IDMX 2026 · INFLEXIÓN DIGITAL MÉXICO — MÉRIDA, YUCATÁN

El punto de inflexión de la IA en el SOC y en el proceso DFIR

Amenazas autónomas, escasez de talento y defensa colectiva: por qué 2026 es el año en que el reloj del atacante superó al del defensor.

Alberto Ávalos

CIO — Pryse Cyber CERT · FIRST México, Liaison

LinkedIn
/in/bettoavalos

Escanea para conectar

TLP:CLEAR · 24 de agosto de 2026

EL PUNTO DE INFLEXIÓN

Durante quince años el tiempo fue el presupuesto del defensor. Ese presupuesto se agotó.

EL RELOJ DEL ATACANTE

29 min

Tiempo promedio de breakout en 2025: del acceso inicial al primer movimiento lateral. 65 % más rápido que el año anterior. Récord observado: 27 segundos.

EL RELOJ DEL DEFENSOR

56 min

Tiempo promedio entre que una alerta se dispara y un analista humano la toca por primera vez. Todo el diseño del SOC, turnos, backlog, escalamientos, vive dentro de ese interval de tiempo

El reloj del atacante corre al doble de velocidad que el del defensor. Ese es el punto de inflexión: no la llegada de la IA al SOC, sino la inversión de la relación temporal que sostenía el modelo operativo previo.

Panorama de amenazas 2025–2026 · Métricas de operación de SOC, 2026.

RUTA DE LA CHARLA

Cuatro movimientos, treinta minutos

01

El problema

Qué cambió en la relación temporal entre ataque y defensa.

02

Las cifras

México y América Latina: volumen, ransomware, identidad y sector público.

03

La IA ejecuta

Tres campañas documentadas donde el agente hizo el trabajo táctico.

04

Las soluciones

SOC, DFIR, ensayo de la respuesta y defensa colectiva.

Todas las cifras están atribuidas a su fuente primaria y fechadas. Ventana de datos: noviembre de 2025 – agosto de 2026.

NOTA METODOLÓGICA

Antes de las cifras: cómo hay que leerlas

Intentos no son incidentes

Las cifras de “miles de millones de intentos” cuentan eventos bloqueados: escaneos, sondeos, autenticaciones fallidas. Un solo botnet genera millones. Miden presión ambiental, no compromisos exitosos.

Fabricante ≠ medición independiente

Cuando una cifra proviene de un proveedor con interés comercial en el resultado, se señala de forma explícita. No representan una linea base

Lo que sí cuenta víctimas

Las víctimas publicadas en sitios de extorsión son un conteo verificable de compromisos exitosos. Ese es el indicador al que hay que hacer caso.

Presentar intentos como sinónimo de ataques exitosos degrada la credibilidad de la postura de Seguridad.

EL PROBLEMA · VOLUMEN

México dentro de una región bajo presión industrial

58.1 billones

Intentos de ciberataque registrados en México durante 2025.

843.3 billones

Intentos acumulados en toda América Latina en el mismo periodo.

30 billones

Escaneos activos contra objetivos mexicanos en 2025.

≈36,000

Sondeos por segundo en el primer semestre de 2025.

2.º y 3.er lugar de América Latina

Detrás de Brasil y junto a Colombia. Los casos en empresas mexicanas crecieron 38 % interanual. No es actividad artesanal: es reutilización de herramientas, compra de accesos, botnets y agentes de IA.

