Deloitte.

THE DAWN OF GENERATIVE AI: INFINITE POSSIBILITIES

Deloitte.

THE DAWN OF GENERATIVE AI: INFINITE POSSIBILITIES, INFINITE RISKS?

TODAY'S SPEAKERS

JORDAN MCKENZIE
Manager
Deloitte Cyber
jormckenzie@deloitte.de

LUCIE WOLLENHAUPT
Manager
Deloitte Cyber
lwollenhaupt@deloitte.de

MANUEL BOLKART
Manager
Deloitte Cyber
mbolkart@deloitte.de

Deloitte.

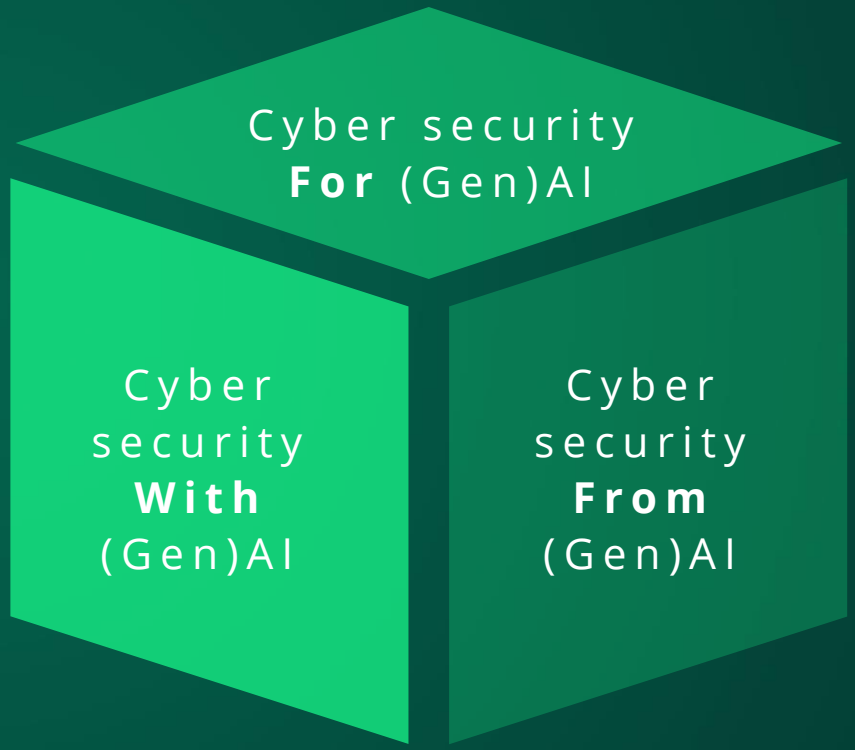
CYBER MEETS (GEN)AI

Deloitte's framework for securing Generative AI and AI Deep Dive: Its impact on scaling cyber services

October 2024

CYBER SECURITY MEETS (GEN)AI

CYBER SECURITY AND GENERATIVE AI AND AI CAN BE VIEWED WITHIN **3 DIFFERENT DIMENSIONS**.



CYBER SECURITY **FROM** (GEN)AI

Adapting to the rapidly evolving cyber security threat landscape due to the evolution of new and more sophisticated types of cyber attacks.

CYBER SECURITY **FOR** (GEN)AI

Protecting (Gen)AI systems from cyber security threats, by providing guidance to secure implemented or planned (Gen)AI use-cases.

OUR GOAL WAS TO **IDENTIFY** THE **THREATS** AND **ATTACK VECTORS** TARGETING (GEN)AI SYSTEMS, AIMING TO UNDERSTAND AND **MITIGATE** THEIR **VULNERABILITIES**.

OUR GOAL WAS TO **IDENTIFY** THE **THREATS** AND **ATTACK VECTORS** TARGETING (GEN)AI SYSTEMS, AIMING TO UNDERSTAND AND **MITIGATE** THEIR **VULNERABILITIES**.

What are the **threats** targeting (Gen)AI systems?

- INPUT INJECTION
- TRAINING DATA POISONING
- MODEL POISONING
- MODEL STEALING
- ADVERSARIAL EXAMPLES
- ...

