Revolutionizing Cybersecurity:The Role of GenAI in Governance,Risk Management and Compliance

Agenda

- Introduction to GenAI
- Key Security Risks in GenAI Models
- Case Studies: Notable Security Breaches
- Risk Management in GenAI
- Ethical and Legal Considerations
- Conclusion

Introduction to GenAI

GenAI is a fascinating field of artificial intelligence that has the remarkable ability to generate new content, such as text, images, and videos, based on the patterns it learns.

GenAI finds applications across various industries such as Finance, Healthcare, Education and Automotive industry to name a few.

One of the critical aspects of GenAI is ensuring the integrity and authenticity of the generated outputs. This is crucial to prevent issues like deepfakes, synthetic identities, and financial fraud.

Attacks: Data Poisoning

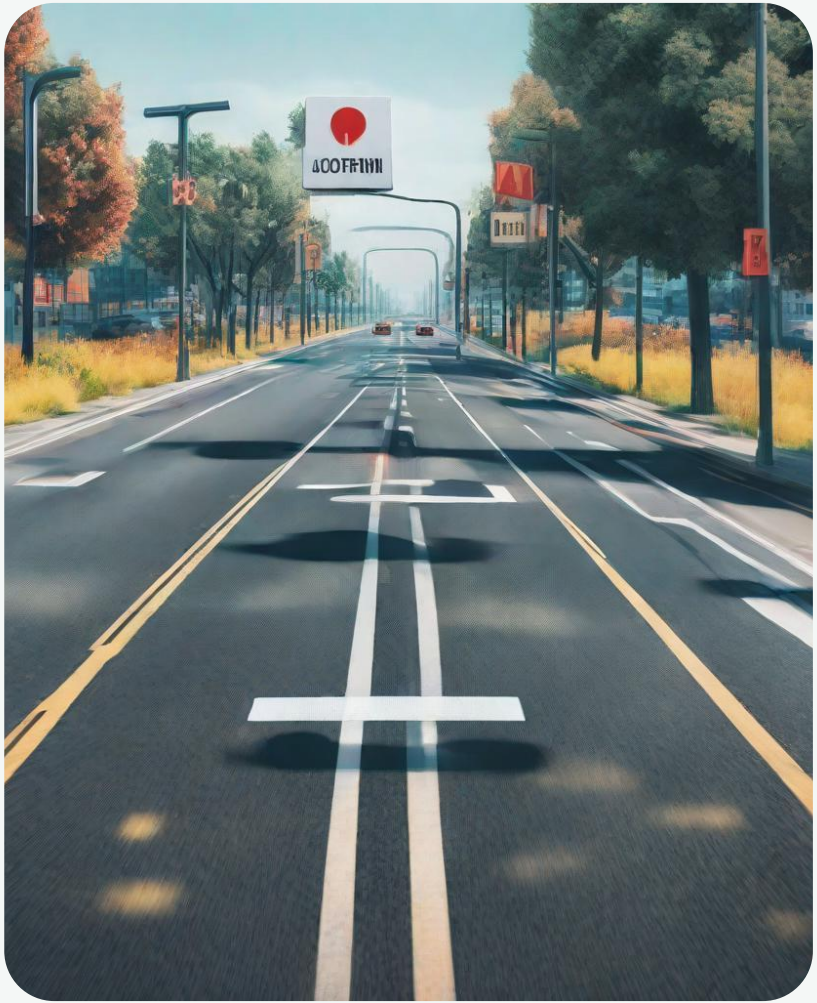
- Involves manipulating the training data of an AI model to produce incorrect, biased, or malicious outcomes
- In the context of GenAI this could mean embedding subtle, harmful biases into data sets, causing the AI to generate prejudiced or flawed content.
- The impact is profound as it can undermine the reliability and fairness of AI systems, potentially leading to loss of user trust and legal implications.

Model Inversion Attacks

- Aims to exploit AI systems to reveal sensitive information about the data they were trained on.
- By reverse engineering a models output an attacker can decipher the underlying properties and relationships between data.
- The attack is harmful for industries that handle sensitive information such as healthcare

Adversarial Attacks

- Designed to confuse AI models into making errors
- This can have great impact because small changes to input can cause incorrect outputs
- Example : A subtly modified image data can cause the AI model to misclassify an item incorrectly.
- Example a traffic sign being incorrectly classified, which could cause accidents.



Backdoor Attacks

- Involves attackers embedding a backdoor into a model while training it.
- This might cause the model to behave differently when specific inputs are provided.
- This causes the model to perform normally except when encountering specific input, causing it to erroneous or malicious output

Case Study One : Deepfake Misuse

1. In one video, a news anchor who looked beautiful heralded a country's role in geopolitical relations at an international summit meeting.

2. This undermines the trust the public has in the media. Since the content looks like human generated content, it is easy to be fooled.

