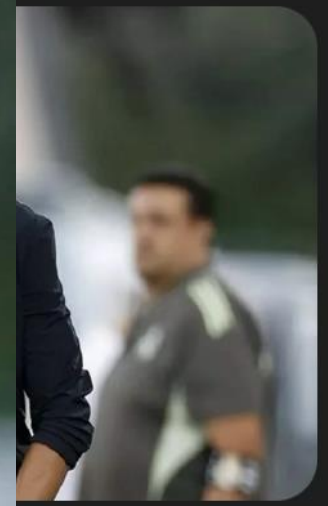
SOCInfo Digital

12 marzo 2026, Santiago de Compostela

TrendAI Galicia AAPP:

Leopoldo_Abascal@trendmicro.com

IA Generativa - Imágenes



TrendAI

OWASP LLM

TrendAI™

Aligned with OWASP Top 10

LLM01:2025 Prompt Injection
LLM02:2025 Sensitive Information Disclosure
LLM03:2025 Supply Chain
LLM04:2025 Data and Model Poisoning
LLM05:2025 Improper Output Handling
LLM06:2025 Excessive Agency
LLM07:2025 System Prompt Leakage
LLM08:2025 Vector and Embedding Weakness
LLM09:2025 Misinformation
LLM10:2025 Unbounded Consumption (DoS)

TrendAI

Ej:#1

Prompt Injection

El atacante introduce un comando como parte de los datos para manipular el comportamiento del modelo.

Ej:#2

Data Model Poisoning

Se introducen datos maliciosos para alterar los datos del modelo y afectar a consultas posteriores.

Ej:#3

Unbounded Consumption (DoS)

Uso abusivo del modelo para generar costes o un DoS.

Otros Riesgos

- Shadow AI.
- Fuga de datos.
- Uso en aplicaciones no autorizadas (ej.: PDF).

Trend... AI

TrendAI™

Historia de innovación impulsada por IA de Trend Micro

Comenzamos en 2005 con soluciones anti-spam y continuamos invirtiendo en todas las formas de IA, incluyendo la más reciente IA Generativa.

