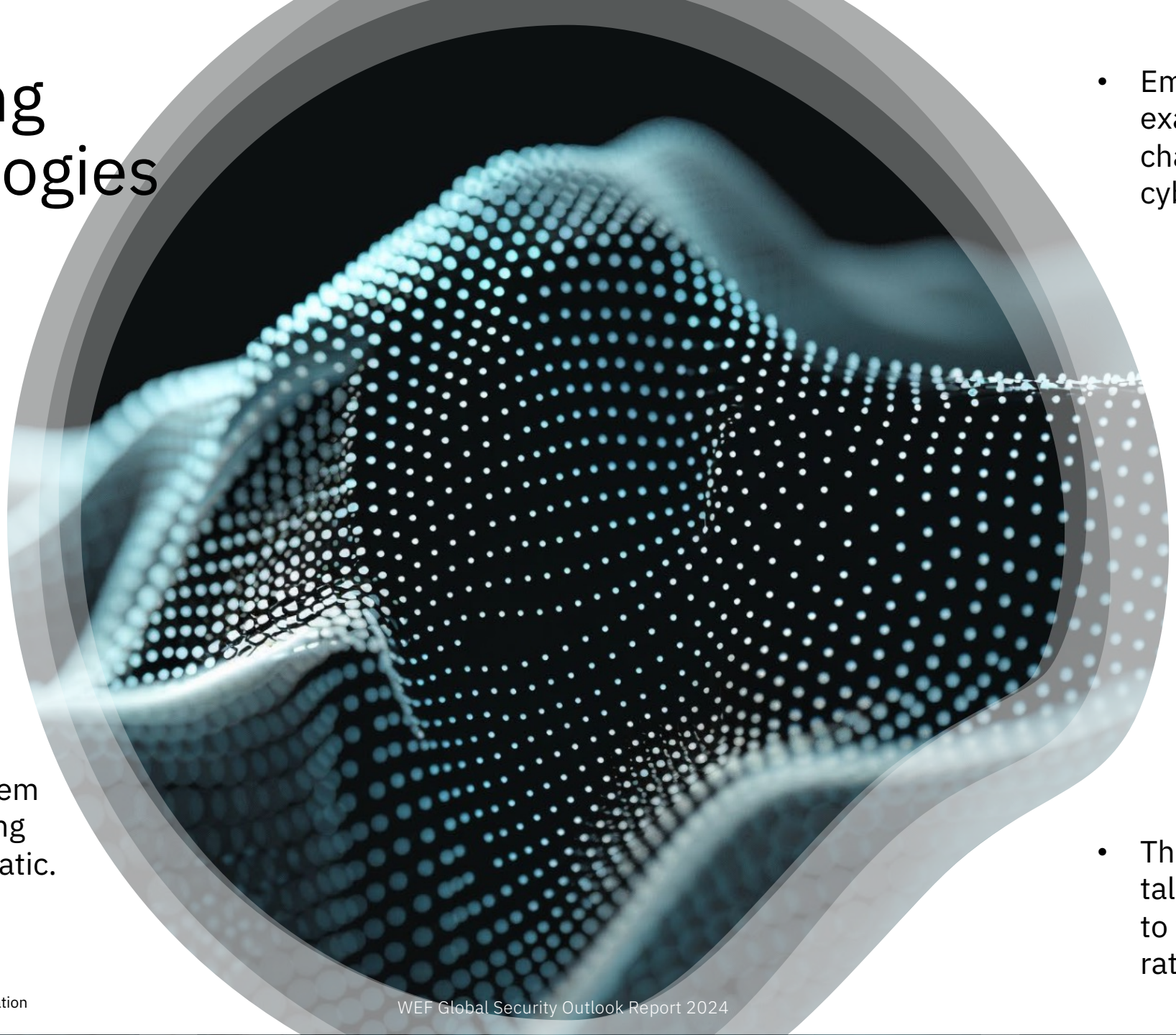
AI in **Security**

a force multiplier for good and bad

Angela Lopes
IBM Security Technical Sales Leader
Latin America

Emerging Technologies



- Emerging technology will exacerbate long-standing challenges related to cyber resilience.

- Cyber ecosystem risk is becoming more problematic.

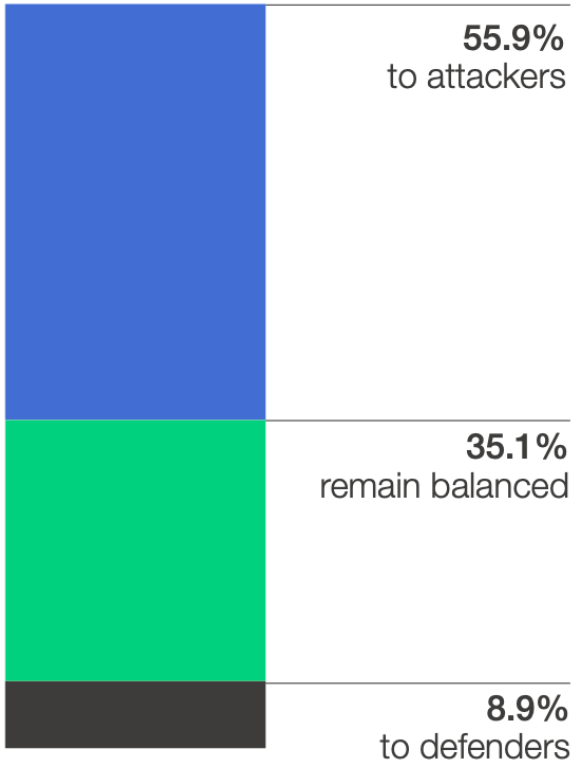
- The cyber-skills and talent shortage continues to widen at an alarming rate.