3. This scenario underscores the urgent need for robust governance mechanisms that ensure the authenticity of media content.

Case Study Two: AI Driven Identity Theft

1. Cybercriminals used a voice changing AI software to mimic the voice of a company CEO,instructing a subordinate to transfer funds to an unauthorized account resulting in significant financial loss

2. The case study shows how cybercriminals can use voice synthesizing AI software to commit fraudulent activities, thereby exploiting trust in human voice authentication

3. The case study shows the importance of MFA and the need for setting up governance frameworks that proactively govern and audit internal processes to prevent such types of fraud, adhering to regulatory standards .

Case Study Three: AI Manipulation

1. An AI system used for trading was manipulated by data poisoning causing it to execute trades at a loss , causing irrecoverable loss to the trading firm

2. The criminals introduced infinitesimally small changes that were difficult to notice, corrupting the data fed to the AI trading algorithm,altering its trading patterns

3. The case study shows the importance of AI governance in Finance & the importance of incorporating security early in the training phase and the need for continuous auditing to ensure the integrity and reliability

GenAI in Governance

INTRO

- In the realm of corporate governance, Generative AI is revolutionizing how decisions are made.
- Transparency is critical in governance to build trust among stakeholders.

EXAMPLE

- By integrating real-time data analysis, AI can provide executives with decision-making support that adjusts dynamically to changing market conditions
- AI systems can track and record all decision paths and processes, making it easy to audit and understand the basis of decisions

BENEFIT

- The primary benefit here is speed and adaptability—qualities essential for staying competitive .
- AI can help meeting compliance requirements but also reinforce stakeholder confidence in the organization's governance practices

GenAI in Risk Management

INTRO	EXAMPLE	BENEFIT
• GenAI is revolutionizing risk management by enabling sophisticated analysis of vast data sets to identify and assess risks.	• Financial institutions are using AI to scan for unusual transaction patterns that might indicate fraud or money laundering.	• This enables organizations to proactively identify risks, often before they manifest, allowing for effective risk management and reduced financial losses.
• GenAI enhances predictive analytics, enabling organizations to foresee and react to potential threats in real time	• Cybersecurity firms employ AI systems that predict and detect cyber attacks by analyzing network traffic patterns	• Real-time threat detection allows organizations to respond to threats instantly, minimizing damage and enhancing overall security posture

GenAI in Risk Management

INTRO

- Generative AI is revolutionizing compliance monitoring and management by automating the detection of compliance issues and managing them efficiently.
- AI technology simplifies and automates compliance processes, including the handling of vast amounts of documentation required by regulatory bodies.

EXAMPLE

- Healthcare providers use AI to monitor patient care and treatment records to ensure compliance with health regulations and patient safety standards.
- A drug company utilizes AI tools to automatically document drug testing processes and outcomes, ensuring compliant records are maintained for regulatory review.

BENEFIT

- This automation not only streamlines the monitoring process but also reduces human errors, ensuring higher compliance and patient safety
- Automation speeds up the documentation process, reduces administrative burdens, and ensures accuracy and compliance in record-keeping

Risk Management in AI

- A comprehensive risk assessment is needed to mitigate attacks against AI models
- This involves analysis of data sources and methodologies used in AI training
    - Vulnerability Assessments
    - Threat Modelling
    - Security Audits

- Utilizing AI security platforms to continuously monitor AI systems in realtime to detect & respond to attacks.
- Implementing Security Controls such as :
    - Encryption and Access Controls
    - Patch Management
    - Incident Response Plan
- Incorporating adversarial examples during training phase of AI models

Ethical Considerations for AI

- Security in AI is also an ethical imperative
- Crucial to maintain public trust and upholding moral responsibilities
- Ethical Considerations include :
  - Transparency
  - Accountability
  - Privacy
  - Fairness

Legal Considerations for AI

- Security in AI is being scrutinized by regulatory bodies worldwide, aiming to create a safe and equitable environment for AI deployment
- Key frameworks include
    - General Data Protection Regulation (GDPR)
    - California Consumer Privacy Act (CCPA)
    - Biometric Information Privacy Act (BIPA)

Challenges

- Being compliant with regulatory frameworks presents challenges for organizations that deploy AI models globally
- Challenges Include
  - Regular Audits
  - Data Management Strategies
  - Cross Functional Compliance Teams
  - Continuous Education/Awareness

Conclusion

In today's evolving technology landscape, the role of security is important to not only protect against threats but also to maintain the trust and confidence of users and stakeholders.

As AI technologies keep advancing at a rapid pace, so will the sophistication of threats to against these systems.Its crucial the community-users,developers & regulators remain vigilant.

The complex landscape of AI and security demonstrates the necessity of a proactive approach—where security is not an afterthought, but a foundational component of all AI initiatives

THANK YOU

linkedin.com/akshaysekar