Deep Dive

OUR GOAL WAS TO **IDENTIFY** THE **THREATS** AND **ATTACK VECTORS** TARGETING (GEN)AI SYSTEMS, AIMING TO UNDERSTAND AND **MITIGATE** THEIR **VULNERABILITIES**.

- What are the **threats** targeting (Gen)AI systems?
- In which stage of the **(Gen)AI Lifecycle and domains** does the threat occur?
- Which **regulations** & **industry standards** already provide security guidance for (Gen)AI?
- How can we **tackle the threats** and **mitigate** the **vulnerabilities**?

Deep Dive

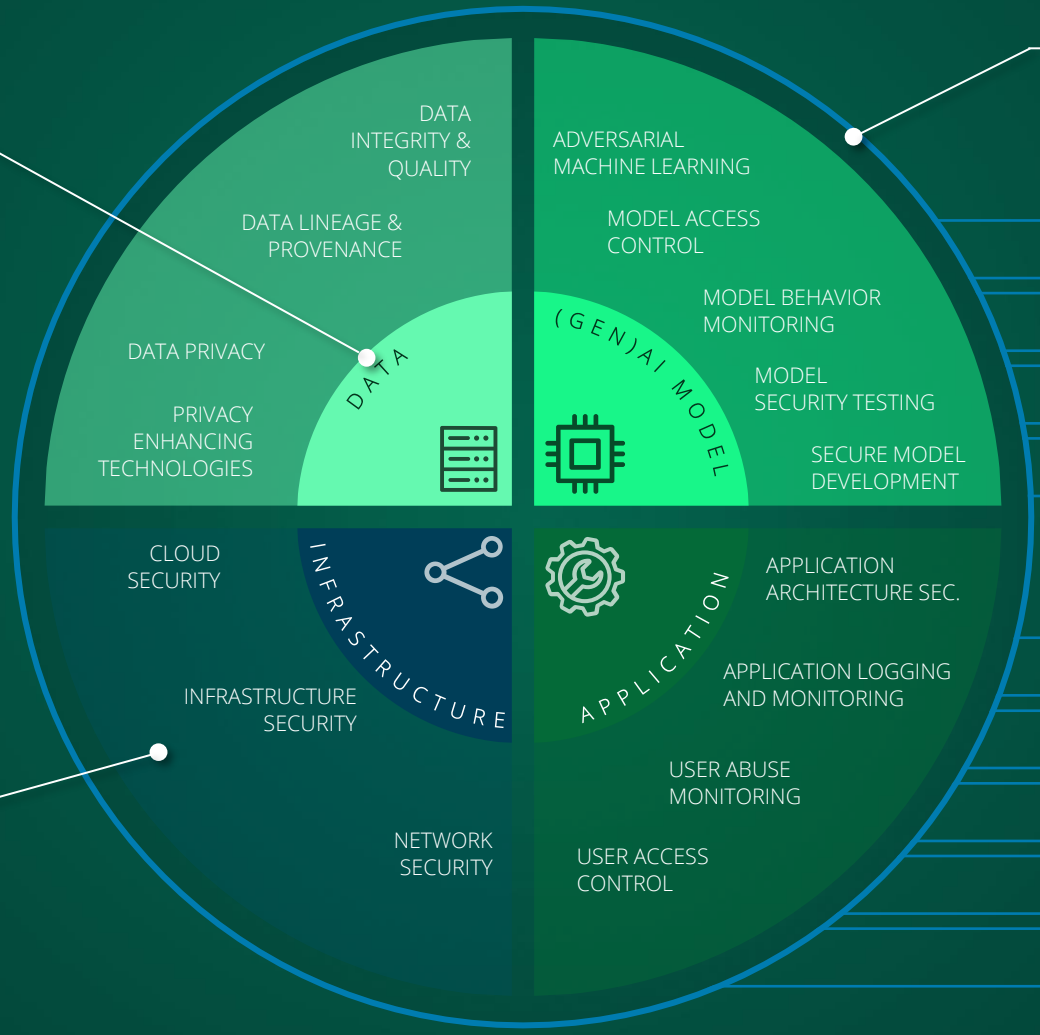
CYBER SECURITY FOR (GEN)AI FRAMEWORK*

(GEN)AI DOMAINS

The Domains constitute the core structure of (Gen)AI systems and are used to cluster security capabilities.