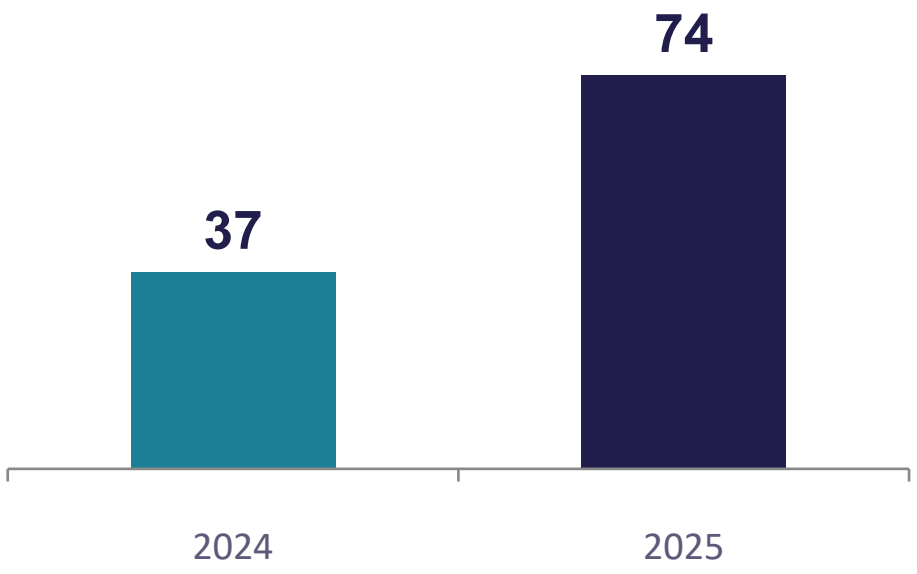
60 % con incidente · 27 % con protección

Seis de cada diez empresas mexicanas reportaron un incidente de seguridad en el último año. Tres de cada diez declaran contar con servicios especializados de protección de activos.

Fuente: FortiGuard Labs, Reporte Global sobre el Panorama de Amenazas 2026 (datos 2025) y Fortinet 1S 2025.

El indicador que sí cuenta víctimas, y se duplicó

Empresas y Organizaciones mexicanas publicadas en sitios de extorsión



2.º

país más atacado por ransomware en América Latina al cierre de 2025.

16 → 11

escaló cinco lugares en el ranking global de víctimas publicadas.

781 / día

≈237,000 intentos de ransomware contra objetivos mexicanos (ago 2024 – jun 2025).

+389 %

víctimas confirmadas a escala global: 7,831 en 2025 frente a ~1,600 en 2024.

Sectores más golpeados: manufactura (1,284) · servicios financieros (824) · comercio (682). Grupos activos en México: Qilin, CL0P, Kazu y LockBit.

Fuentes: FortiRecon 2026; Plan Nacional de Ciberseguridad (155 víctimas mexicanas difundidas entre 2019 y septiembre de 2025).

BLOQUE 01 DE 04

El problema en números

01

Cuatro mediciones que describen el estado real de la superficie de ataque en México.

- Identidad
- Costo de las brechas
- Talento
- Sector público

EL PROBLEMA · IDENTIDAD

Los adversarios ya no irrumpen: inician sesión

82 %

de las detecciones de 2025 fueron malware-free.

CrowdStrike

83 %

de los incidentes en nube y SaaS entraron por identidad.

Google Cloud H1 2026

4,620 M

registros de infostealer negociados en la dark web (+79 %).

FortiRecon 2025

× 12.5

crecieron los ataques de bots impulsados por IA en 2025.

Thales Bad Bot 2026

Implicación para el diseño del SOC

Si el 82 % de las detecciones son malware-free y el 83 % de los incidentes en nube entran por identidad, un SOC que se basa en firmas de antivirus y correlación de eventos de endpoint está mirando al lugar equivocado.

Los bots representaron más del 53 % del tráfico total de internet; 27 % de esos ataques se dirigió a APIs y 46 % de los incidentes de toma de cuentas ocurrió en servicios financieros.

LAS CIFRAS · COSTO

Lo que cuesta una brecha, y lo que ahorra automatizar

– USD 1.93 M

menos por brecha en las organizaciones que usan IA y automatización en sus operaciones de seguridad.

80 días

más corto el ciclo de vida de la brecha con uso extensivo de IA.

USD 4.99 M

promedio global: récord histórico, +12 % interanual.