2005
Anti-spam using ML

2000's
SVM/Deep Learning for phishing & malicious URLs

2000's
NLP for DGA

2010's
Belief Propagation & Markov Model for Lateral Movement

2010's
Network Exploits using SVM

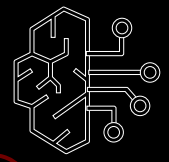
2010's
Writing Style DNA for BEC

2010's
Computer Vision for phishing

2010's
TrendX pre\post execution file detection

2020's
Agentic AI to build AI SOAR & SIEM & Shift Left

2020's
GenAI Companion for AI SOC & CREM



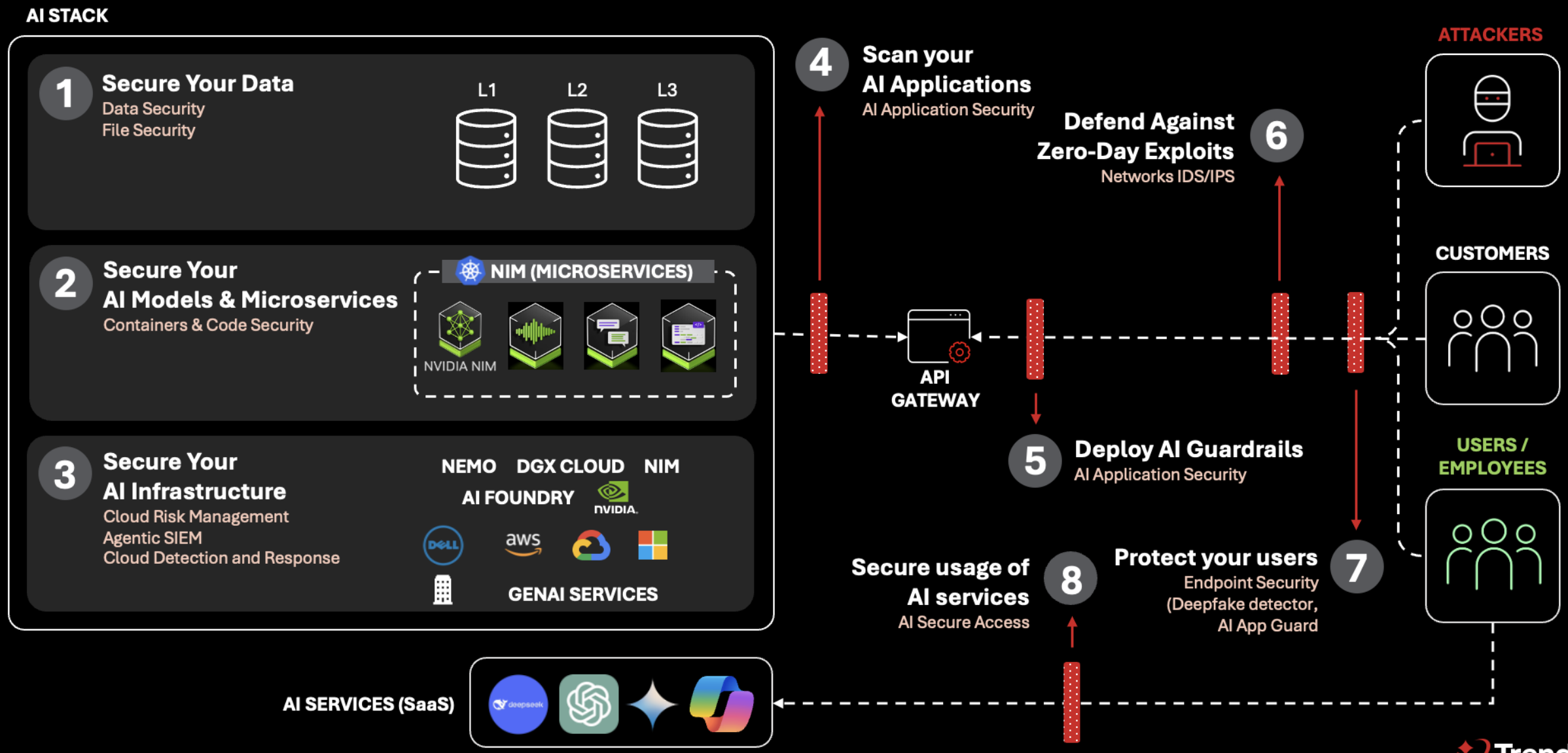
2020's
Vulnerability prioritization models used in CREM

2020's
Isolation forest used in anomalous network traffic detection

2020's
Zero Trust AI Regression analysis for account compromise detection

TrendAI

AI Security – Blueprint

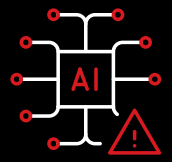


TrendAI

Trend Vision One™ AI Security

AI Risk Insights

Visibility and guidance to transform AI risks into secure, compliant adoption.



AI Risk Insights

AI Security Blueprint

AI Secure Access

Governance and threat protection for third-party generative AI use.

AI Secure Access

Deep Fake Detection

Detect of deep fake risk on server and endpoint.

Endpoint and Server Protection

AI Application Security

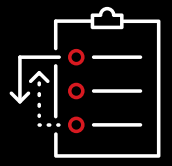
Secures AI apps from development to production using proactive testing & real-time threat prevention.

AI Scanner

AI Guard

Agentic SIEM

Turn complex logs into actionable insights through automated, natural language workflows.

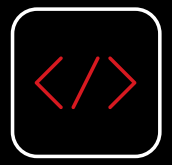


Agentic SIEM

XDR for Cloud

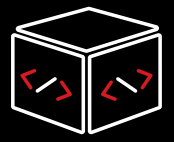
Cloud Risk Management

Safeguards code, data, containers, APIs, and AI workloads by detecting vulnerabilities, controlling exposures, supporting compliance, and enabling rapid threat response.

