CAZADORES DIGITALES
Vulnerabilidades y Ataques bajo la Lupa de la IA

TICAL 2025 I San José, Costa Rica

Jorge Mora-Flores
LATAM Sales Account Manager I DeepSeas

LA NUEVA REALIDAD DIGITAL

MAGNITUD DE LA AMENAZA:

- Ataques cibernéticos:
  - +58% en dos años
  - Promedio: 1,984 ataques semanales/organización (2025)

- TRANSFORMACIÓN POR IA:
  - 47% de organizaciones: GenAI es su mayor preocupación
  - 42% sufrieron ataque exitoso de ingeniería social (último año)
  - 223% aumento en comercio de herramientas deepfake (2023-2024)

- PARADOJA:
  - 93% organizaciones reportan resiliencia cibernética insuficiente

Referencias: World Economic Forum. (2025). *Global Cybersecurity Outlook 2025*. https://www.weforum.org/publications/global-cybersecurity-outlook-2025/
Unión Internacional de Telecomunicaciones. (2024). *Riesgos de ciberseguridad e IA*. ITU.

IA: ARMA Y DEFENSA

Referencias: World Economic Forum. (2025). *Artificial Intelligence and Cybersecurity: Balancing Risks and Rewards*. https://reports.weforum.org/docs/WEF_Artificial_Intelligence_and_Cybersecurity_Balancing_Risks_and_Rewards_2025.pdf

NUEVAS FRONTERAS DEL ATAQUE

Nuevas Fronteras del Ataque Cibernético

RAG Poisoning
- Contaminación de Bases de Conocimiento
- Información Empresarial Comprometida

Inyección de Prompts
- Instrucciones Ocultas
- Manipulación de Asistentes IA

IA Agéntica Weaponizada
- Sistemas IA Autónomos
- Ciberataques Complejos

Ataques Multimodales
- Imágenes Maliciosas
- Audio Malicioso
- Calendarios Maliciosos

Referencias: Kaspersky. (2025). *Formas de vulneración a los LLM en 2025*. https://www.kaspersky.es/blog/new-llm-attack-vectors-2025/31427/

CASO: EXTORSIÓN MASIVA CON IA AGÉNTICA

IA identifica organizaciones objetivo

Reconocimiento automatizado con Claude Code

IA determina montos de rescate

Análisis de datos financieros para cálculo de rescates

IA exige pagos

Rescates de hasta $500,000 USD

Decisiones tácticas y estratégicas tomadas por IA

IA planifica ataque y rescate

Notas de rescate psicológicamente diseñadas

IA elabora mensajes persuasivos

TICAL 2025

Referencias: Anthropic. (2025). *Detecting and countering misuse of AI: August 2025*. https://www.anthropic.com/news/detecting-countering-misuse-aug-2025

CASO: EXTORSIÓN MASIVA CON IA AGÉNTICA

Translate

[SELL] Fully Private FUD Ransomware + Crypter Sources For Windows 10/11.

Found 1 message

Go to post

January 17th 2025, 11:29 pm

TssXX25
SELLING
=======

- Sources for ransomware DLL & exe (Lite + Full Version), Decrypter, Key Generator and DLL launcher, Fresh RSA keypair and 1x unique stub. Limited to 10 Copies. $400 USD.

- Source for Ransomware-as-a-Service (RaaS) kit. Secure PHP console & C&C tools. $800 USD.

- Source for Windows 10/11 FUD Crypter designed for native Windows binaries, written in C++. Uses AES-256 for payload encryption. Currently limited to 10 builds, each build includes 1x private stub. For educational and research use only. $1200USD.

Please get in touch and find out more via our website:

- techscckl72ibnfg2ksj5aqlanwgzw32asr6ml37aojnyw4nardojyid.onion -

TICAL 2025

Referencias: Anthropic. (2025). *Detecting and countering misuse of AI: August 2025*. https://www.anthropic.com/news/detecting-countering-misuse-aug-2025

EXPOSICIÓN DE INFRAESTRUCTURA

CVE-2025-32711 (Microsoft 365 Copilot):

- Gravedad: CVSS 9.3 (CRÍTICA)
- Inyección de comandos de IA
- Divulgación de información sensible

SERVIDORES EXPUESTOS:

- Ollama: 10,000+ servidores sin autenticación en Internet
- Chroma DB: 200+ servidores completamente desprotegidos
- Bases de datos vectoriales: inyección de información maliciosa

