Deloitte.

Overcoming identity and access management risks associated with generative AI

MAKING AN IMPACT THAT MATTERS since 1845

Introduction

The rapid evolution of generative AI technologies presents both opportunities and challenges for Identity and Access Management (IAM) systems, including Privileged Access Management (PAM) and Identity Governance and Administration (IGA). This paper explores the associated risks and outlines strategies to enhance the current IAM systems and mitigate risks where applicable.

The audience for this discussion on the business considerations of adopting generative AI in Identity and Access Management (IAM) include:

C-suite executives

IT and security professionals

Compliance and risk management officers

Data privacy officers

C-suite executives who are responsible for strategic decision-making and overseeing the adoption of new technologies within the organization.

IT and security professionals who implement and manage IAM systems and are involved in evaluating new security technologies.

Compliance and risk management officers responsible for ensuring that the organization complies with regulatory requirements and manages risks associated with data protection and privacy.

Data privacy officers tasked with ensuring that the organization's use of data complies with privacy laws and ethical standards.

AI and machine learning specialists

Enterprise architects

Investors, board members, and other stakeholders

Technology consultants and advisors

AI and machine learning specialists such as data scientists, machine learning engineers, and AI researchers interested in the practical applications and implications of generative AI in security contexts.

Enterprise architects who design and manage the IT infrastructure of an organization, ensuring that new technologies integrate seamlessly with existing systems.

Investors, board members, and other stakeholders who need to understand the potential benefits and risks associated with investing in generative AI for IAM.

Technology consultants and advisors who provide guidance to organizations on the adoption of new technologies and strategies for improving security and efficiency.

Academic and research institutions

Policy makers and regulators

Academic and research institutions studying the intersection of AI, security, and privacy, as well as the implications of these technologies on business and society.

Policy makers and regulators oversee the compliance of businesses with laws and regulations related to data protection, privacy, and AI ethics.

Generative AI in IAM

Generative AI is poised to revolutionize IAM by introducing innovative solutions that enhance security, efficiency, and user experience. As enterprises navigate an increasingly complex digital landscape, the integration of generative AI into IAM systems offers significant competitive advantages.

Enhancing security protocols

Generative AI models, such as GPT-3 and Jasper, can create synthetic data and simulate potential breaches, providing invaluable insights into vulnerabilities within IAM frameworks. This capability enables enterprises to proactively fortify their security protocols, anticipating and mitigating threats before they materialize. By leveraging AI-driven simulations as a precursor to purple-teaming, organizations can identify weak points in their security infrastructure and develop robust defences, ensuring comprehensive protection against evolving cyber threats.

Streamlining user authentication

Generative AI can significantly enhance user authentication methods, moving beyond traditional passwords to more secure, efficient alternatives. AI-driven biometric authentication, including facial and voice recognition, provides a seamless and secure user experience. These advanced methods reduce the risk of unauthorized access and streamline the authentication process, minimizing friction for legitimate users while maintaining stringent security standards.

Improving access control and monitoring

Generative AI excels in analyzing vast amounts of data to identify patterns and anomalies. This capability is crucial for PAM and IGA, where precise access control and continuous monitoring are essential. AI can dynamically adjust access privileges based on real-time analysis of user behavior, ensuring that only authorized individuals have access to sensitive information. This adaptive approach enhances security while optimizing operational efficiency.

Supporting regulatory compliance

Compliance with regulatory requirements is a critical aspect of IAM. Generative AI can assist organizations in maintaining compliance by continuously monitoring and analyzing ever growing access logs, identifying potential violations, and ensuring adherence to industry standards and legal obligations. This automated oversight reduces the burden on compliance teams and minimizes the risk of costly regulatory breaches; alerting security teams of potential or actual threats the moment they occur.