Fuente: IBM Cost of a Data Breach Report 2026 y 2025 (Ponemon Institute). Salud: sector más caro por decimotercer año consecutivo.

El cuello de botella no es el presupuesto ($): es la gente

77,000

perfiles que hoy simplemente no existen en el mercado laboral mexicano.

1.2 %

de la brecha cubre la generación 2025 completa

129

programas académicos de ciberseguridad, frente a 547 de software

3 %

de los profesionales tiene el nivel de inglés que exigen las vacantes

77 / 23

distribución por género en la matrícula (hombres / mujeres)

Fuente: IQSEC y Select, "Radiografía del talento digital", mayo de 2026; matrícula de educación superior y ENOE-INEGI. Salario promedio del especialista: MXN 51,386 mensuales.

EL PROBLEMA · SECTOR PÚBLICO

El diagnóstico no es alentador

7 de cada 10

dependencias federales presentan vulnerabilidades de alto riesgo.

- Seis dependencias fueron víctimas de incidentes entre 2022 y 2025 — y cinco de ellas recortaron sus recursos de informática y seguridad digital después del incidente.
- El presupuesto de informática de la SICT cayó más de 85 %: de ~1,800 millones de pesos a poco más de 260 millones, pese a haber sido blanco de ataques que comprometieron datos críticos.

Plan Nacional de Ciberseguridad 2025–2030

Contempla un Centro Nacional de Operaciones de Ciberseguridad (CNSOC) y un CSIRT con alerta temprana. La crítica técnica más consistente: no hay asignación presupuestal específica.

3,852 mdp

ATDT para el ejercicio 2026: +20 % real, pero no es un monto exclusivo para ciberseguridad.

>3,000 mdd

CISA (EE. UU.) en 2025, destinados exclusivamente a la función de ciberseguridad.

El mercado mexicano de ciberseguridad se estima entre 3,050 y 3,120 millones de dólares para 2026 (CAGR 9.2 %–11.4 % a 2031). El cuello de botella no es la inversión: es la capacidad de absorberla.

LA IA COMO ARMA

¿Cuánto de esto ya es inteligencia artificial?

No existe un contador universal: el uso adversarial de IA solo es visible cuando deja rastro. Pero hay mediciones convergentes desde ángulos distintos, y todas apuntan en la misma dirección.

1 de cada 4

brechas maliciosas ya fue habilitada por IA. Crecimiento interanual: +56 %.

IBM / Ponemon, jul 2026

+89 %

aumentaron las operaciones de adversarios que emplean IA.

CrowdStrike GTR 2026

78 %

de las organizaciones tuvo ataques con IA confirmados (45 %) o sospechosos (33 %).

SANS AI Survey 2026

94 %

de los líderes ven la IA como el mayor motor de cambio del panorama.

WEF Outlook 2026

Qué tipo de ataque con IA se ve en la práctica

Suplantación por deepfake (cerca de la mitad) · malware habilitado por IA · phishing generado con IA.

Y del lado defensivo

87 % reporta más vulnerabilidades relacionadas con IA, la categoría de crecimiento más rápido, y el uso de IA en red team pasó de 33 % a 61 % en un año.

BLOQUE 02 DE 04

Casos documentados

02

Tres operaciones reales, con víctimas reales e indicadores de compromiso publicados.

GTG-1002

Campaña contra firewalls

Malware que consulta un LLM

CASO 1 · CIBERESPIONAJE

GTG-1002: la primera campaña orquestada por IA

80–90 %

de las operaciones tácticas ejecutadas por la IA sin intervención humana. Solo 4 a 6 puntos de decisión estratégica por campaña.

Atribución	Grupo patrocinado por el Estado chino (alta confianza)
Detección / publicación	Mediados de septiembre de 2025 · publicado 13–14 de noviembre de 2025
Objetivos	≈30 entidades: tecnología, banca, manufactura química y gobierno
Arquitectura	Claude Code como orquestador, vía MCP, sobre herramientas de pentesting open source
Herramientas	Nmap, Metasploit, sqlmap. Sin 0-days. Sin malware propietario.