Emerging Technologies

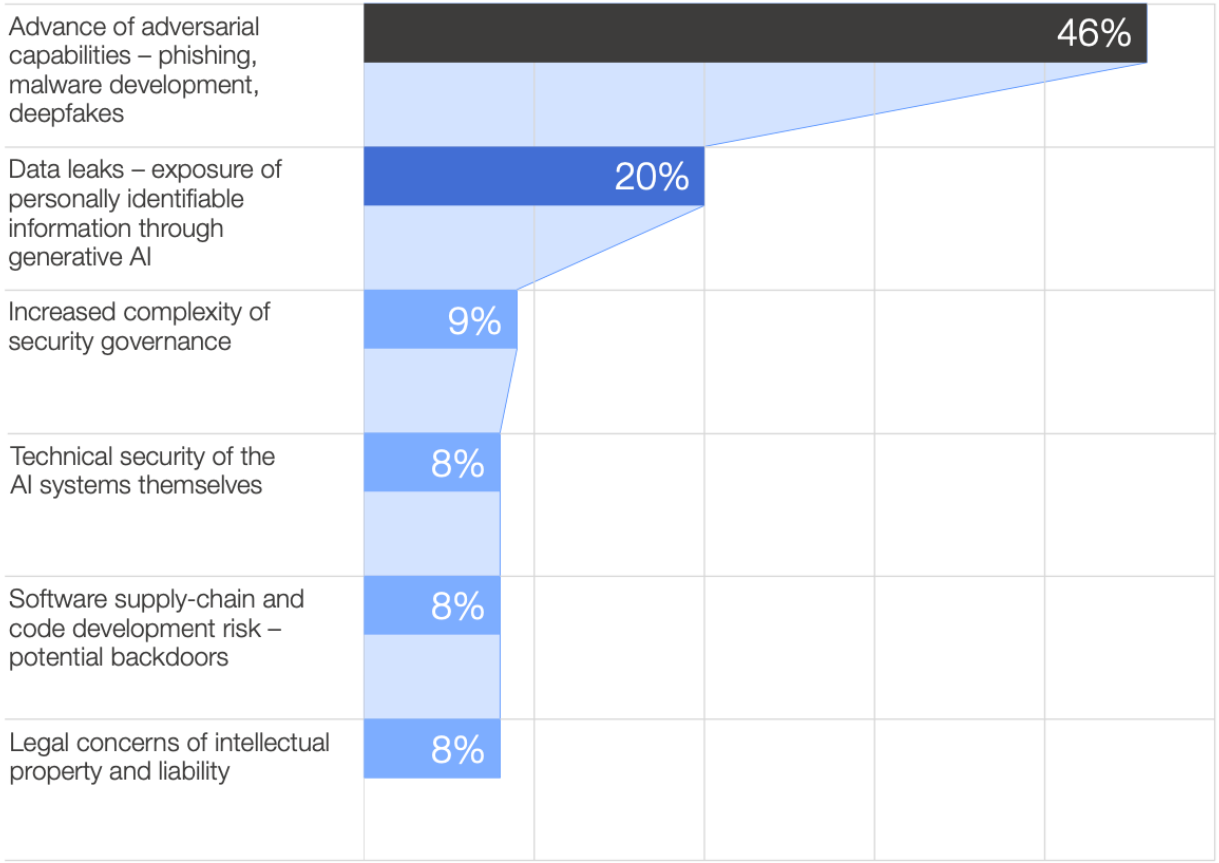
Generative AI's impact on Cyber

Emerging technologies will exacerbate long-standing challenges related to cyber resilience

In the next two years, will generative AI provide overall cyber advantage to attackers or defenders?



What are you most concerned about in regards to generative AI's impact on cyber?



Cyber vulnerabilities

New tools and capabilities will open new markets for criminal networks, with cybercrime offering an increasingly low-risk and low-cost revenue stream for organized crime.

Phishing attacks, for example, can now be easily and accurately translated into minority languages using generative AI.

Latin America:
What we heard from CISOs and Security leaders in 2H2023

- "… Identify incidents quickly and respond faster. Automation with AI is in fashion."
- "…Don't see AI replacing people in security operations."
- "…Cybersecurity teams are reduced against millions of attackers."
- "…New technologies must be allies in order to accelerate the detection process."
- "…Automate, orchestrate, generate efficiency, with AI and reduce time."
- "…The attacker is going to use AI to attack them. If solutions do not use AI to facilitate incident response and feed them more information to respond to that, they will be behind."
- "…Challenging balance between corporate use of generative AI vs user exposing sensitive data."