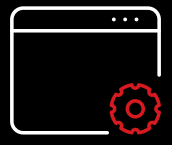


Code Security

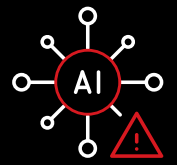
DSPM



Container Security



API Risk Visibility



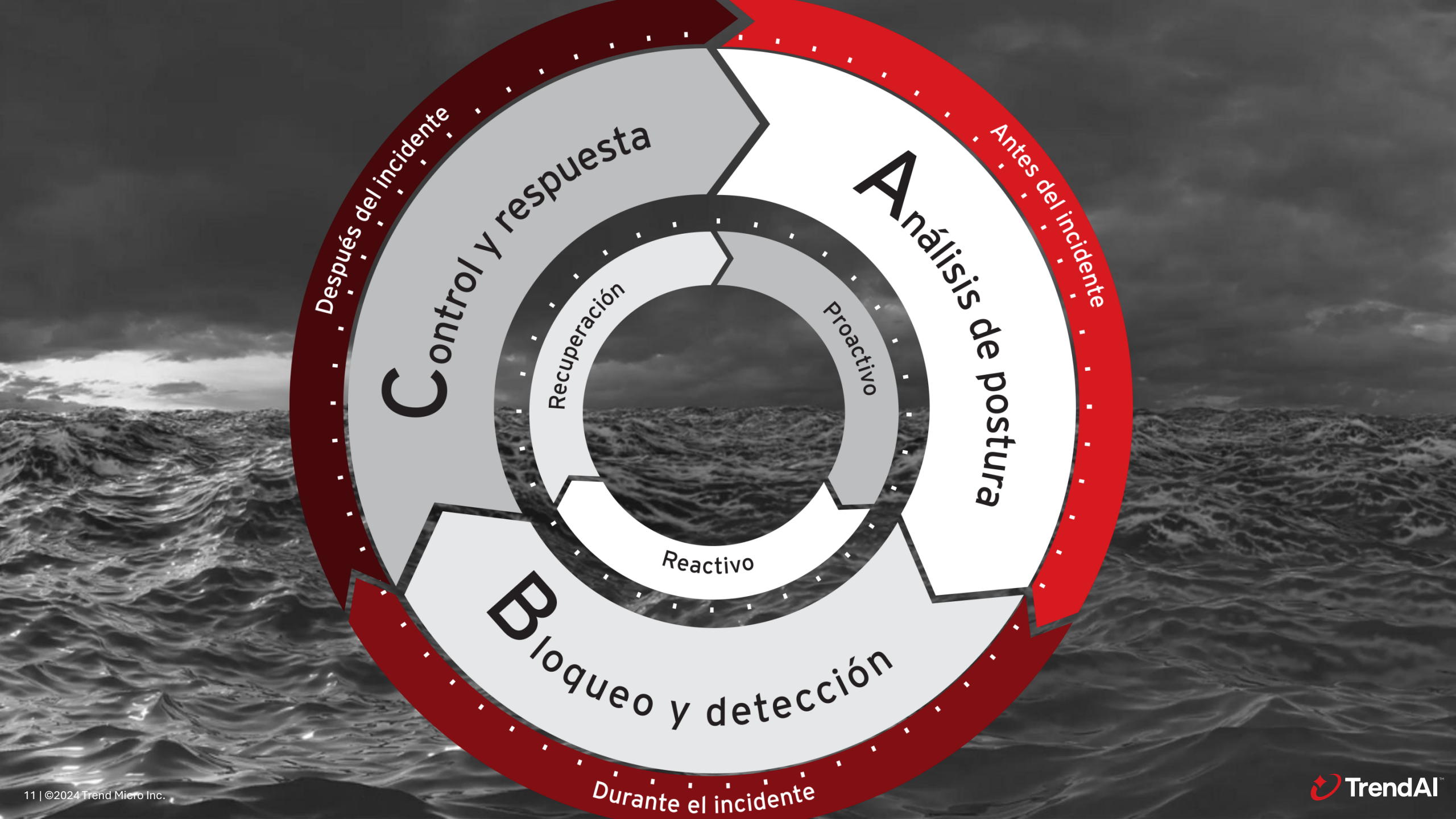
AI - SPM

NDR

Turn complex flows into actionable insights through automated, network detection and response.

