Cybersecurity, Capgemini
marjorie.bordes@capgemini.com

Marjorie previously spent three successful years as a Director of Cyberdefense Operations for Group Cybersecurity, during which she demonstrated exceptional expertise and leadership and led a team of more than a hundred cyber experts across the globe. Her previous roles were in cyber crisis management at a French bank (Société Générale), and she has ten years of experience in crisis management at the French Ministry of Defense and the French Ministry of Interior, in particularly complex and critical environments. She holds a PhD in Political Science from Descartes University and a master's degree in enterprise management.

Geert van der Linden
Executive Vice President, Group Offer leader Cybersecurity, Capgemini
geert.vander.linden@capgemini.com

Geert has served as the CISO for Cloud Infrastructure Services since 2021. He joined Capgemini in 2008, initially managing the application outsourcing practice in the Netherlands, before becoming the CIO of the Infrastructure Strategic Business Unit (SBU) in 2012. Early in his career, Geert also worked with the Dutch government. He holds degrees in Informatics, Business Informatics, and Organization, and is a qualified Chartered Accountant and Chartered IT Auditor.

Authors

Meet the experts

Jerome Desbonnet
CTIO, Cybersecurity and Chief cybersecurity architect for Cloud and Infrastructure Services and Insights & Data, Capgemini
jerome.desbonnet@capgemini.com

Prior to his current role, Jerome served as the Head of Security Solutions and Operations at Euroclear from 2018 to 2021. He also held the position of Global Cybersecurity CTO at Capgemini and SOGETI. His background includes roles as a security CTO for various organizations, project lead, and engineer, specializing in consulting, security engineering, architecture, identity and access management (IAM), privileged access management (PAM), Security Operations Center (SOC) management, threat intelligence, and threat hunting.

Steve Jones
Executive Vice President, Data Driven Business and Generative AI, Capgemini
steve.g.jones@capgemini.com

Steve brings over 21 years of innovation experience. He was a pioneer in mobile applications within the Capgemini group and played a key role in advancing Capgemini's early initiatives in SaaS and Cloud. Since then, Steve has established the Master Data Management and Big Data businesses at Capgemini. Currently, he leads the Group's efforts in Trusted AI and Collaborative Data Ecosystems and is a prominent writer and conference presenter on these topics.

Robert Engels
Vice President and Head of Generative AI Lab, CTO,Insights & Data, Capgemini
robert.engels@capgemini.com

Robert (Dr. Bob) Engels is also serving as CTO, AI for the Insights & Data Business Line. He has a long track record in the fields of AI, cognitive psychology, and knowledge presentation. Before joining Capgemini, he worked for startups, angel investors, and the Oslo municipality, had his own startup, oversaw radio and television production infrastructures with AI, and built a digital (AI-based) experience center for popular music. Robert holds a master's degree in cognitive psychology and AI, and a PhD in AI and reasoning.

Dr Mark Roberts
Deputy Director of Capgemini AI Futures Lab, CTO - Applied Sciences, Capgemini Engineering
mark.roberts@capgemini.com

Mark holds a PhD in AI and has over 25 years of experience in both academia and industry. He has applied advanced data science, AI, and custom software development for leading R&D companies. With a diverse background in technical roles, consultancy, project management, and strategic technology leadership, he is currently responsible for identifying emerging technologies and building capabilities to drive innovation in R&D. He is also serving as a Deputy Director of Capgemini's Group AI Lab, a hub for next-generation AI research and thought leadership.

Authors

Meet the experts

Babu Mauze
Executive Vice President – Cloud Infrastructure Services, Financial Services Strategic Business Unit, Capgemini
babu.mauze@capgemini.com

Babu has over 25 years of consulting experience in strategy development, solution design, and systems integration, with an extensive background in leading global teams in Banking & Insurance to help clients enable their digital transformation agendas. He holds an MBA in Finance and Information Systems and a bachelor's degree in Computer Engineering.

Serge Dujardin
Vice President, Cybersecurity, Capgemini
serge.dujardin@capgemini.com

Serge is a seasoned industry veteran who has previously held various executive positions, including leading the GTM strategy and professional services practice for military and commercial-grade encryption technologies at Thales Information Systems Security. He served as Senior Vice President of Technology & Business Development in the communications and connectivity sector for Aerospace & Defense at Cobham plc and held various global executive positions at Alcatel. He holds a BSc degree in computer science with a specialization in systems analysis.