Delivering business value

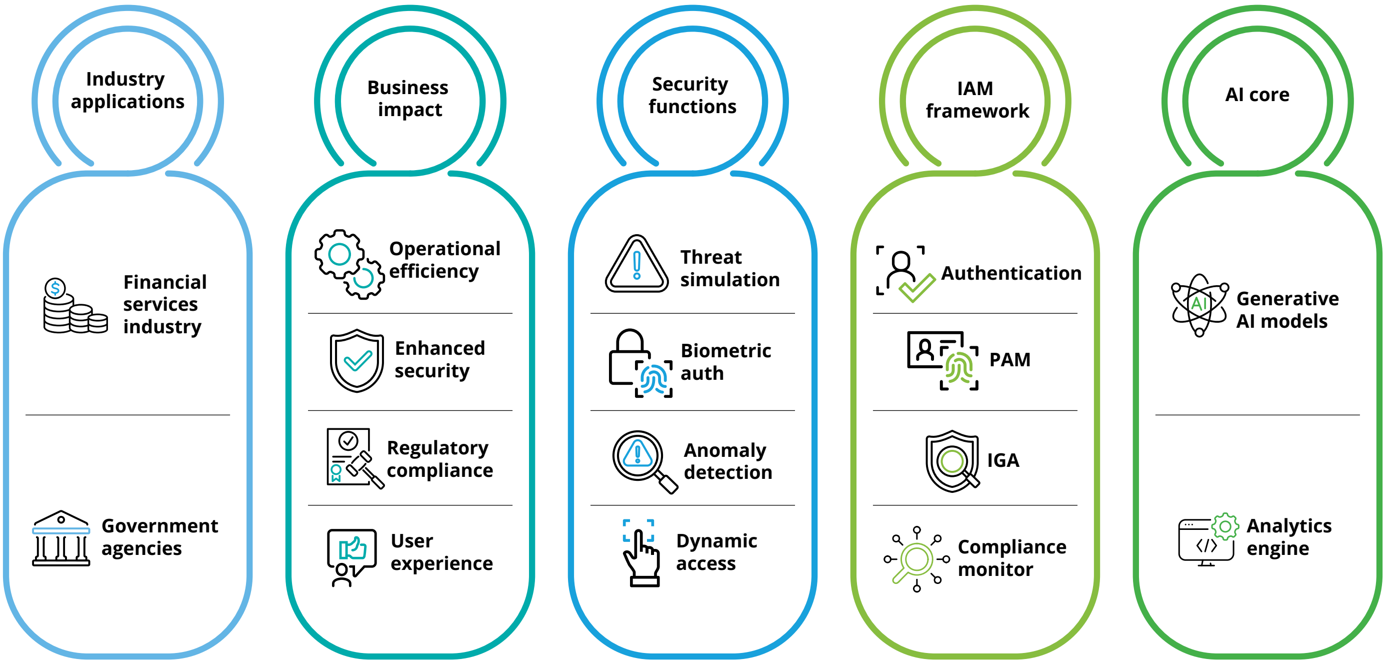
Integrating generative AI into IAM systems not only enhances security and compliance but also delivers tangible business value. By improving operational efficiency, reducing security incidents, and ensuring a seamless user experience, AI-driven IAM solutions can significantly impact the bottom line.

Real-world application examples

In the Financial Services Industry (FSI), generative AI can secure sensitive financial data, detect and prevent fraud, and streamline customer authentication processes.

For government agencies, AI-driven IAM solutions ensure the protection of critical infrastructure, secure citizen data, and facilitate compliance with stringent regulatory requirements. Real-world examples demonstrate how AI integration can enhance security, efficiency, and trust in these sectors.

By adopting generative AI in IAM, enterprises can transform their security posture, streamline operations, and deliver superior user experiences. This innovative approach not only mitigates current risks but also prepares organizations for future challenges, ensuring long-term resilience and success in a rapidly evolving digital landscape.



Risks of Generative AI

Data-driven risks

Data poisoning: Malicious actors might manipulate the training data for generative AI models used in IAM, leading to compromised authentication or authorization mechanisms. For example, inserting fake user profiles or access logs could trick the AI into granting access to unauthorized individuals.

Privacy breaches: Generative AI models trained on sensitive user data, such as biometric information or behavioral patterns, pose privacy risks. Accidental leakage or unauthorized access to this data could have serious consequences for individual privacy and security.

Deepfake attacks: Advanced generative AI techniques could be used to create convincing deepfakes of authorized users, enabling unauthorized access through biometric authentication or video verification.

Phishing attacks and social engineering: Generative AI can be used to create highly convincing phishing emails, messages, or other social engineering attempts based on the victim's behavior and social network. These attacks may be more difficult for traditional security measures to detect, as the content generated can closely mimic legitimate communications.

Checklist for data-driven risks mitigation

- Implement robust data validation and cleansing processes
- Use differential privacy techniques during model training
- Use biometric technologies that are deepfake resistant.
- Regularly update training datasets to remove outdated or compromised data
- Educate employees on recognizing and responding to phishing and social engineering attacks

Security loopholes

Adversarial attacks: Malicious actors might exploit vulnerabilities in generative AI models used for IAM, such as those designed to detect anomalies or fraudulent behaviour. This could allow them to bypass security measures and gain unauthorized access.