El ciclo completo, sin humano

Reconocimiento · descubrimiento de vulnerabilidades · generación y ejecución de exploits · recolección de credenciales · movimiento lateral · análisis · exfiltración · y hasta la documentación de la propia operación.

No explotaron el modelo: lo convencieron

Ingeniería social sobre el propio modelo: le hicieron creer que trabajaba para una firma de seguridad en pruebas autorizadas, y fragmentaron las tareas para que cada solicitud pareciera legítima. La superficie de ataque de un agente incluye su comprensión de la situación.

Límite observado: el modelo alucinó. Fabricó credenciales que no funcionaban y afirmó haber extraído datos que no había extraído. Precedente: "vibe hacking", junio 2025 — 17 organizaciones extorsionadas en un mes.

CASO 2 · CIBERCRIMEN FINANCIERO

La IA como línea de ensamblaje

600+

dispositivos Firewall comprometidos

55+

países afectados en cinco continentes

5 semanas

de actividad: 11 de enero – 18 de febrero de 2026

CERO

vulnerabilidades explotadas. Ningún 0-day.

El vector real

Interfaces de administración expuestas a internet (puertos 443, 8443, 10443 y 4443) y credenciales débiles con autenticación de un solo factor. Después: configuración completa, credenciales SSL-VPN recuperables, políticas de firewall, topología y parámetros IPsec.

Qué hizo la IA

Planes de ataque paso a paso, herramientas propias de reconocimiento en Go y Python, parseo de configuraciones, triaje analítico. El código delata al modelo: comentarios redundantes.

"Una línea de ensamblaje de IA para el cibercrimen, que ayuda a trabajadores menos calificados a producir a escala." — Amazon Threat Intelligence, 20 de febrero de 2026

CASO 3 · AUTO-MODIFICACIÓN

Malware que consulta un LLM mientras se ejecuta

Google Threat Intelligence Group documentó en noviembre de 2025 la primera fase operativa del abuso de IA: familias que invocan un LLM durante su ejecución, no solo durante su desarrollo.

PROMPTSTEAL

OPERATIVO

Minero de datos en Python. Consulta Qwen2.5-Coder vía Hugging Face para generar comandos de Windows que recolectan información y documentos, y los ejecuta. Atribuido a APT28 en operaciones contra Ucrania.

PROMPTFLUX

EN DESARROLLO

Dropper en VBScript. Su módulo “Thinking Robot” pide a la API de Gemini código nuevo de evasión de antivirus y reescribe su código fuente completo cada hora.

QUIETVAULT

OBSERVADO

Ladrón de credenciales en JavaScript dirigido a tokens de GitHub y NPM; usa interfaces locales de IA para localizar y exfiltrar secretos.

FRUITSHELL / PROMPTLOCK

PoC / PÚBLICO

Reverse shell en PowerShell con prompts embebidos para evadir el análisis de seguridad basado en LLM. Y ransomware en Go que genera scripts Lua en tiempo de ejecución.

Fuente: GTIG AI Threat Tracker, Google Threat Intelligence Group, 5 de noviembre de 2025. La técnica: “auto-modificación just-in-time”.

BLOQUE 03 DE 04

SOC, DFIR e IA

03

Qué se movió de verdad cuando la inteligencia artificial entró en la operación.

Adopción

MTTA · MTTR · MTTC

Forense digital

La discusión sobre adoptar IA en el SOC ya terminó

50 % → 78 %

Uso activo de IA en la estrategia de ciberseguridad, en un solo año. Es el mayor movimiento interanual que la encuesta ha registrado en su historia.

SANS AI Survey 2026 — 536 practicantes y 57 líderes de seguridad

4 de cada 5

practicantes recurren a herramientas de IA o de aprendizaje automático en su trabajo diario. La IA dejó de ser un proyecto piloto en el SOC: es el conjunto de herramientas por defecto.