Leonardo Carissimi
Senior Director, Cybersecurity Practice, Capgemini
leonardo.carissimi@capgemini.com

Leonardo has over 25 years of experience in the cybersecurity field, holding various leadership roles in both local and multinational companies. Throughout his career, he has developed a unique blend of solid technical background, along with business, managerial, and leadership skills. Leonardo holds a bachelor's and master's degree in Computer Science, an MBA in Corporate Finance, and key industry certifications like CISSP and ISO 27001 Lead Auditor.

Authors

Meet the Capgemini Research Institute

Jerome Buvat
Head,
Capgemini Research Institute
jerome.buvat@capgemini.com

Jerome is the head of the Capgemini Research Institute. He works closely with industry leaders and academics to help organizations understand the business impact of emerging technologies.

Ramya Krishna Puttur
Associate Director,
Capgemini Research Institute
ramya.puttur@capgemini.com

Ramya has over twelve years of experience in consulting and digital transformation and is an Indian Institute of Technology alumnus with gold medals at both post-graduation and under-graduation levels. She has co-authored numerous publications for the Capgemini Research Institute on the impact of digital technologies across industries and developed proprietary tools for assessing clients' data and digital maturity.

Hiral Shah
Manager, Capgemini Research Institute
hiral.shah@capgemini.com

Hiral is a research manager at the Capgemini Research Institute, where she collaborates with industry and business leaders to harness the power of digital technologies. Her work focuses on driving organizational transformation and optimizing business operations through innovative, technology-driven solutions.

Siva Chidambaram
Manager, Capgemini Research Institute
siva.chidambaram-s@capgemini.com

Siva is a research manager at the Capgemini Research Institute, where he collaborates with industry leaders to drive modern tech-enabled solutions. His work specializes in strategic research, advisory, and improving business operations through intelligence that aids staying ahead in new age technology landscapes.

The authors would like to thank the following people for their contributions to the research:

Subrahmanyam KVJ
Senior Director, Capgemini Research Institute

Victoria Otter
Security Analyst, Cybersecurity, Capgemini

Ben Dickson
Enterprise Architect, Cybersecurity, Capgemini

Thierry Daumas
Executive Vice President, Apps Practice, Capgemini

Stephen Hilton
Executive Vice President, Cloud & Infrastructure Services, Capgemini

Jeanne Heure
Director Digital Trust & Security, Future of Technology, Capgemini

Michael Wasielewski
Head of Cloud Security and Next-Gen Secure Architectures, Cloud & Infrastructure Services, Capgemini

Christophe Menant
Cybersecurity Strategy & Governance, Capgemini

Eric Fradet
Vice President, Cloud & Infrastructure Services, Capgemini

Ayan Bhattacharya
Vice President, Cloud & Infrastructure Services, Capgemini

Vincent Fokke
Vice President, Global Head of Enterprise Architecture for Financial Services, Capgemini

Nicolas Cabridain
Vice President, Future of Technology, Capgemini

Didier Appell
Head of OT/IoT Cybersecurity, Capgemini

Swadesh Dash
Senior Director, Financial Services, Capgemini

Noel Bonnet
Principal Business Analyst, Apps Practice, Capgemini

Sree Vadakkepat
Principal, Corporate & Common Functions, Capgemini

Pierre-Adrien Hanania
Head of Business Development, Public Sector

Will Matthews
Marketing & Communication, Public sector, Capgemini

Luca Bordonaro
Consultant, Future of Technology, Capgemini

Philippine Carle
Consultant, Insights & Data, Capgemini

Andy Talbot
Vice President & Global Head Cyber Managed Services, Capgemini

Oliver Jones
Director Digital Trust & Security, Future of Technology, Capgemini

Camille Maindon
Consultant, Future of Technology, Capgemini

Giordano Orchi
Managing Consultant, Cloud & Infrastructure services, Capgemini

Sebastian Lagana
Senior Manager, Corporate & Common Functions, Capgemini

Damien Stulemeijer
Strategy and Transformation Consultancy, Capgemini

Debarati Ganguly
Director, Industry and Innovation, Capgemini

Nicolas Dignoire
Security Director, Cybersecurity, Capgemini

The authors would like to thank all industry executives who participated in this research. They would also like to thank Tricia Stinton, Vijayalakshmi K, Punam Chavan, Suparna Banerjee, Jaydeep Neogi, and Amitabha Dutta for their contributions to the research.