False positives and negatives: Generative AI models can make mistakes, leading to false positives (denying access to authorized users) or false negatives (granting access to unauthorized users). This can disrupt operations and compromise security.

Bias and discrimination: Generative AI models trained on biased data can perpetuate existing biases in identity and access decisions, unfairly affecting certain groups of individuals.

AI attacks: Malicious actors can use AI driven tools and scripts that can detect and exploit weaknesses or unpatched systems in the environment. As well as dormant accounts and over-privileged accounts.

Checklist for security loopholes risks mitigation

- Identify AI model vulnerabilities by conducting regular security audits
- Implement anomaly detection by monitoring for false positives/negatives
- Train AI on diverse datasets to mitigate bias and discrimination
- Regular patching and automated cleanup of identity data is crucial to minimize loophole

Ethical considerations

Transparency and accountability: it is crucial to understand how generative AI models make decisions, especially when they deny access or flag suspicious activity. Without transparency, users may feel unfairly targeted or discriminated against.

Human oversight and control: While generative AI can automate parts of IAM, human oversight and control remain essential. Decisions with critical consequences, like account termination or emergency access, should ultimately involve human judgment.

Content moderation challenges: In platforms where user-generated content is prevalent (e.g., forums, social media), generative AI can be used to create inappropriate or harmful content. IAM systems may need to adapt to new challenges in content moderation to maintain a secure and safe environment.

Checklist for ethical considerations risk mitigation

- Ensure transparency with users on the AI decision-making process
- Maintain detailed logs of AI-driven decisions
- Integrate human oversight to ensure accountability
- Establish protocols for human review of critical decisions (such as blocking of users or access).

Strategies for AI risk mitigation

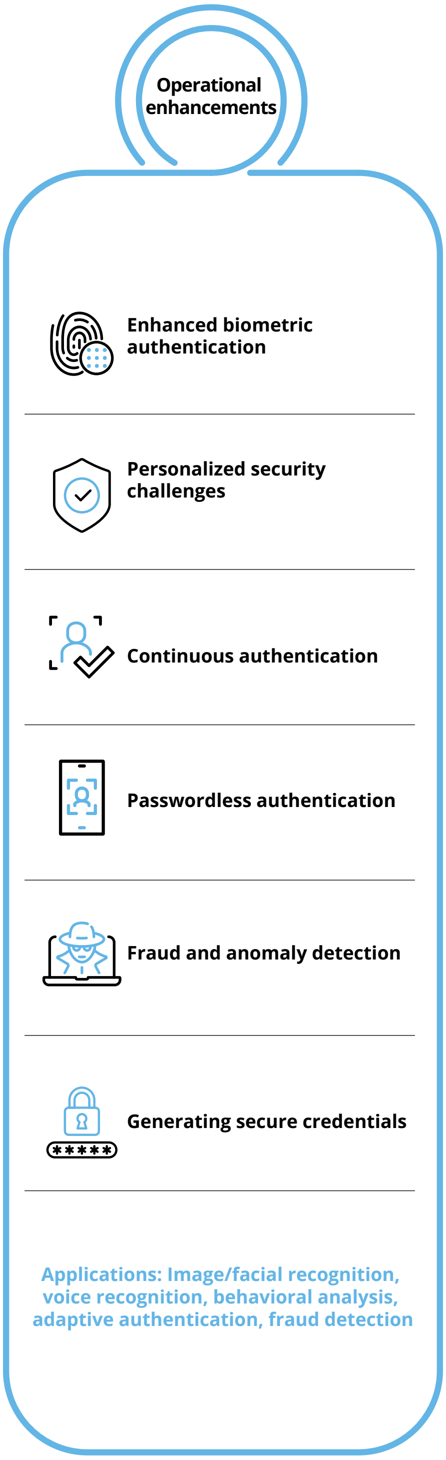
Regular updates to IAM policies: Enterprises need to continuously update their IAM policies and procedures to address emerging threats and vulnerabilities associated with the emerging AI technologies.

User education: Enterprises should promote user education and awareness campaigns to recognize AI driven attacks such as customised phishing attacks and deepfakes.

Advanced authentication: Enterprises should invest in advanced authentication methods to counteract the potential threats posed by generative AI. This may include incorporating multi-factor authentication, behavioral analytics, location-based policies, and other advanced security measures to strengthen the resiliency of their systems.