(GEN)AI SECURITY CAPABILITIES

A Security Capability is a category for grouping of controls which address specific cyber security threats in every domain.



OVERARCHING SECURITY CAPABILITIES

Capabilities to ensure the security of (Gen)AI solutions over all domains.

LIFECYCLE SECURITY

- SECURE DEVELOPMENT PROCESS
- SECURE SUPPLY CHAIN
- ASSET MANAGEMENT
- DATA LOSS PREVENTION (DLP)

GOVERNANCE, RISK & COMPLIANCE

- THIRD-PARTY RISK MANAGEMENT
- (GEN)AI SECURITY RISK MANAGEMENT
- (GEN)AI SPECIFIC POLICIES, STANDARD & ARCHITECTURE
- BUSINESS CONTINUITY MANAGEMENT
- REGULATORY COMPLIANCE

Deep Dive

*based on ISO standard (ISO/SAE42001:2023), OWASP (threats for LLM and ML) and enisa publications, MITRE-ATLAS insights

CYBER SECURITY **WITH** (GEN)AI

Improving cyber security capabilities and boosting cyber security processes by including (Gen)AI.

CYBER SECURITY **WITH** (GEN)AI

Improving cyber security capabilities and boosting cyber security processes by including (Gen)AI.

For Example

(Gen)AI enhanced Threat Intelligence

Cyber Security Chatbots

AI supported Incident Handling

CYBER SECURITY **WITH** (GEN)AI

Improving cyber security capabilities and boosting cyber security processes by including (Gen)AI.

AI supported Incident Handling

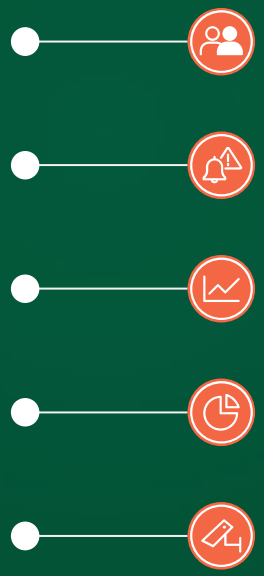
Focus on: overcoming Challenges in Security Operations with the Power of GenAI

PROVIDING **SECURITY OPERATIONS** CAN BE **CHALLENGING** IN VARIOUS WAYS INDEPENDENT FROM THE SIZE OR SECTOR OF YOUR ORGANIZATION.

Deep Dive

PROVIDING **SECURITY OPERATIONS** CAN BE **CHALLENGING** IN VARIOUS WAYS INDEPENDENT FROM THE SIZE OR SECTOR OF YOUR ORGANIZATION.

These challenges are...



- SKILLED PERSONNEL SHORTAGE
- CONSTANT EMERGING THREAT LANDSCAPE
- HIGH COST
- ALERT FATIGUE
- REAL-TIME MONITORING

Deep Dive

(GEN)AI SUPPORTED ALERT INVESTIGATION & INCIDENT HANDLING

SKILLED PERSONNEL SHORTAGE	AUTOMATED THREAT INVESTIGATION
CONSTANT EMERGING THREAT LANDSCAPE	PREDICTIVE ANALYTICS
HIGH COST	COST EFFECTIVE AUTOMATION
ALERT FATIGUE	ALERT FILTERING
REAL-TIME MONITORING	ANOMALY DETECTION

Deep Dive

THANK YOU FOR YOUR ATTENTION AND WE LOOK FORWARD TO CONNECTING WITH YOU AFTERWARDS!

Meet us at Hall 7A, Booth 520!

JORDAN MCKENZIE
Manager
Deloitte Cyber Strategy
jormckenzie@deloitte.de

LUCIE WOLLENHAUPT
Manager
Deloitte Cyber Strategy
lwollenhaupt@deloitte.de

MANUEL BOLKART
Manager
Deloitte Cyber Defense
mbolkart@deloitte.de