Why Partner with Capgemini?

At Capgemini, we understand that securing Generative AI isn't one-size-fits-all. Our Gen AI Security Suite offering covers the entire lifecycle—from design and deployment to continuous monitoring. We assess your security posture, offer best-practice guidance, and build secure, traceable solutions tailored to your business needs.

Unlock GenAI's potential with Capgemini's Gen AI Security Suite

With our expertise in both Generative AI and cybersecurity, you can harness the transformative power of Gen AI while minimizing risks. Whether your focus is innovation, optimization, personalization, or digital transformation, our solution integrates trust and security into every phase of AI development and deployment. Our approach includes:

- Ethical and secure AI development: Ensuring models, data, and outputs are secure and ethical.
- Protection of corporate assets and reputation: Mitigating risks associated with the deployment and use of GenAI.
- Enhanced decision-making: Leveraging AI for faster, informed security choices.
- Proactive defense: Guarding against AI-driven threats.

With Capgemini, you can:

- Adopt GenAI confidently: Securely use AI for customer engagement and process optimization.
- Ensure data security: Protect data across the AI lifecycle.
- Meet compliance: Adhere to regulations and manage risks.
- Maximize value: Optimize your GenAI investments securely.

Leverage our unparalleled expertise in AI development, seamless application delivery, and advanced threat mitigation. Our integrated security framework guarantees authentic GenAI security, turning uncertainty into unwavering confidence.

For more information:

Visit: Cybersecurity Services | Cybersecurity Strategy & Transformation | Capgemini | Generative AI & LLM: Transforming Industries for Business Excellence| Capgemini

Follow us on LinkedIn: Capgemini Cybersecurity

For more information, please contact:

Global

Geert van der Linden
EVP & Cybersecurity Group offer Leader, Capgemini
geert.vander.linden@capgemini.com

Marco Pereira
EVP, Global Head Cybersecurity, Cloud and Infrastructure Services, Capgemini
marco.pereira@capgemini.com

Jerome Desbonnet
CTIO, Cybersecurity, Cloud and Infrastructure Services, Capgemini
jerome.desbonnet@capgemini.com

Michael Wasielewski
Head of Cloud Security and Next-Gen Secure Architectures, Cloud and Infrastructure Services, Capgemini
michael.wasielewski@capgemini.com

Regional

France
Abdembi Miraoui
Co-Head of Cloud, Endpoint & Infrastructure Security
abdembi.miraoui@capgemini.com

Brazil
Leonardo Silva Carissimi
Director, Cybersecurity operations
leonardo.carissimi@capgemini.com

APAC
Keith Betts
Senior Director, Cybersecurity
keith.betts@capgemini.com

Germany
Adivitya Mahajan
Cloud and platforms security specialist
adivitya.mahajan@capgemini.com

Spain
Manuel Felipe Castillo Guerrero
manuel.castillo@capgemini.com

North America
Cedric Thevenet
Head of Cyber Sales & Solutioning
cedric.thevenet@capgemini.com