Enhancing IAM through Generative AI

While it is important to consider the risks above when planning your IAM strategy; it is just as important to leverage generative AI technologies to create a more secure IAM environment.



Operational enhancements

Enhanced biometric authentication

Personalized security challenges

Continuous authentication

Passwordless authentication

Fraud and anomaly detection

Generating secure credentials

Applications: Image/facial recognition, voice recognition, behavioral analysis, adaptive authentication, fraud detection

Operational enhancements

By using generative AI enabled products, clients can enhance the capability and strength of their environment by using advanced techniques that would have been cost prohibitive in the past due to the laborious process it would have had to go through without using AI.

Enhanced biometric authentication

Image and facial recognition: Generative AI models can be utilized to improve the accuracy and robustness of facial recognition systems. This technology can be integrated into authentication processes for secure access to physical locations or digital systems.

Voice and speech recognition: Generative AI can enhance voice and speech recognition systems by creating more accurate models that understand and verify the uniqueness of an individual's voice. This can contribute to the development of secure voice-based authentication methods.

Generating unique, unpredictable challenges: AI can generate random, one-time biometric challenges, such as asking users to blink a specific pattern or speak a unique phrase. This makes it difficult for attackers to predict or replicate these challenges using traditional methods.

Analyzing behavioral patterns: AI can analyze subtle behavioural patterns, such as typing rhythm, mouse movements, and eye gaze, to create a unique "behavioral fingerprint" for each user. This fingerprint can be used to continuously authenticate users in a transparent and non-intrusive way.

Detecting anomalies: AI can detect anomalies in biometric data, such as inconsistencies in facial features or voice patterns, which could indicate spoofing attempts, deep fakes, or other methods of attacks that rely on user input.

Personalized security challenges

Tailoring challenges to individual users: AI can tailor authentication challenges to individual users based on their preferences and cognitive abilities, making the authentication process more user-friendly and secure.

Generative AI models: can be used to create more effective and secure Captcha systems for human verification during the authentication process. These systems can adapt to evolving strategies employed by malicious actors attempting to bypass authentication.

Generative AI-powered chatbots: can be integrated into authentication processes to provide secure and user-friendly support. These chatbots can engage in natural language conversations to verify user identity or assist with account recovery.

While these applications showcase the potential of generative AI in enhancing authentication processes, it's essential to carefully consider privacy and ethical considerations. Striking a balance between security and user experience is crucial to ensure that authentication measures are effective without causing undue friction for legitimate users.

Continuous authentication

Monitoring user behavior: AI can monitor user behavior throughout their sessions to detect any unusual activity that might signal unauthorized access. This includes analyzing keystroke patterns, mouse movements, application usage, and geolocation data.

Adaptive authentication: AI can dynamically adjust the level of authentication required based on risk factors, such as the user's location, device, and recent activity. For example, it might prompt for additional verification if a user logs in from an unfamiliar device or location.

Passwordless authentication

Tailoring challenges to individual users: AI can tailor authentication challenges to individual users based on their preferences and cognitive abilities, making the authentication process more user-friendly and secure.

Generative AI models: can be used to create more effective and secure Captcha systems for human verification during the authentication process. These systems can adapt to evolving strategies employed by malicious actors attempting to bypass authentication.

Fraud and anomaly detection

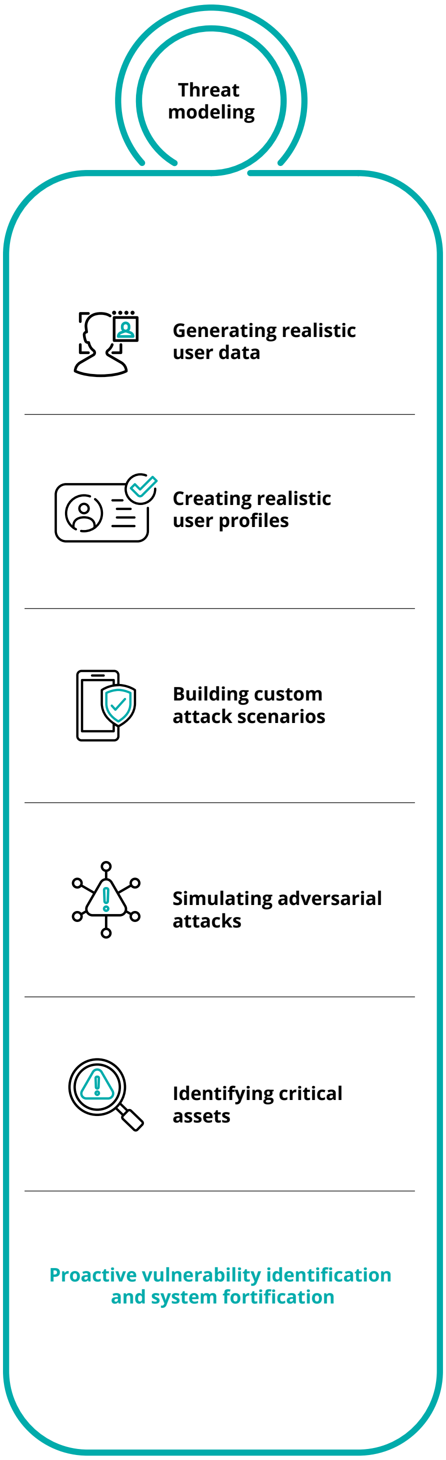
Identifying fraudulent patterns: AI can analyse large datasets of user behavior and authentication attempts to identify patterns that might indicate fraudulent activity, such as unusual login times or locations, or attempts to use multiple identities.