Interviews and activities executed in 2H2023 around Latin American countries

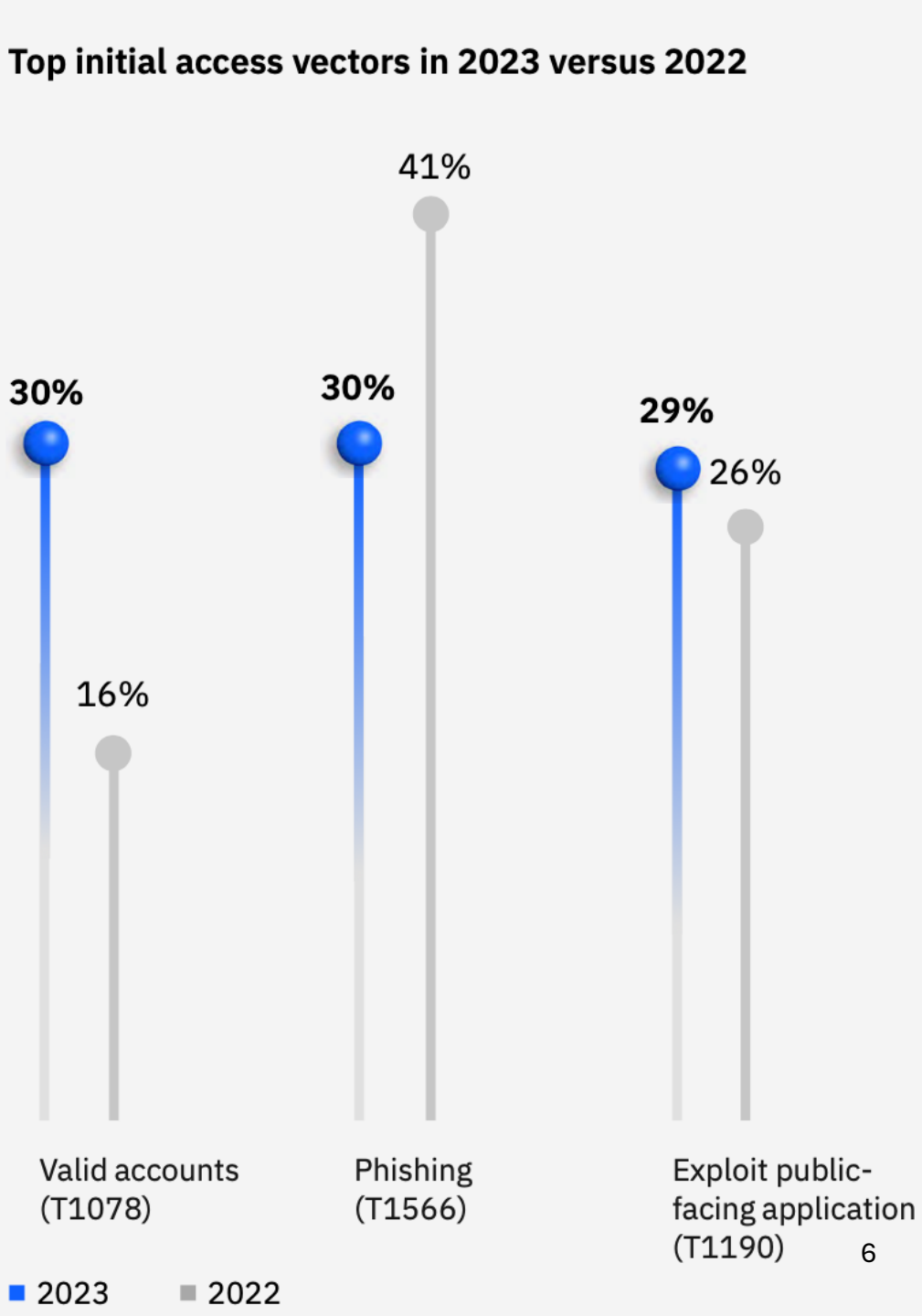
Cyberthreat landscape

71%

Valid accounts as a preferred initial access technique among cybercriminals – tying with phishing for the first time...

266%

... accompanied by an upsurge in malware designed to steal information, known as infostealer malware.



Top initial access vectors X-Force observed in 2022 and 2023.
Sources: X-Force and MITRE ATT&CK Matrix1 for Enterprise framework

Exploiting the human attack surface

Attackers are logging in not hacking in...

... using harvested credentials.

Attackers are monetizing the human attack surface...

... moving to human attack surface to not depend on or rely on vulnerabilities.

Regarding Ransomware attack, they are focus on extortion.

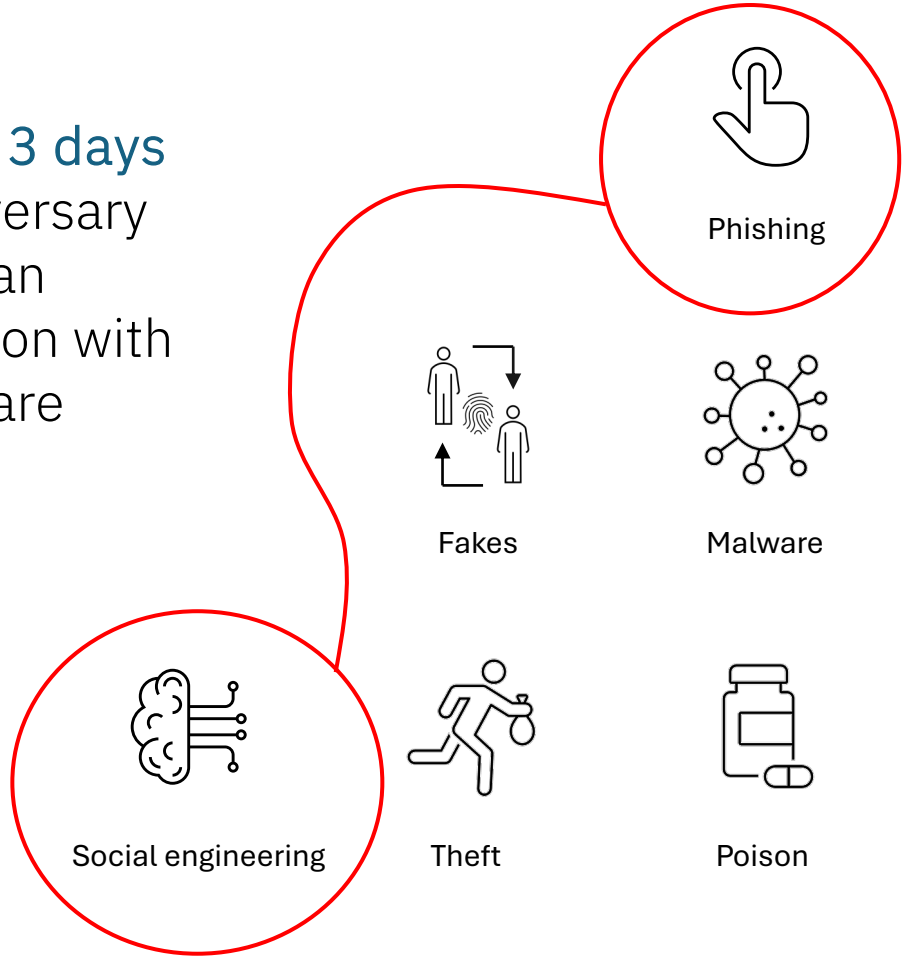
11.5%

Drop in enterprise ransomware incidents

- Log in
- Steal the data
- Do not encrypt it
- Intimidate
- Threat to expose something that will shame the victim
- Extorsion