RIESGO: Infraestructura de IA académica potencialmente vulnerable

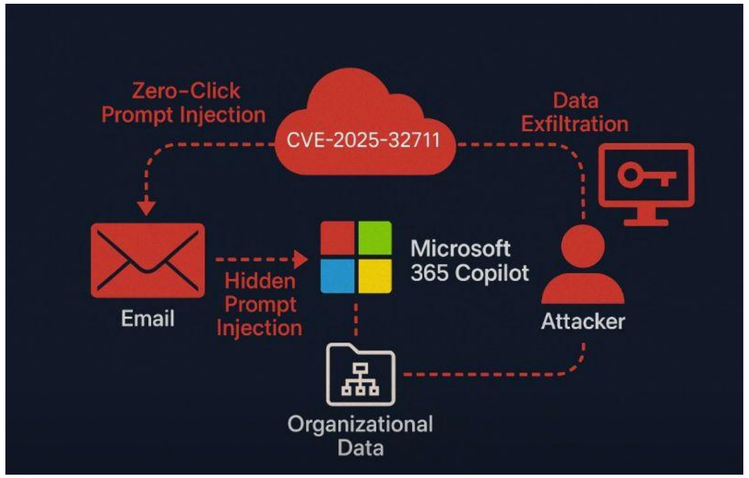


Imagen: https://www.linkedin.com/posts/marialuisaredondo_cve-2025-32711-is-a-turning-point-last-activity-7344644567255691265-8Ju8/

IA-BIOSEGURIDAD

CUANDO LA IA SALE DEL CIBERESPACIO

Entrenar Sistema IA

El sistema IA es entrenado para identificar moléculas tóxicas.

Identificar Compuestos Tóxicos

El sistema IA identifica 40,000 compuestos tóxicos.

Mejorar Virus con IA

La IA se utiliza para mejorar virus.