Proactive threat detection: AI can proactively detect and prevent potential threats, such as account takeover attempts or data breaches, by analyzing real-time data and identifying suspicious patterns.

Generating secure credentials

Creating strong, unique passwords: AI can create passwords that are both strong and easy to remember for users, reducing the likelihood of password reuse or compromise.

Generating decoy passwords: AI can create realistic-looking decoy passwords that can be used to trap attackers who have stolen a user's password database.



Threat modeling

Generating realistic user data

Creating realistic user profiles

Building custom attack scenarios

Simulating adversarial attacks

Identifying critical assets

Proactive vulnerability identification and system fortification

Threat modelling

Additionally, generative AI can be a powerful ally in simulating and predicting potential vulnerabilities associated with Identity and Access Management within an organization's security infrastructure. By mimicking potential threat scenarios, companies can proactively identify weaknesses and fortify their systems before actual breaches occur.

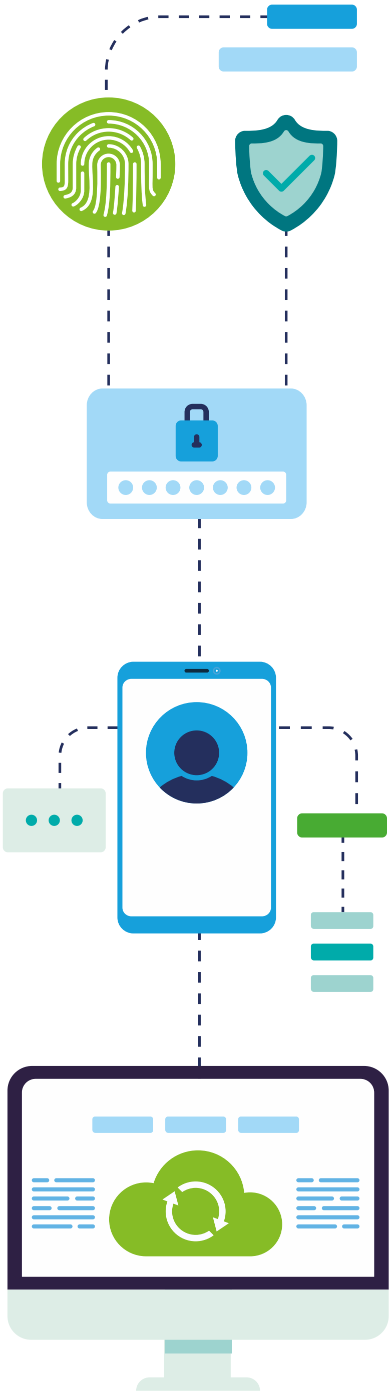
Generating realistic user data: AI can generate synthetic user data, including passwords, biometric information, and behavioral patterns, to train and test authentication systems without putting real user data at risk.