XDR for Network

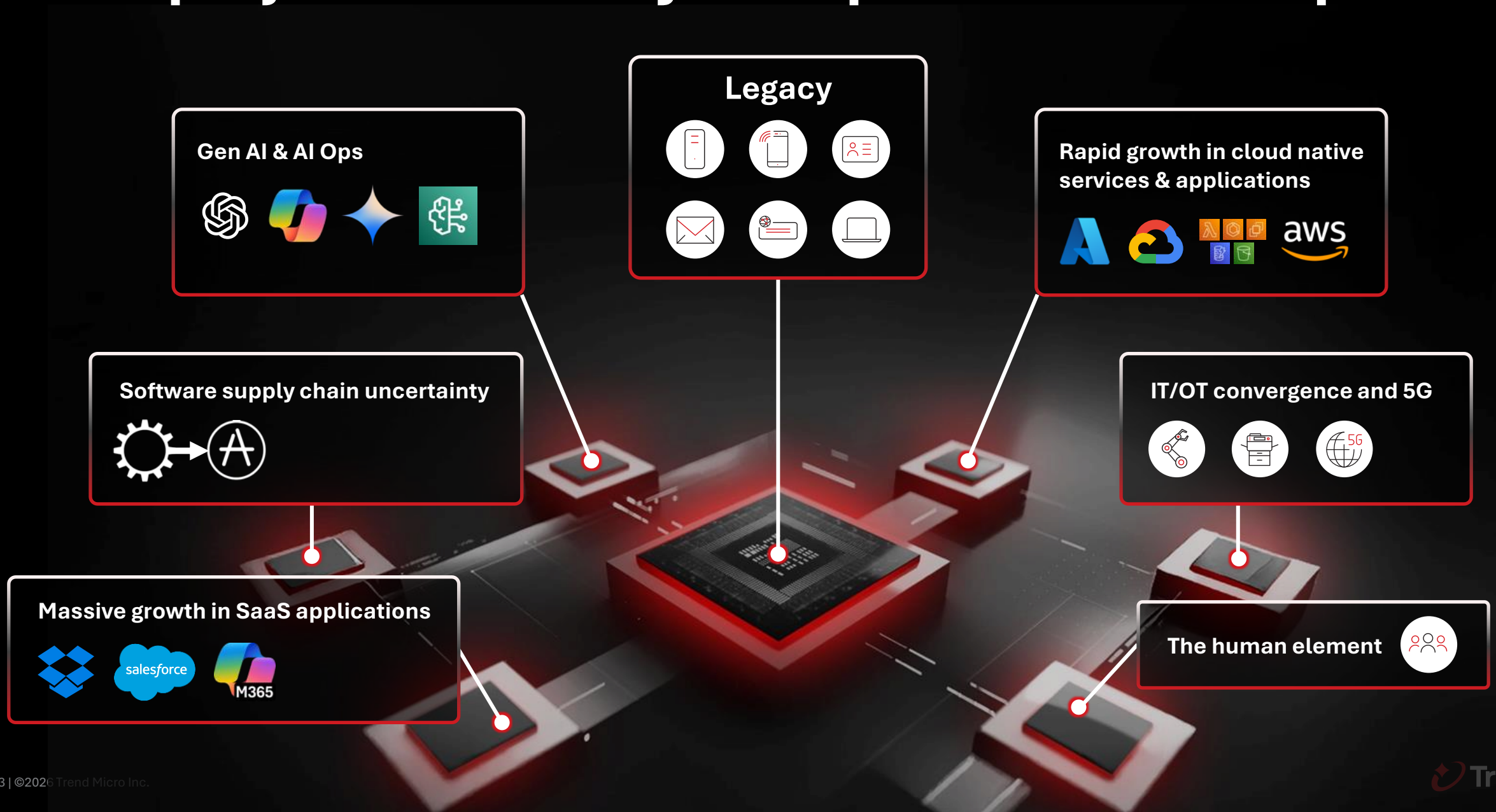
TrendAI™


Después del incidente
Antes del incidente
Durante el incidente
Control y respuesta
Análisis de postura
Bloqueo y detección
Recuperación
Proactivo
Reactivo

TrendAI

Análisis del Ciber Riesgo

Complejidad con Mayor Superficie de Ataque



TrendAI™

¿Qué es el Riesgo?

RIESGO = **PROBABILIDAD** x **IMPACTO**

impacto = g (criticidad de negocio)

Attack Path

=

¿Cuál es el punto más vulnerable?