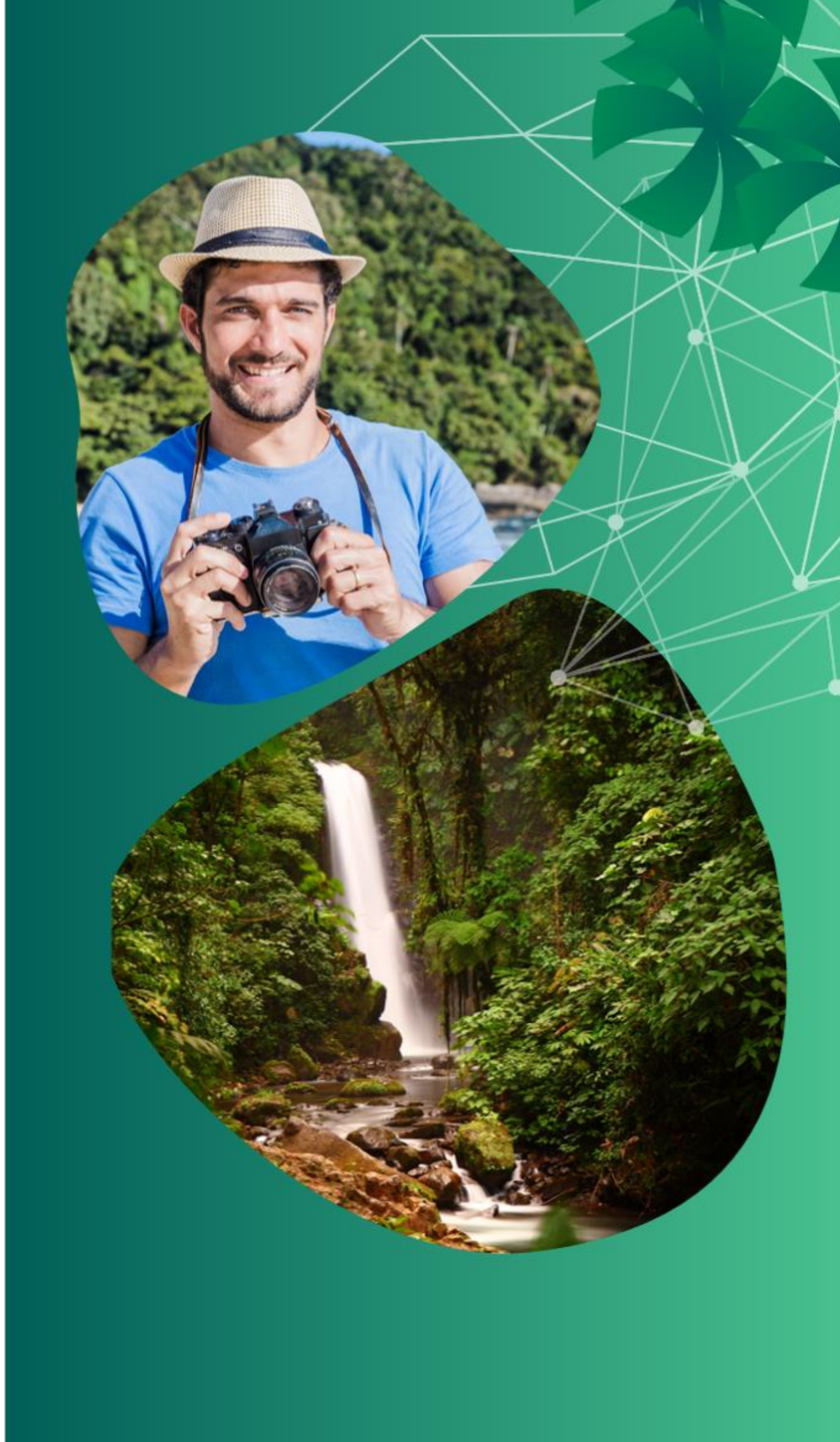
Hackear con ADN Malicioso

El ADN malicioso se utiliza para hackear sistemas.

Acelerar Desarrollo de Bioarmas

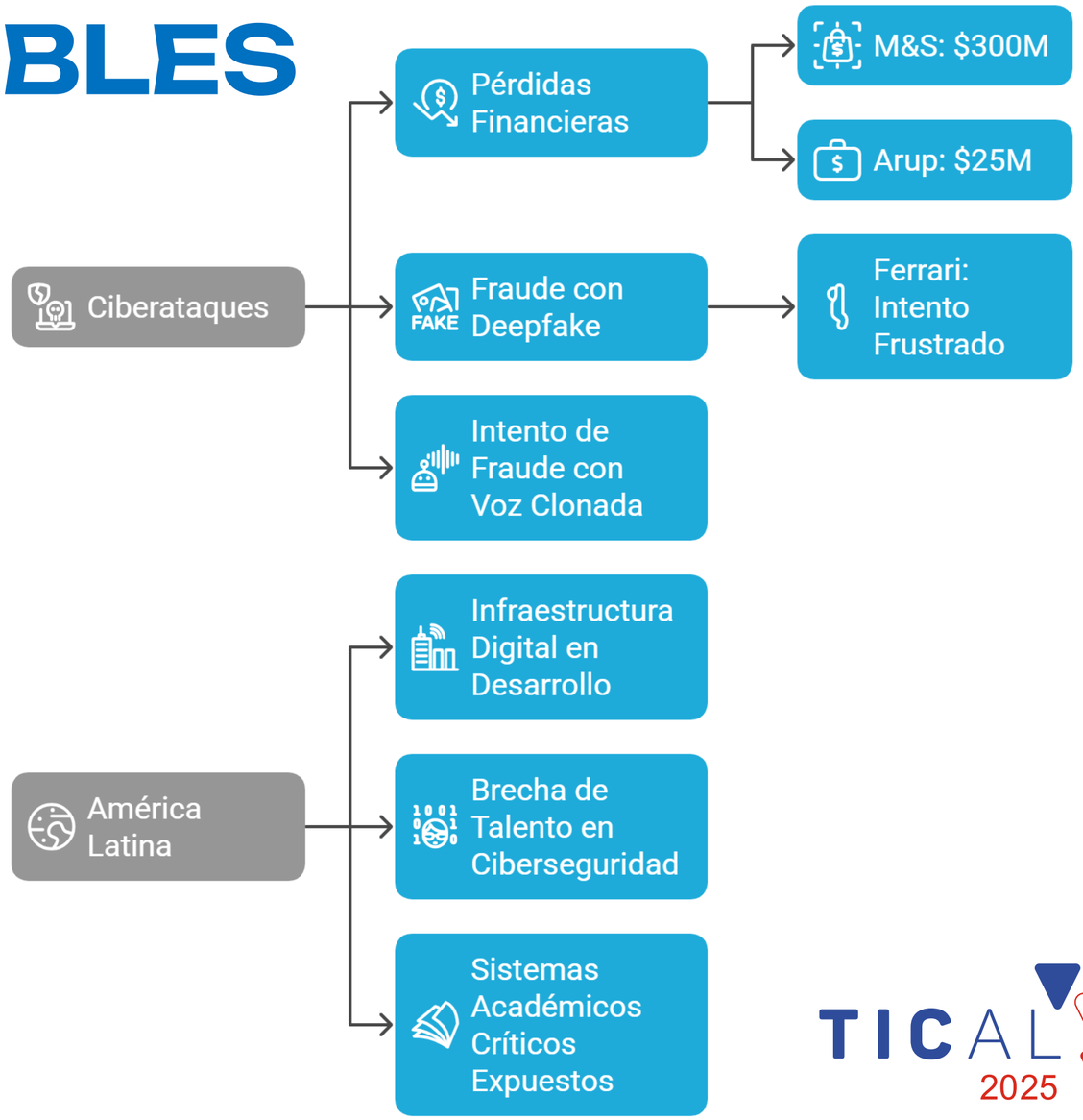
La IA acelera el desarrollo de bioarmas.