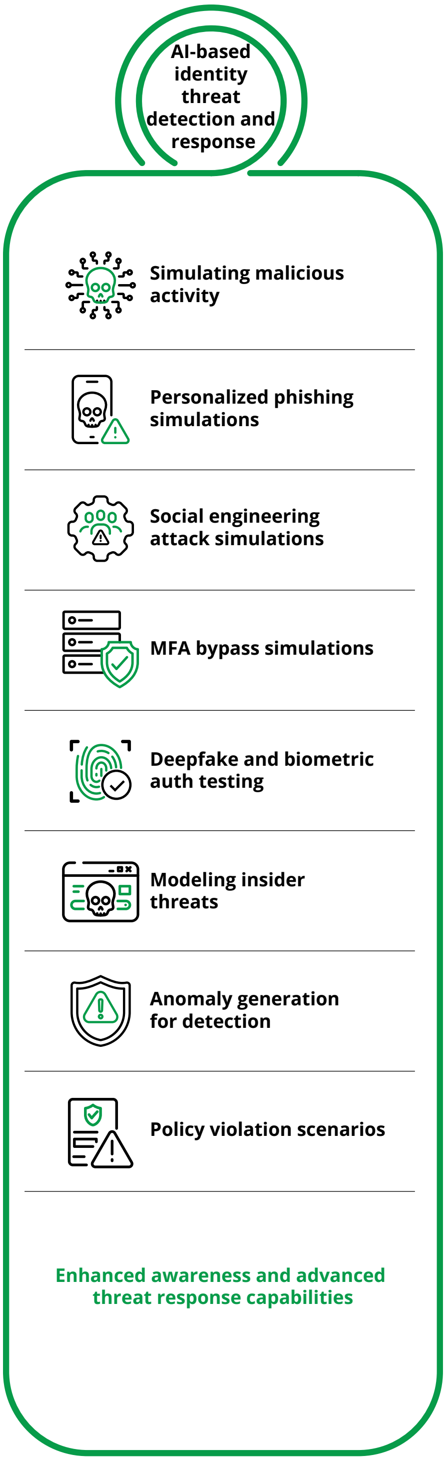
Creating realistic user profiles: AI can generate diverse user profiles with varying access levels, permissions, and behaviours, mimicking a real-world user base. This allows for testing a wide range of attack scenarios.

Building custom attack scenarios: By feeding the AI model specific information about the organization’s IAM systems and potential vulnerabilities, researchers can create tailored attack scenarios for targeted testing and mitigation strategies.

Simulating adversarial attacks: It can generate data that simulates various adversarial attacks, helping models become more robust against attempts to fool them.

Identifying critical assets: AI can analyze IAM systems and user access to identify critical assets and data that require extra protection. This helps prioritize security efforts and focus resources on the most sensitive areas.





AI-based identity threat detection and response

Simulating malicious activity

Personalized phishing simulations

Social engineering attack simulations

MFA bypass simulations

Deepfake and biometric auth testing

Modeling insider threats

Anomaly generation for detection

Policy violation scenarios

Enhanced awareness and advanced threat response capabilities

AI Based ITDR

Enhancing Identity Threat Detect & Response using generative AI will take an organisation's awareness and abilities to the next level.

Simulating malicious activity: AI can generate synthetic data representing various attack attempts, like phishing emails, malware payloads, or brute-force login attempts. This helps assess the effectiveness of existing security measures and identify potential weaknesses.

Personalised phishing simulations: AI-powered phishing simulations that are targeted and personalized for every user can be created to test the effectiveness of employee training programs and the resilience of IAM systems against evolving phishing techniques.

Simulating social engineering attacks: AI can generate realistic conversations mimicking social engineering tactics. This can help security teams train employees to identify and resist attacks as well as provide a more realistic experience to a real-world attack.

Multi-factor authentication bypasses: AI can simulate various bypass attempts for different MFA methods, like phone calls, SMS, or hardware tokens, revealing potential weaknesses in implementation or user behavior.

Deepfake and biometric authentication: Generative AI, particularly in the domain of deepfakes, can simulate attempts to bypass biometric authentication systems. By generating synthetic facial images or voice recordings, organizations can evaluate the susceptibility of their IAM infrastructure to these emerging threats.

Modelling insider threats: AI can create scenarios where authorized users misuse their access or attempt to escalate privileges, helping organizations develop effective mitigation strategies for insider threats.

Anomaly generation for detection testing: AI models can generate anomalous patterns of user behavior to test the effectiveness of anomaly detection mechanisms within IAM systems. This can include simulating unusual login times, locations, or access patterns to assess the IAM system's ability to detect and respond to abnormal activities.

Policy violation scenarios: AI models can generate scenarios that violate IAM policies, such as unauthorized access attempts, privilege escalation, or improper data sharing. This allows organizations to assess the IAM system's ability to enforce and detect policy violations.

Optimizing response plans: By simulating attacks and their potential consequences, AI can help organizations optimize their incident response plans and improve their overall readiness for security breaches.