Punto de Entrada+ Activos afectados

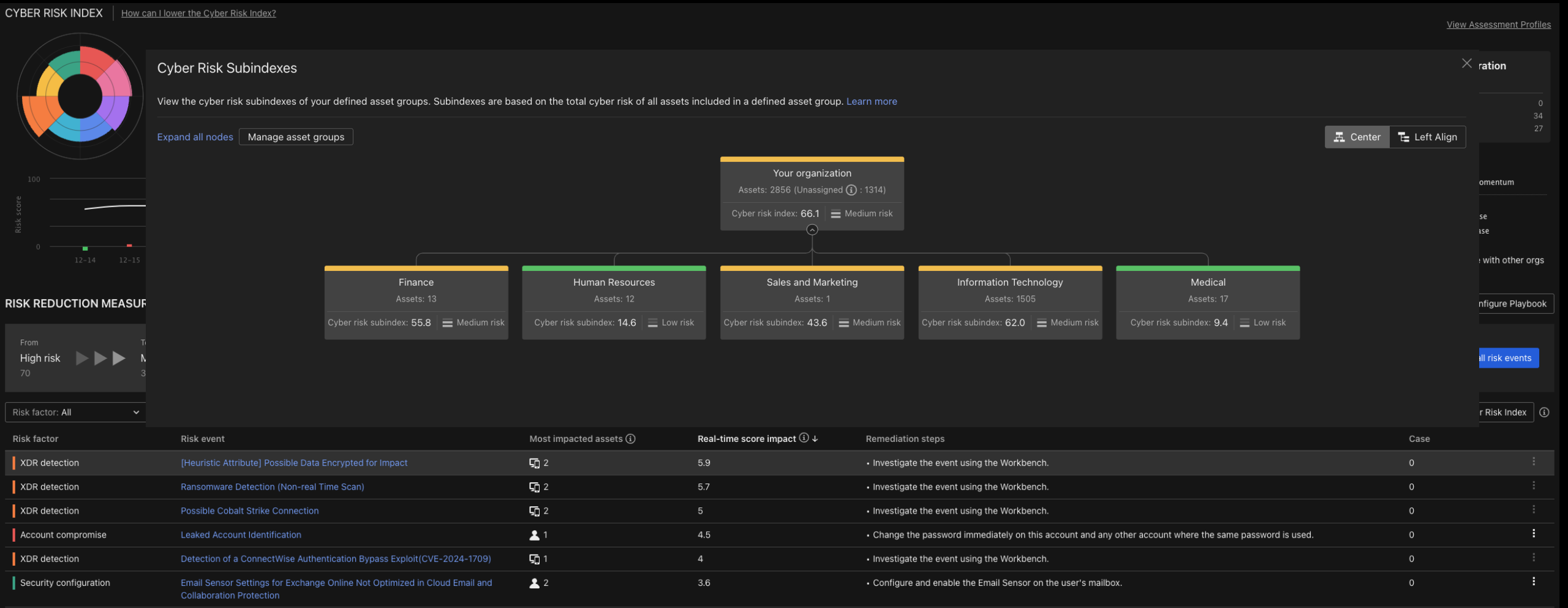
x

Joyas de la Corona

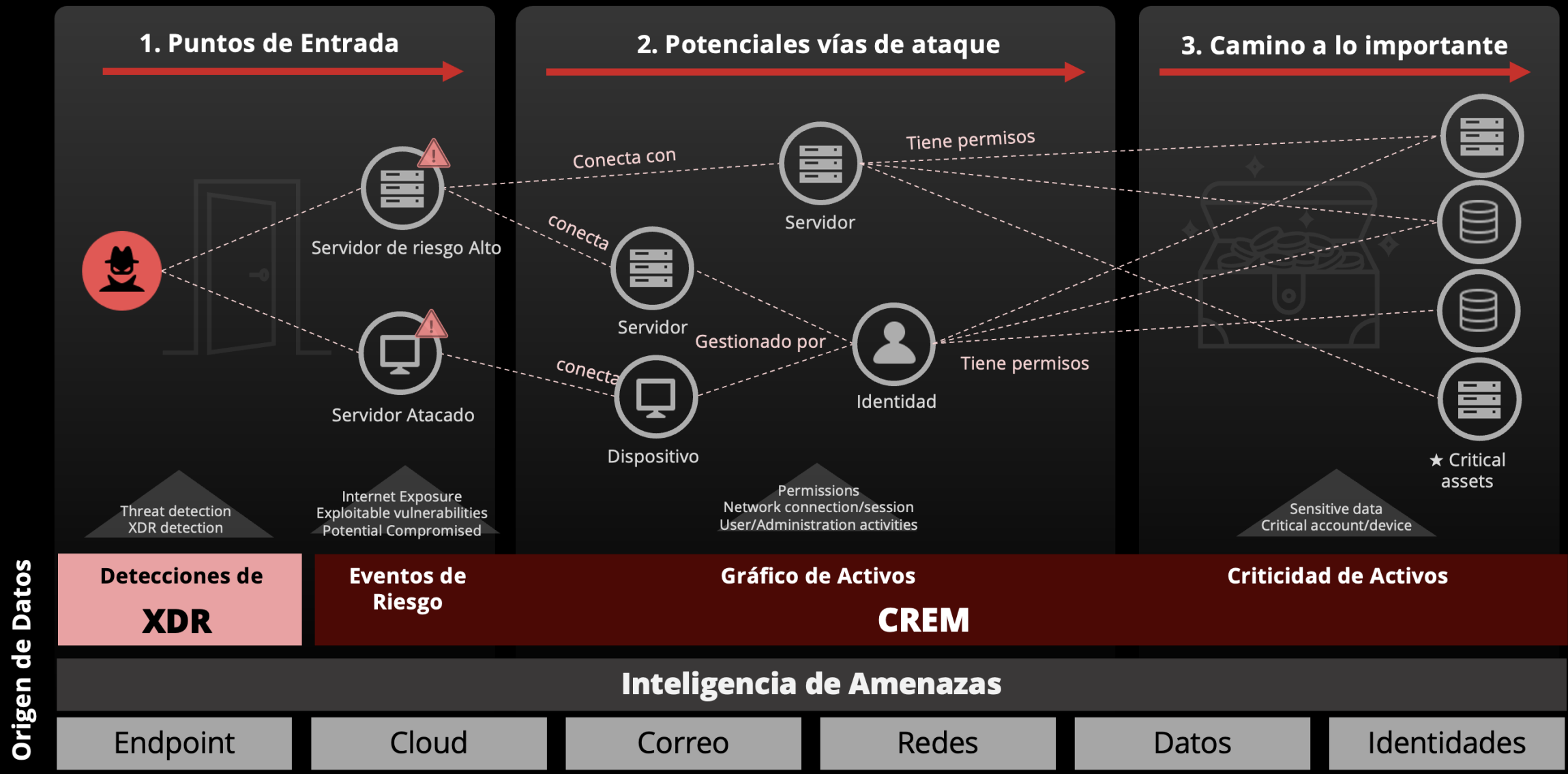
¿Cuál es el valor de este active?
¿Cuántos activos puedo comprometer?

Probabilidad = f (vulnerabilidades, exposición, amenazas, controles de mitigación)