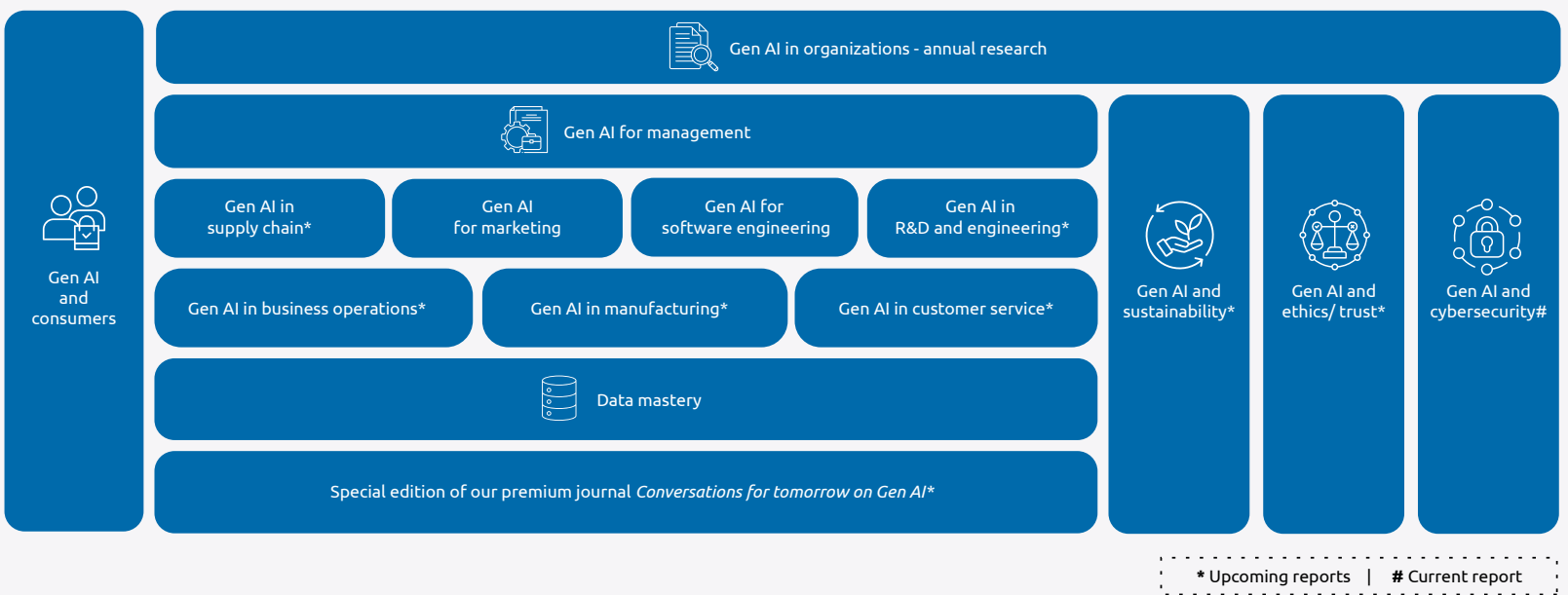
Netherlands
Ruben Tienhooven
ruben.tienhooven@capgemini.com

UK
Wayne Reid
Solution Architect
wayne.reid@capgemini.com

Michael Linster
Director Architect, Delivery
michael.linster@capgemini.com

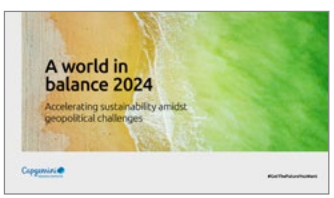
Our generative AI research series

This report is a part of Capgemini Research Institute's series on generative AI.

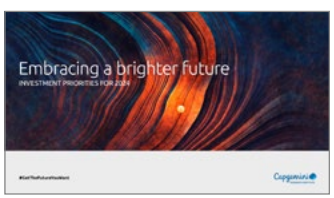


To find out more, please go to https://www.capgemini.com/insights/research-institute/

More Capgemini Research Institute publications

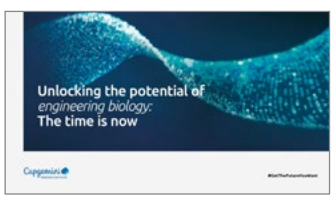


A World in balance 2024: Accelerating sustainability amidst political challenges



Embracing a brighter future: Investment priorities for 2024

The Eco-Digital Era™: The dual transition to a sustainable and digital economy



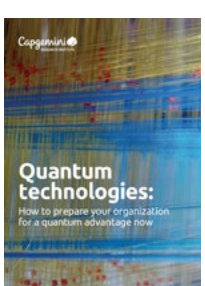
Unlocking the potential of engineering biology: The time is now

Connected products: Enhancing consumers' lives with technology

The resurgence of manufacturing: Reindustrialization strategies in Europe and the US

Data for net zero: Why data is bridging the gap between net zero ambition and action