Gen AI is making adversaries faster and more sophisticated

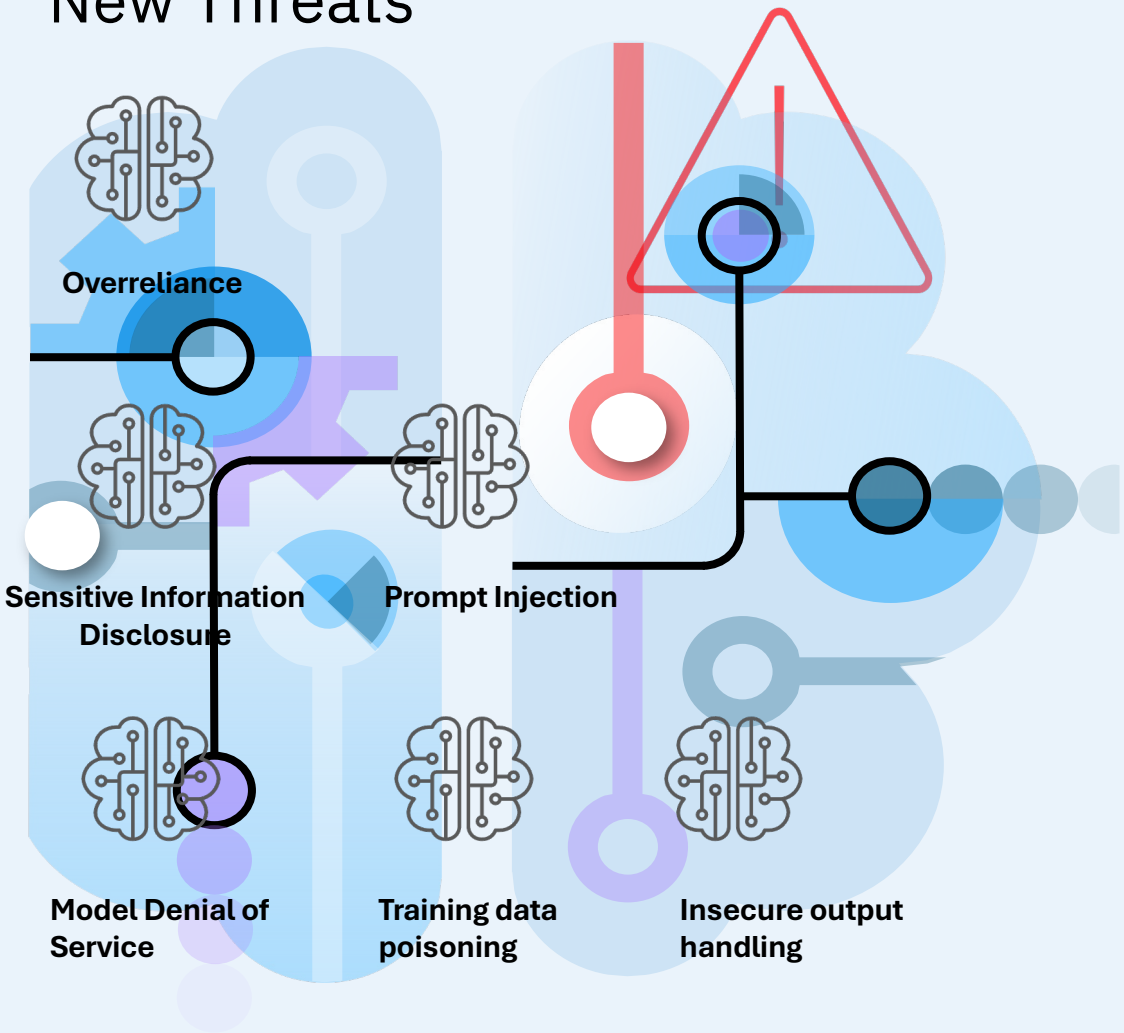
Less than 3 days for an adversary to attack an organization with ransomware