Penetration testing and red teaming: AI can assist in automating penetration testing and red teaming activities. AI-powered tools can simulate the behaviour of malicious actors by attempting to exploit vulnerabilities in the IAM system, helping organizations identify and remediate potential weaknesses

> While it is important to consider the risks above when planning your IAM strategy; it is just as important to leverage generative AI technologies to create a more secure IAM environment.

Using AI to enhance AI

Training anomaly detection systems

Expanding diversity and representation

Balancing minority classes

Controlling data characteristics

Debiasing existing datasets

Mitigation: Data validation + differential privacy + deepfake detection + employee training

Using AI to enhance AI

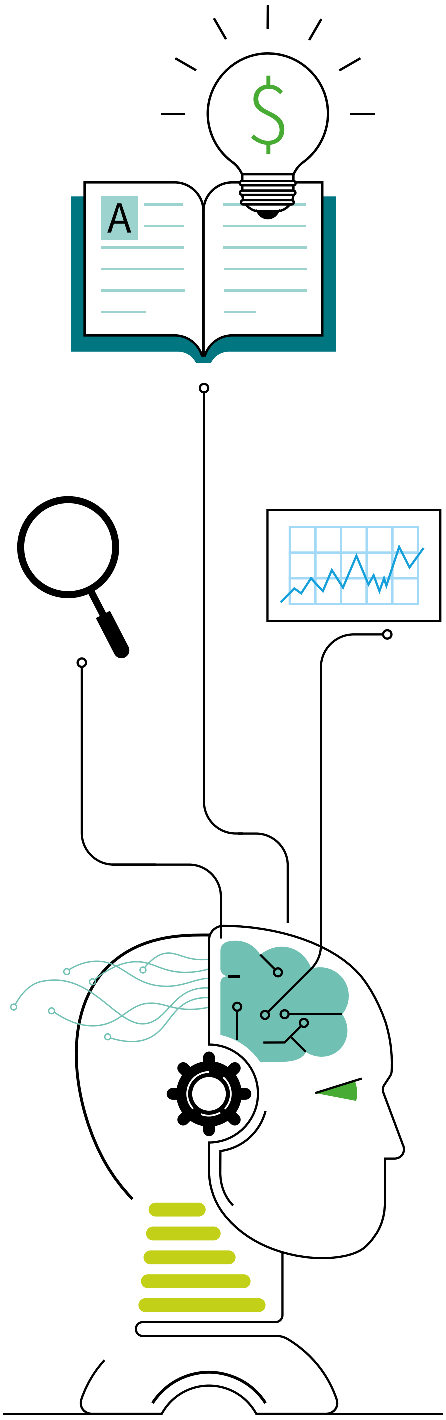
Training anomaly detection systems: AI can be used to generate massive datasets of anomalous user behavior and authentication attempts. This data can be used to train anomaly detection systems to identify real-world security threats in real-time. This ensures that the simulated scenarios remain relevant and effective in identifying potential vulnerabilities as threat landscapes evolve.

Expanding diversity and representation: Generative AI can create new, synthetic data points that fill in gaps in existing datasets, ensuring better representation of diverse identities and potential fraud scenarios. This includes creating additional samples with variations in lighting conditions, poses, facial expressions, and background scenarios. Augmenting datasets helps improve the model's ability to generalize across different conditions and avoid biases by ensuring inclusivity in the model's recognition capabilities.

Balancing minority classes: It can generate additional examples of minority classes (e.g., rare fraud cases) to address class imbalance and improve model performance.

Controlling data characteristics: Developers can control the specific characteristics of synthetic data to ensure it aligns with real-world patterns and addresses specific verification challenges.

Debiasing existing datasets: AI can identify and correct imbalances or unfair representations in existing datasets, promoting model fairness and inclusivity. It can also generate diverse synthetic data to train models that are less likely to perpetuate biases and discriminate against underrepresented groups.



Recommendations

While generative AI can be a valuable tool in simulating and predicting IAM vulnerabilities, it's important to supplement these simulations with regular security audits, testing, and proactive measures to address identified weaknesses and enhance overall security posture.

Additionally, organizations should consider ethical considerations and ensure that simulated attacks do not pose actual risks to the security and privacy of their systems and users.

Moreover, generative AI might assist in creating more robust training datasets for machine learning models used in identity verification. This can help in improving the accuracy and efficiency of identity recognition systems, reducing false positives and negatives.

It is important to note that while generative AI can provide valuable support in creating robust training datasets, careful consideration should be given to the ethical and legal implications of generating synthetic data, especially when dealing with sensitive identity-related information. Additionally, the performance of the model on real-world data should always be thoroughly evaluated to ensure its effectiveness in practical applications.