TICAL 2025



Referencias: Unión Internacional de Telecomunicaciones. (2024). *Riesgos de ciberseguridad e IA*. ITU.
World Economic Forum. (2025). *Artificial Intelligence and Cybersecurity: Balancing Risks and Rewards*.
https://reports.weforum.org/docs/WEF_Artificial_Intelligence_and_Cybersecurity_Balancing_Risks_and_Rewards_2025.pdf

CONSECUENCIAS TANGIBLES



TICAL 2025

Referencias: World Economic Forum. (2025). *Cybersecurity awareness: AI threats and cybercrime in 2025*. https://www.weforum.org/stories/2025/09/cybersecurity-awareness-month-cybercrime-ai-threats-2025/

DEMOCRATIZACIÓN DEL CIBERCRIMEN: Cibercrimen para todos(as)

Cibercrimen tradicional

Años de entrenamiento especializado

Capacidades de IA

Capacidades avanzadas con IA

Ransomware-as-a-Service

Venta de ransomware funcional

Cibercrimen moderno

Barrera técnica desaparecida

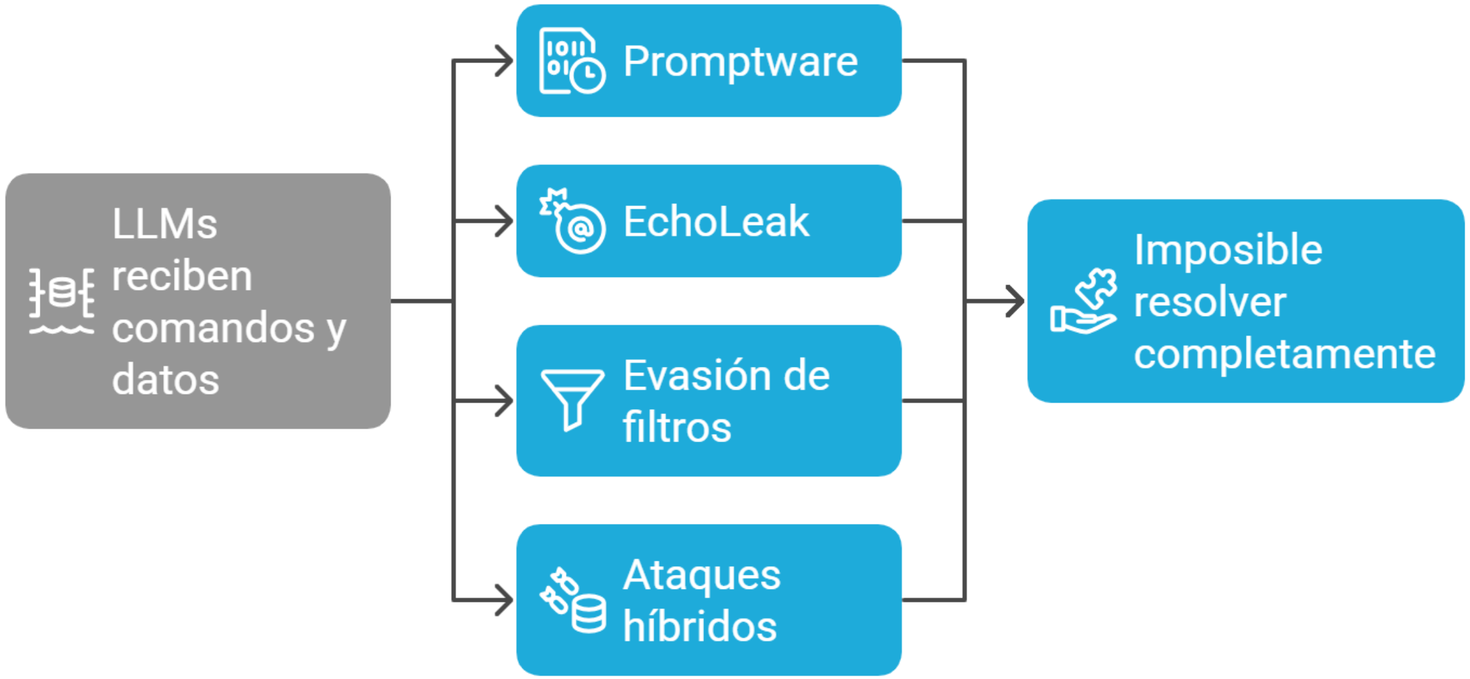
Referencias: Anthropic. (2025). *Detecting and countering misuse of AI: August 2025*. https://www.anthropic.com/news/detecting-countering-misuse-aug-2025