5 minutes X 16 hours

Click rate: 11% AI X 14% X-Force

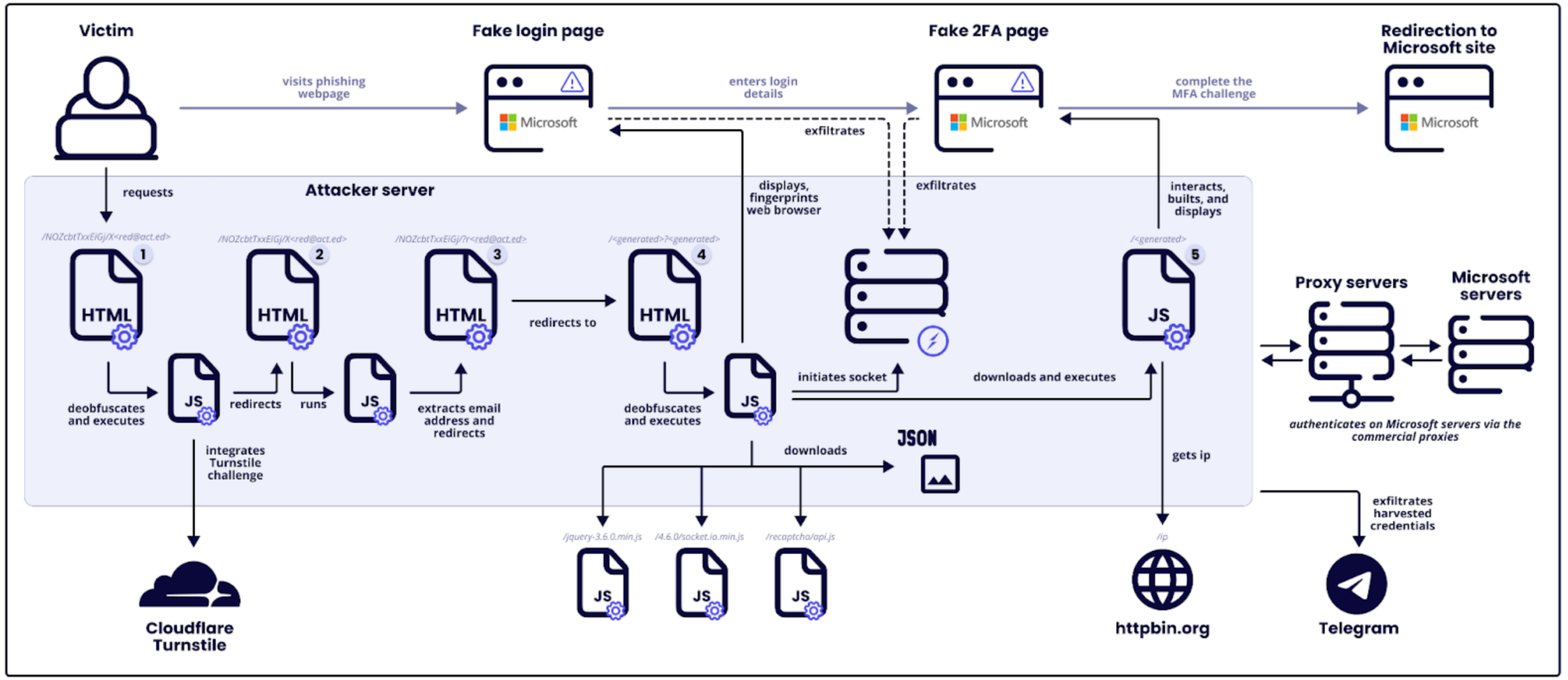
New Threats



By tricking a Gen AI model with just 5 simple instructions

Phishing-as-a-service (PhaaS) platform

sekoia | Main operations of the Tycoon 2FA phishing kit, as of March 2024

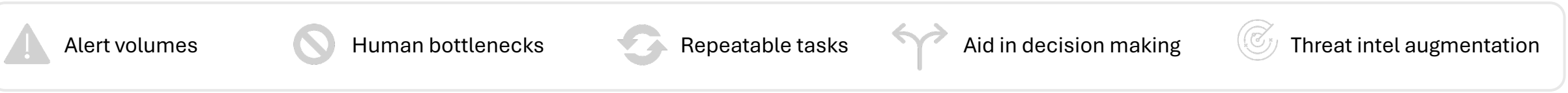


New MFA-bypassing phishing kit targets Microsoft 365, Gmail accounts

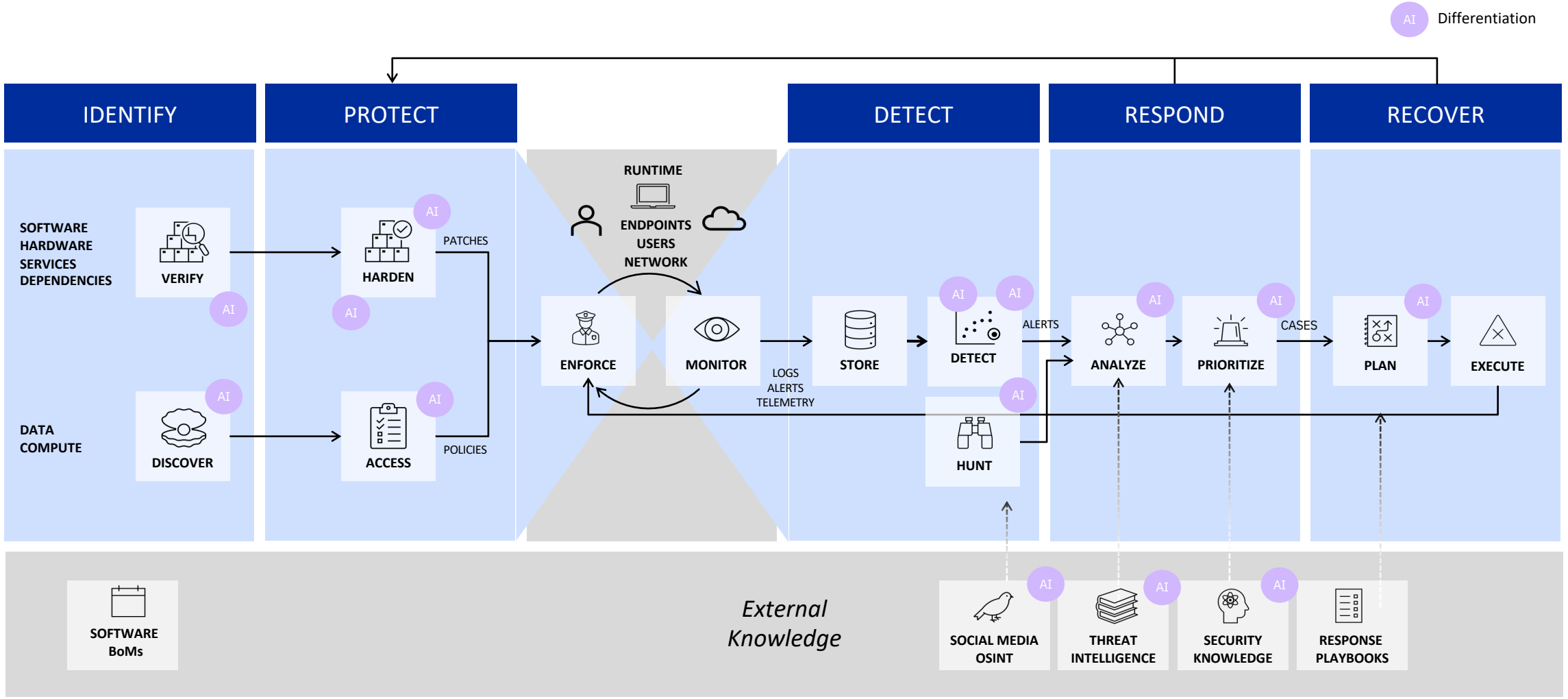
AI for Security

AI & Gen AI > Use cases we hear from clients

Use cases tend to be centered around these themes

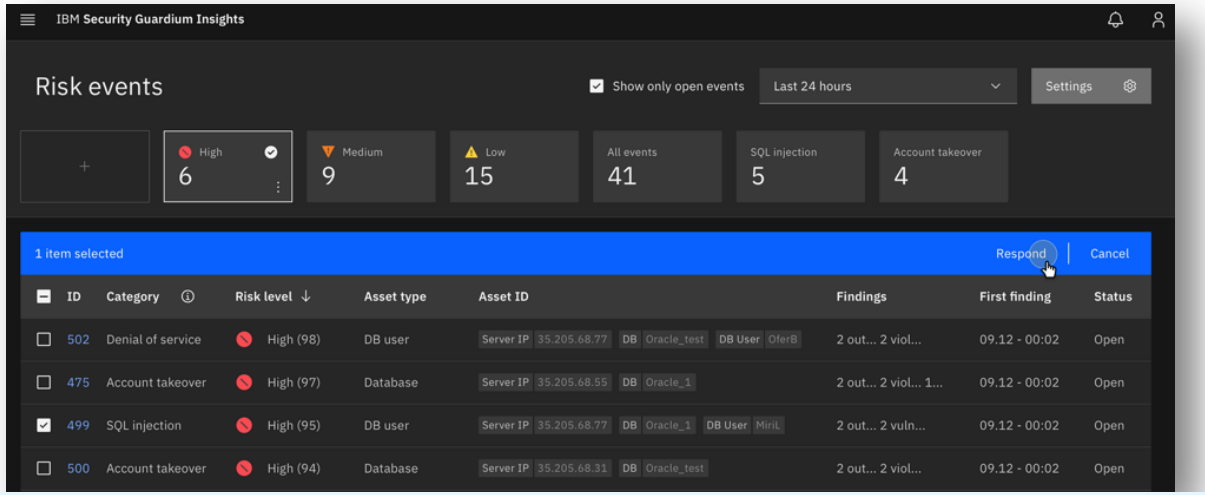


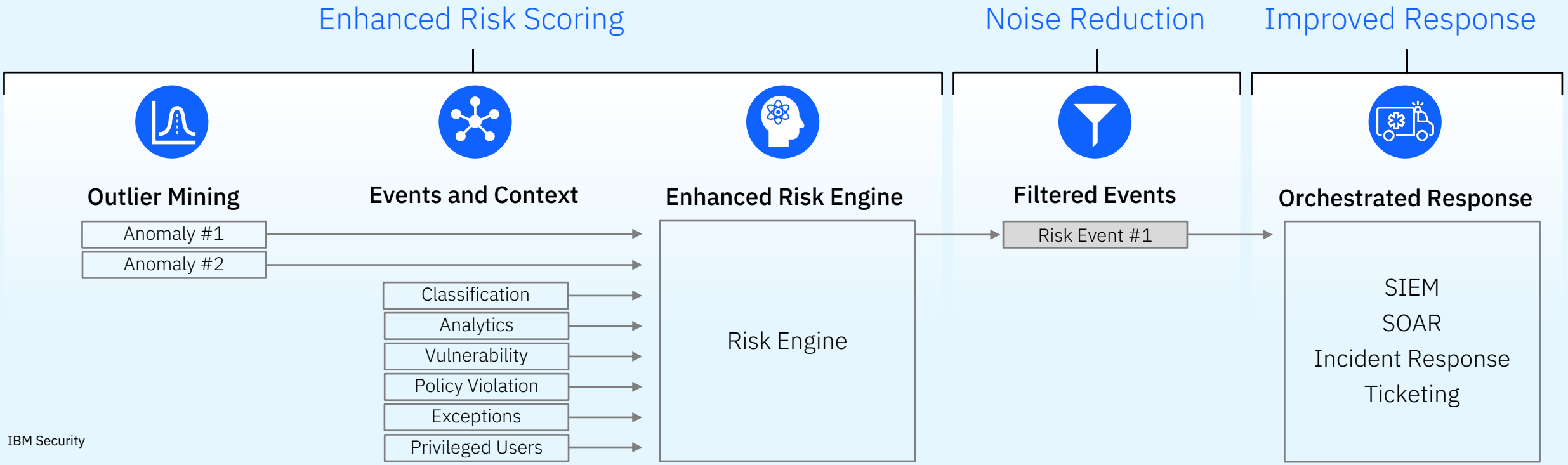
Implementing AI across the five dimensions of the NIST Cybersecurity Framework is critical for improving cybersecurity operations



Actionable intelligence

Automate remediation efforts with context-based risk scoring and SIEM event pre-filtering to help reduce costs and increase efficiency





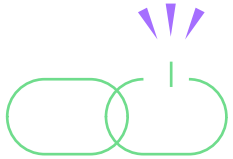
AI-fueled Identity Threat Detection and Response (ITDR) and Identity Security Policy Management (ISPM)

“Hackers don’t break in, they login"

Continuous identify security posture and threats detection

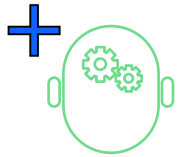
Discover Security Blind Spots

Shadow access, directories
Unauthorized local accounts
MFA bypass
Dormant service accounts



Expose Infrastructure Gaps

Dangerous identities
Misconfigurations and hazardous deviations from security policy