Índice de Riesgo



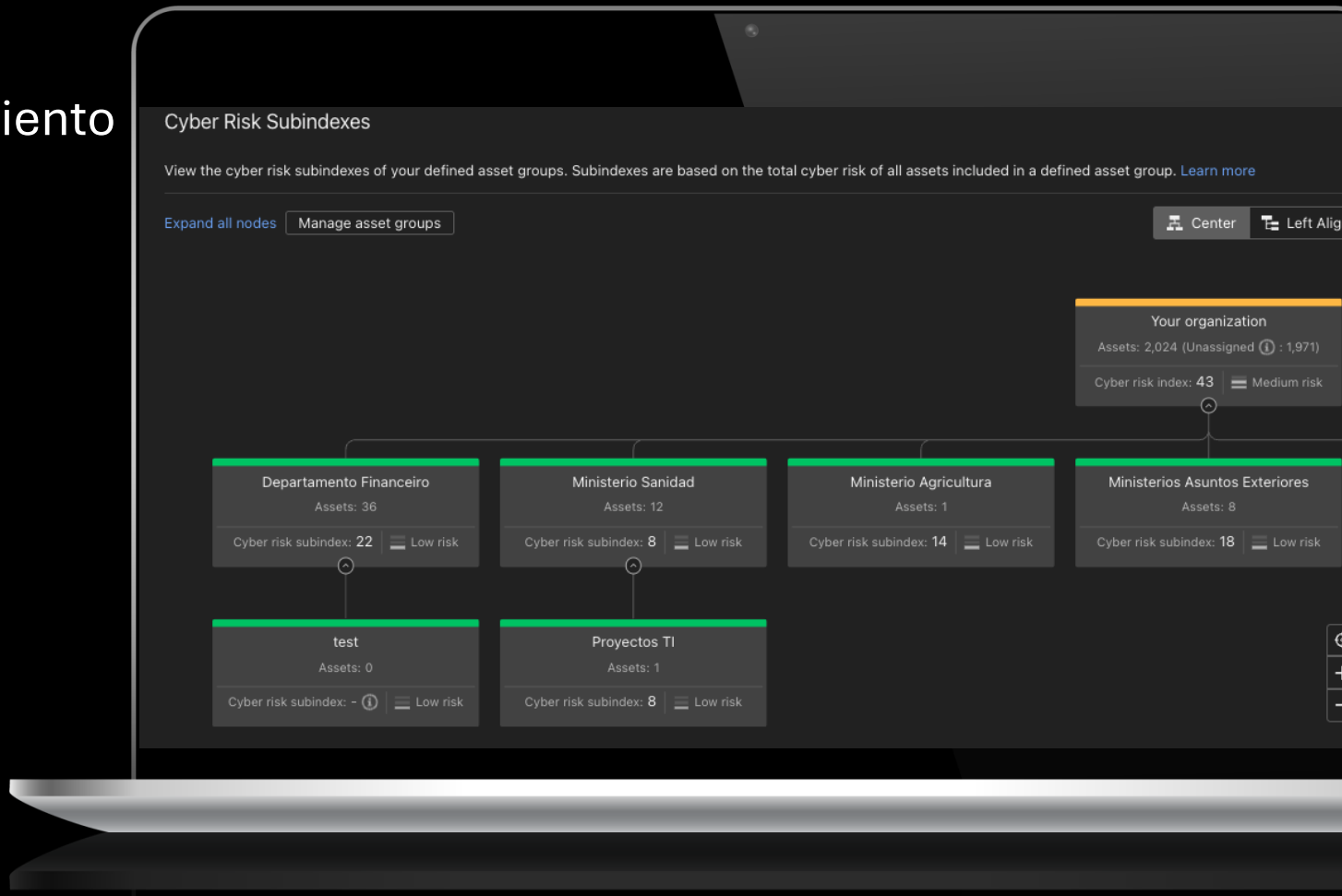
Identificar potenciales caminos de Ataque

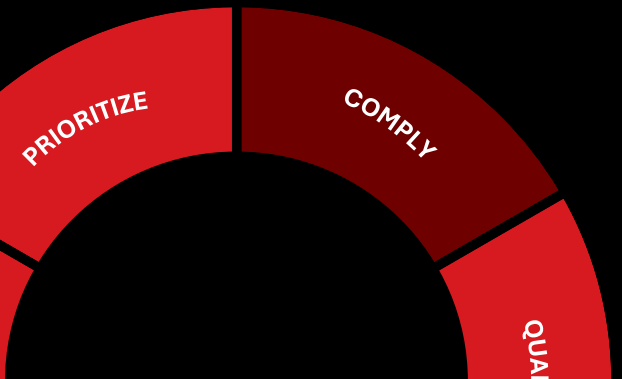


TrendAI

Cumplimiento

- NIS2, DORA, ENS,etc
- Priorizar esfuerzos de cumplomiento
- Generar informes de Auditoría.





TrendAI™


Después del incidente
Control y respuesta
Antes del incidente
Análisis de postura
Recuperación
Proactivo
Reactivo
Bloqueo y detección
Durante el incidente

TrendAI™