CUANDO NO PUEDES CONFIAR EN TUS OJOS

Referencias: World Economic Forum. (2025). *Cybersecurity awareness: AI threats and cybercrime in 2025*. https://www.weforum.org/stories/2025/09/cybersecurity-awareness-month-cybercrime-ai-threats-2025/
Unión Internacional de Telecomunicaciones. (2024). *Riesgos de ciberseguridad e IA*. ITU.

TICAL 2025

EL TALÓN DE AQUILES DE LA IA



- Promptware: Instrucciones ocultas en calendarios
- EchoLeak: Exfiltración vía enlaces de imágenes
- Evasión de filtros mediante codificación visual
- Ataques híbridos (IA + XSS/CSRF)

Referencias: Kaspersky. (2025). *Formas de vulneración a los LLM en 2025*. https://www.kaspersky.es/blog/new-llm-attack-vectors-2025/31427/

LOS TRES PILARES DE DEFENSA

Framework de MIT para defensa contra ataques potenciados por IA

Higiene de seguridad automatizada

Código auto-reparable, sistemas auto-parchado, arquitectura zero-trust y gestión continua de superficie de ataque.

1

Sistemas defensivos autónomos

Moving-target defense, tácticas de engaño automatizadas y ML para identificación de amenazas.

2

Supervisión aumentada

Análisis de riesgo automatizado, inteligencia en tiempo real e insights predictivos para ejecutivos.

3

TICAL 2025

Referencias: Siegel, M. (2025). AI cyberattacks and three pillars for defense. *MIT Sloan Management Review*. https://mitsloan.mit.edu/ideas-made-to-matter/ai-cyberattacks-three-pillars-defense

COLABORACIÓN REGIONAL - BELLA II

Necesidad LAC: Public-Private Partnerships
- Desarrollo de Talento
- Inversión en Educación

Ecosistema Digital LAC-Europa
- Conexión Digital
- Colaboración Regional

Éxito Reciente: INTERPOL + AFRIPOL
- Arrestos
- Recuperación de Fondos

Proyecto BELLA II + REDCLARA

Redes Académicas Interconectadas
- Investigación Conjunta
- Intercambio de Conocimiento

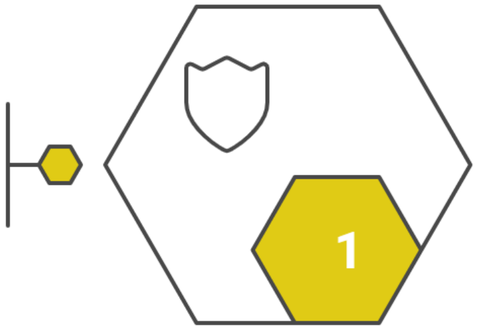
Inteligencia de Amenazas Compartida
- Detección de Amenazas
- Respuesta Coordinada

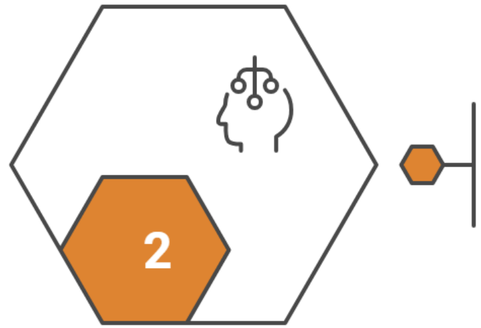
TICAL 2025

RECOMENDACIONES PARA INSTITUCIONES ACADÉMICAS

Gobernanza

Modelado detallado de amenazas, RBAC para herramientas de IA y confirmación humana en acciones críticas.