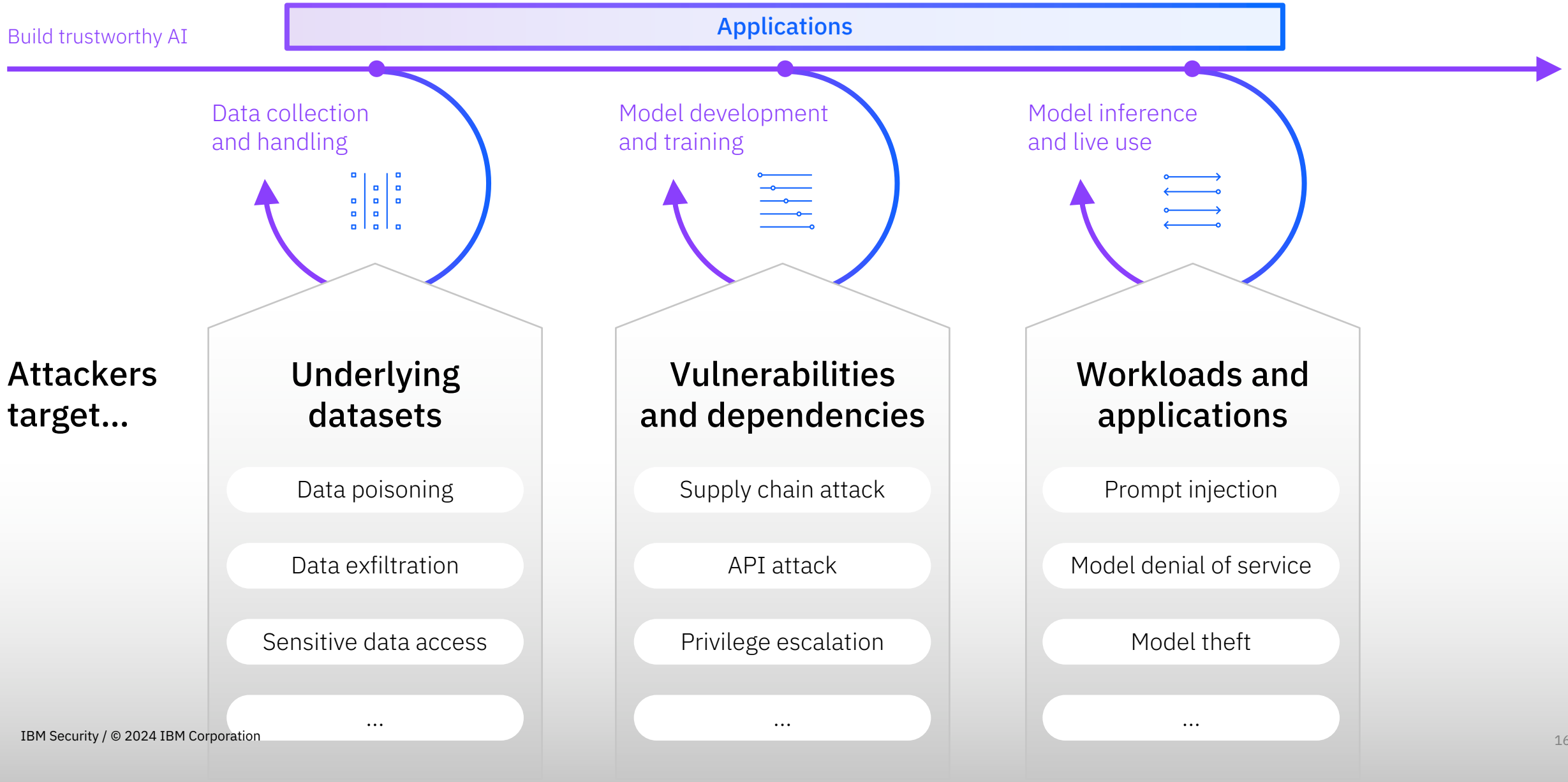


Detect Risky Activity

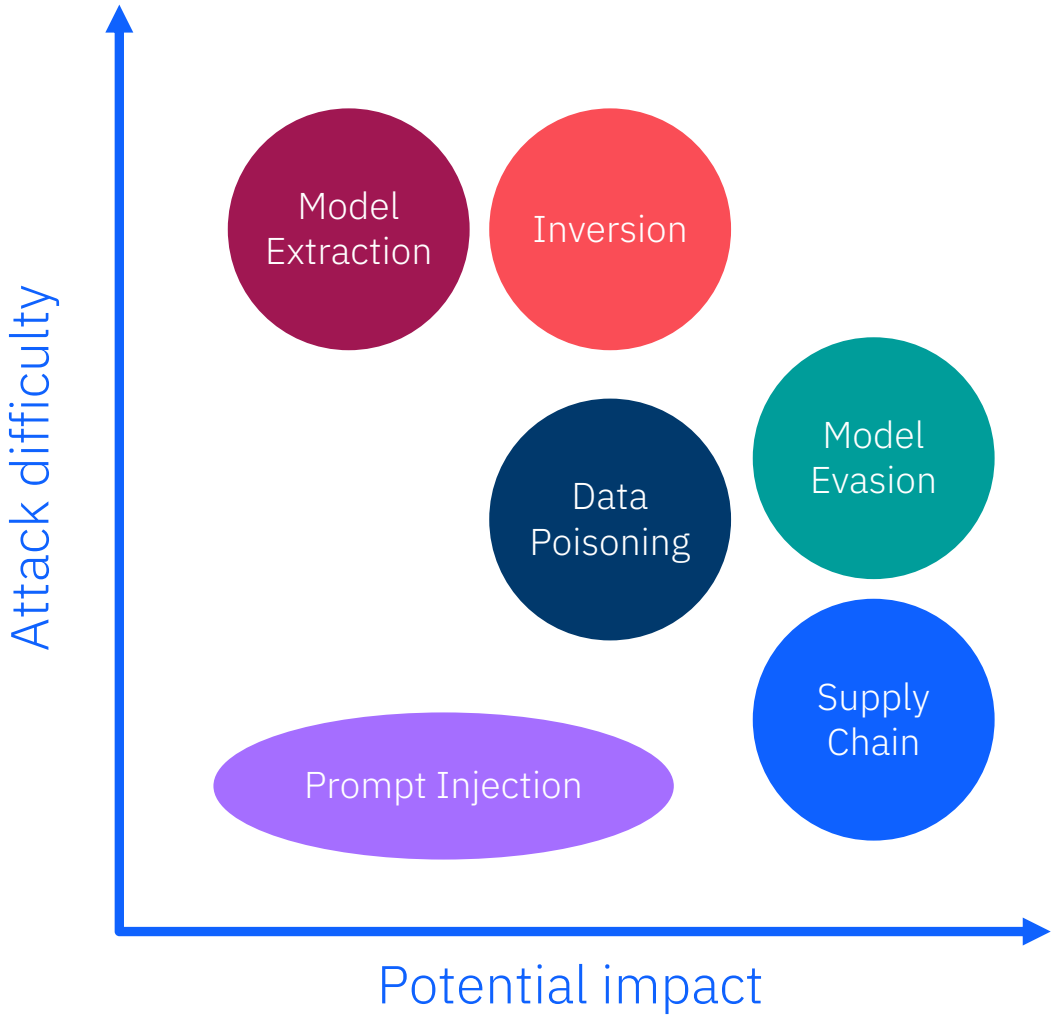
Identify exposures, threats and attacks in real-time with detailed contextual analysis

Security for AI

AI pipeline security risks



Understanding adversarial risks to AI



Prompt Injection

Manipulate AI models into performing unintended actions, by dropping guardrails and limitations put in place by the developers

Data Poisoning

Changing the behavior of AI models by altering the data used to train them

Model Evasion

Circumventing the intended behavior of an AI model by crafting inputs that trick them

Model Extraction

Steal a model's behavior by observing the relationships between inputs & outputs

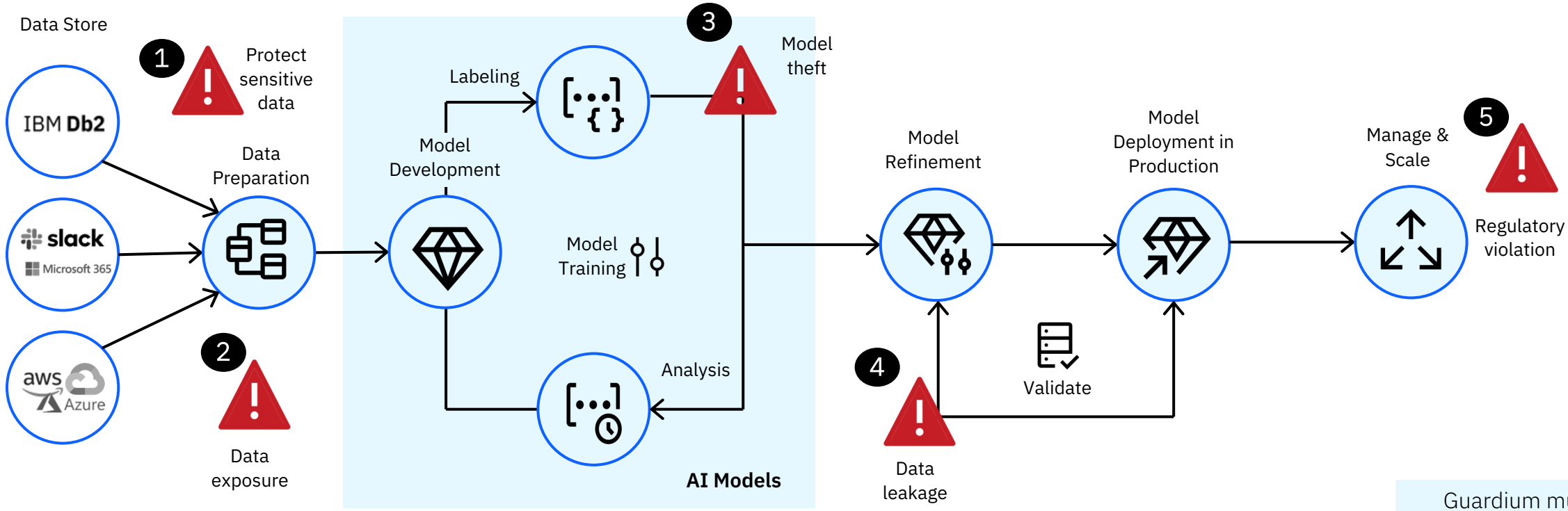
Inversion Attacks

Reveal information on the data used to train a model, despite only having access to the model itself

Supply Chain Attacks

Generate harmful models that hide malicious behavior, or target vulnerabilities in systems connected to the AI models.