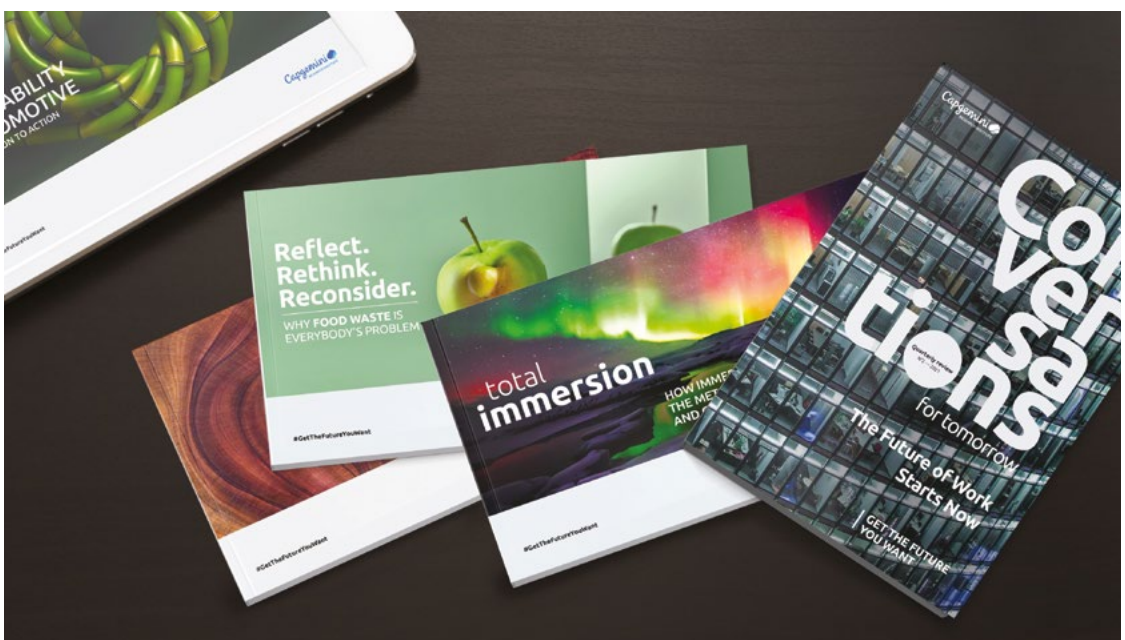
Digital Twins: Adding intelligence to the real world



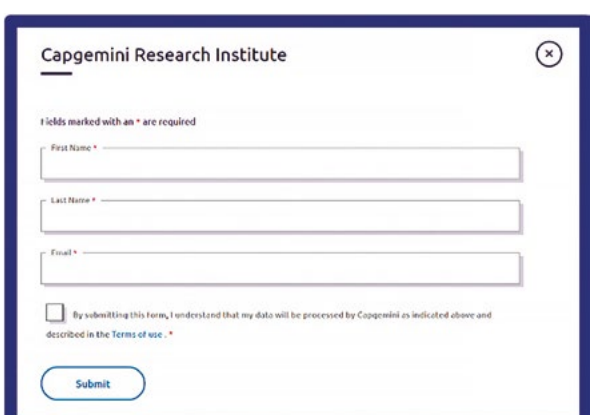
Quantum technologies: How to prepare your organization for a quantum advantage no

Reinventing cybersecurity with artificial intelligence

Subscribe to latest research from the Capgemini Research Institute



Receive copies of our reports by scanning the QR code or visiting
https://www.capgemini.com/capgemini-research-institute-subscription/



Notes

About Capgemini

Capgemini is a global business and technology transformation partner, helping organizations to accelerate their dual transition to a digital and sustainable world, while creating tangible impact for enterprises and society. It is a responsible and diverse group of 340,000 team members in more than 50 countries. With its strong over 55-year heritage, Capgemini is trusted by its clients to unlock the value of technology to address the entire breadth of their business needs. It delivers end-to-end services and solutions leveraging strengths from strategy and design to engineering, all fueled by its market leading capabilities in AI, cloud and data, combined with its deep industry expertise and partner ecosystem. The Group reported 2023 global revenues of €22.5 billion.

Get the Future You Want | **www.capgemini.com**

Copyright © 2024 Capgemini. All rights reserved.