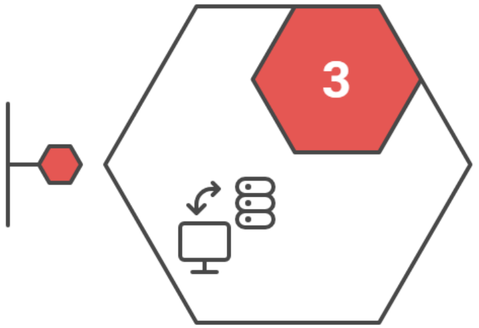


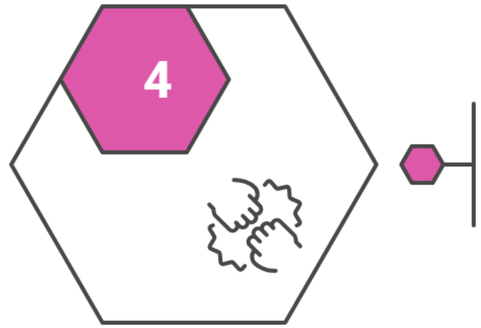
Capacitación

Formación en seguridad de IA para todo el personal, programas especializados para CISOs y "red teaming" con IA adversarial.

Tecnología

Logging robusto (integración SIEM + MXDR), mínimo privilegio por defecto y auditorías de seguridad regulares.





Colaboración

Compartir indicadores de compromiso y participación en CSIRTs regionales (REUNA, RNP, CEDIA, RENATA).

RBAC: Role-Based Access Control
MXDR: Managed Extended Detection & Response

DE REACTIVO A PREDICTIVO

"¿Podemos romper la asimetría de la guerra cibernética? El atacante necesita un punto de entrada; el defensor debe proteger todos"

- Michael Siegel, MIT

CONCLUSIÓN

CAZADORES Y CAZADOS

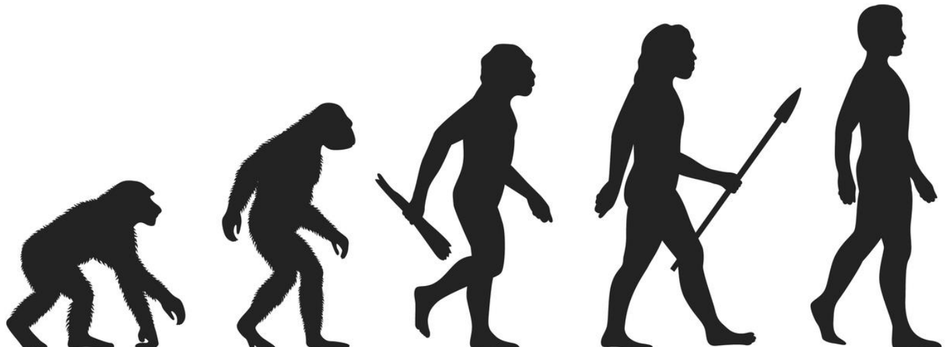
REALIDAD ACTUAL:

- La IA ha democratizado el cibercrimen
- Los ataques son más sofisticados y frecuentes
- Solo 14% de organizaciones tiene talento adecuado (WEF, 2025)

NUESTRA RESPUESTA:

- Adopción de marcos defensivos multicapa
- Inversión en capacitación continua
- Colaboración regional fortalecida
- Supervisión humana siempre presente

"El atacante evoluciona; como defensores también debemos hacerlo"



TICAL 2025

"66% DE LAS ORGANIZACIONES ANTICIPAN QUE LA IA TENDRÁ EL MAYOR IMPACTO EN CIBERSEGURIDAD.

SOLO EL 37% TIENEN PROCESOS PARA EVALUAR SU SEGURIDAD ANTES DE DESPLEGARLA.

LA BRECHA ENTRE ANTICIPACIÓN Y PREPARACIÓN ES LA VULNERABILIDAD MÁS PELIGROSA."

World Economic Forum, 2025

TICAL 2025

¡GRACIAS!
THANKS!

¿Alguna pregunta? Any questions?

Jorge Mora-Flores
LATAM Sales Account Manager DeepSeas

jorge.mora@deepseas.com
jorge@moraflores.com

DEEP seas