Securing the AI pipeline with Guardium



1 **Protect sensitive data**
Prevent usage of sensitive information in AI model found in big data (ex. Parquet) and meet compliance

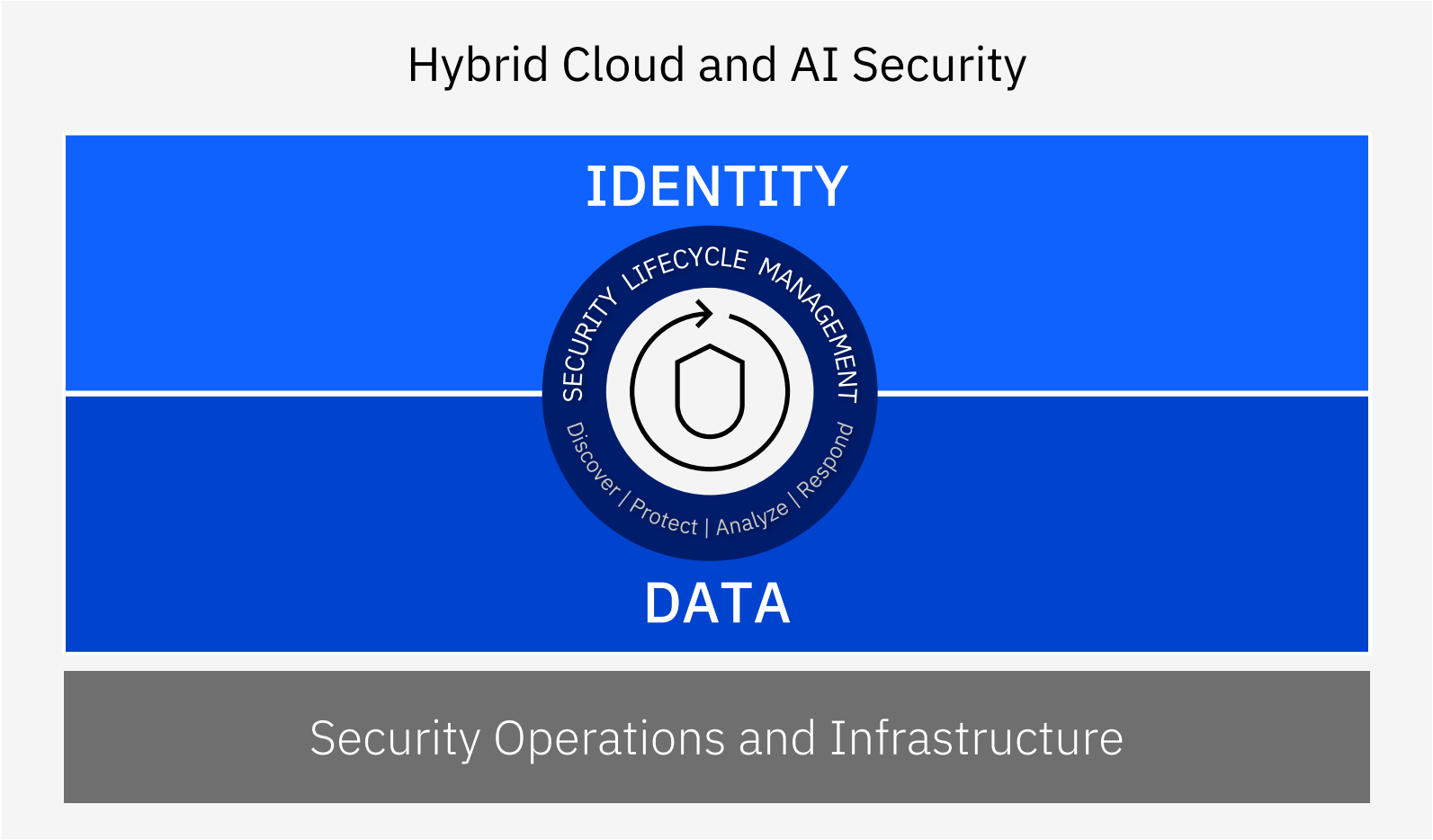
2 **Unintentional exposure**
Monitor copies of sensitive data across apps and clouds to prevent exposure to a 3rd party (ex. unapproved open source models)

3 **Keys or tokens found in Slack**
Detect if keys are exposed in shared slack channels, which can give access to attackers who aim to steal IP

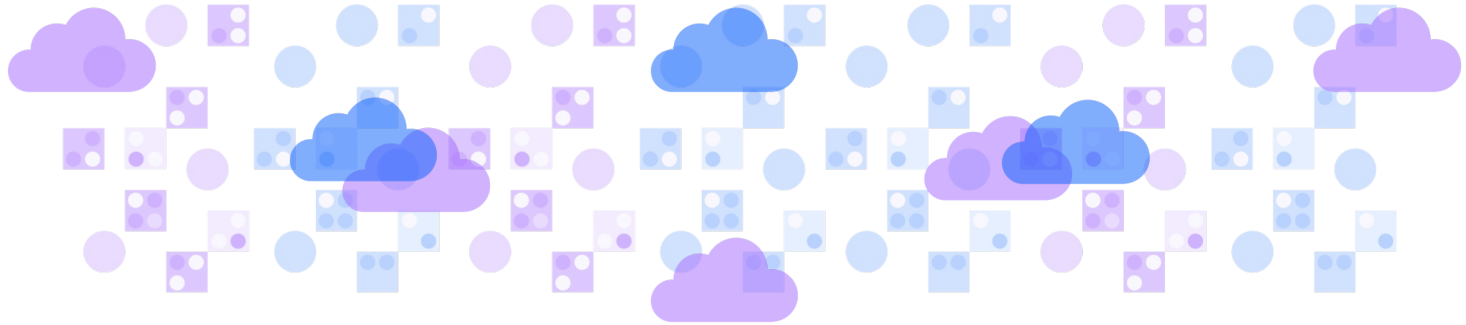
4 **Prevent data leakage**
Detect if sensitive data (ex. PII) has leaked during model validation (could be from insider threat)

5 **Data Transfer Violation**
Monitor data flow to ensure compliance across SaaS applications in various geographies (ex. GDPR)

Secure hybrid cloud and AI



IBM’s unique approach to cybersecurity



Simplify the complexity of securing
HYBRID CLOUD

Data is expanding across hybrid cloud platforms

Discover, protect, and govern sensitive data, regardless of location

Managing access across hybrid clouds is complex

Simplify identity management across environments with an identity fabric approach

Trust the secure adoption of
ARTIFICIAL INTELLIGENCE

AI is the new unprotected attack surface

Ensure only trusted employees and consumers have access

GenAI adoption is outpacing trusted security approaches

Safeguard training data and large language model integrity

Thank you

Follow us on:

ibm.com/security

securityintelligence.com

ibm.com/security/community

xforce.ibmcloud.com

@ibmsecurity

youtube.com/ibmsecurity


IBM