This potential progress raises ethical concerns regarding the misuse of generated data, privacy issues, and the need for stringent regulations to govern its usage in identity and access management. Enterprises will need to strike a balance between innovation and safeguarding user privacy and security.

Privacy concerns: The accuracy of generative AI models vary, leading to the creation of synthetic data that does not faithfully represent real-world diversity. If the generated data is biased or lacks diversity, it can result in machine learning models that are themselves biased and less effective in diverse environments.

Data accuracy and bias: The accuracy of generative AI models vary, leading to the creation of synthetic data that does not faithfully represent real-world diversity. If the generated data is biased or lacks diversity, it can result in machine learning models that are themselves biased and less effective in diverse environments.

Security risks: Generated data could be misused for malicious purposes, such as creating synthetic identities for fraudulent activities or exploiting vulnerabilities in identity verification systems. To minimise the risk of misuse, the training data should be secured, and the model should be trained and configured to not answer direct questions about individual people or attributes that could expose weaknesses in the environment.

Consent and informed decision-making: Individuals represented in generated data may not have given explicit consent for their synthetic identities to be created and used. This raises ethical questions about consent, informed decision-making, and the potential impact on individuals whose synthetic attributes may be associated with sensitive information. Thus, it is critical that as part of on-boarding, consent should be thought (even if optional) to using the user's data and information in future models and scenarios.

Regulatory compliance: There may be a lack of clear regulations governing the creation and use of generated data, especially in the context of identity and access management. The absence of robust regulations can contribute to misuse and potentially harmful applications of synthetic data. It is important to self-regulate in the absence of regulations and laws governing AI and make use of best practice rules that will ensure the organisations reputation and ethical standards are met.

Unintended consequences: The use of generated data may have unintended consequences, including reinforcing existing biases, creating new privacy risks, or inadvertently influencing decision-making processes. Evaluating and mitigating these unintended impacts is crucial for responsible use.

Educational and Awareness Gaps: There may be a lack of awareness and understanding among individuals, organizations, and policymakers about the implications of using generated data in IAM. Education and awareness programs are essential to inform stakeholders about the ethical considerations and potential risks.

Author

Mike Arakilo | Partner
Digital Identity Middle East Leader
miarakilo@deloitte.com

Contributors

Guus van Es | Partner | Cyber
guvanes@deloitte.nl

Hany Elkady | Director | Digital Identity
helkady@deloitte.com

This publication has been written in general terms and therefore cannot be relied on to cover specific situations; application of the principles set out will depend upon the particular circumstances involved and we recommend that you obtain professional advice before acting or refraining from acting on any of the contents of this publication.

Deloitte & Touche (M.E.) (DME) is an affiliated sublicensed partnership of Deloitte NSE LLP with no legal ownership to DTTL. Deloitte North South Europe LLP (NSE) is a licensed member firm of Deloitte Touche Tohmatsu Limited.

Deloitte refers to one or more of DTTL, its global network of member firms, and their related entities. DTTL (also referred to as "Deloitte Global") and each of its member firms are legally separate and independent entities. DTTL, NSE and DME do not provide services to clients. Please see www.deloitte.com/about to learn more.

Deloitte is a leading global provider of audit and assurance, consulting, financial advisory, risk advisory, tax and related services. Our network of member firms in more than 150 countries and territories, serves four out of five Fortune Global 500® companies. Learn how Deloitte's approximately 457,000 people make an impact that matters at www.deloitte.com.

DME would be pleased to advise readers on how to apply the principles set out in this publication to their specific circumstances. DME accepts no duty of care or liability for any loss occasioned to any person acting or refraining from action as a result of any material in this publication.

DME is a leading professional services organization established in the Middle East region with uninterrupted presence since 1926. DME's presence in the Middle East region is established through its affiliated independent legal entities, which are licensed to operate and to provide services under the applicable laws and regulations of the relevant country. DME's affiliates and related entities cannot oblige each other and/or DME, and when providing services, each affiliate and related entity engages directly and independently with its own clients and shall only be liable for its own acts or omissions and not those of any other affiliate.

DME provides services through 23 offices across 15 countries with more than 7,000 partners, directors and staff. It has also received numerous awards in the last few years such as the 2023 & 2022 Great Place to Work® in the UAE, the 2023 Great Place to Work® in the KSA, and the Middle East Tax Firm of the year.