SANS SOC Survey, 10.ª edición, junio de 2026 — 444 practicantes y 69 líderes

La pregunta ya no es si adoptar IA en el SOC, sino cómo medir que realmente sirvió y quién audita la diferencia entre lo que promete el fabricante y lo que ocurre en su operación.

Qué se movió de verdad en MTTA, MTTR y MTTC

Evidencia sólida

USD 1.93 M	de ahorro por brecha con IA y automatización (IBM)
−80 días	de ciclo de vida de la brecha con uso extensivo de IA (IBM)
26–36 %	de mejora de MTTD y 22 % de MTTR en piloto con controles
−16 pp	de falsos positivos en ese mismo piloto
25–40 %	menos tiempo medio de triaje con ML sobre datos históricos (Forrester)

Rangos publicados por fabricantes

MTTA → 0	la IA inicia la investigación al dispararse la alerta
MTTD −40/−60 %	solo en plataformas de alto desempeño
MTTR −45/−80 %	el 80 % únicamente si se automatizan las cuatro fases
MTTC −75/−95 %	con playbooks preaprobados que contienen al confirmar
USD 25–50 → 2–5	costo por investigación cuando la IA enriquece y correlaciona

LAS SOLUCIONES · DFIR

En el SOC la IA ataca el volumen; en DFIR, el tiempo de reconstrucción

247 días

para identificar y contener una brecha: +6 días y el primer retroceso tras cinco años de mejora.

Cuando el equipo de DFIR llega, el adversario lleva meses dentro, ha movido lateralmente, ha exfiltrado y ha limpiado buena parte de sus huellas.

El límite del perito

Muchos modelos son cajas negras: producen conclusiones sin explicar cómo llegaron a ellas. Un tribunal puede rechazar evidencia que el investigador no sabe explicar.

Triaje de evidencia

La aplicación más madura: seleccionar los datos relevantes en lugar de procesarlos todos.

Análisis de malware

Clasificación por comportamiento en vez de por firma, y asistencia en ingeniería inversa.

Línea de tiempo

Correlación por grafos de eventos de Windows, MFT, registro, EDR y nube: de 45 minutos a segundos.

Contención

Aislamiento, revocación de credenciales y bloqueo en segundos, sin demora de aprobación.

Documentación

Informes listos para auditoría en tiempo real: la fase que más calendario consume.

BLOQUE 04 DE 04

Para ir cerrando

04

De la evidencia a la decisión: qué hacer con todo esto.

Preparación | Defensa colectiva | Qué hacer | Conclusiones

La mayoría compra más detección. Casi nadie ensaya la respuesta.

49 %

planea invertir en seguridad tras una brecha. El año previo era 63 %.

43 %

de esas organizaciones invertirá en detección de amenazas.

35 %

invertirá en probar su respuesta a incidentes. Es el rubro más bajo.

USD 2.66 millones

menos por brecha en las organizaciones que tenían un equipo de respuesta a incidentes y que además habían probado su plan, frente a las que no tenían ninguno de los dos.

Y el 37 % de las organizaciones no tiene plan de respuesta, pese a ser un requisito normativo.

Por qué no se prueban: ninguna razón es técnica

- El plan se trata como documento de cumplimiento, no como capacidad operativa.
- La dirección no comprende su propio papel en el plan ni el valor de ensayarlo.
- Presupuesto, falta de personal con experiencia y prioridades en competencia.
- La organización se concentra en prevenir incidentes en lugar de prepararse para ellos.

Fuente: IBM Cost of a Data Breach (2022 y 2025); HIPAA Journal.

LAS SOLUCIONES · DEFENSA COLECTIVA

La asimetría no es de presupuesto ni de tecnología: es de información compartida

FIRST FELLOW

Forum of Incident Response and Security Teams. Fundado en 1989 tras el gusano WANK y constituido en 1995 para coordinar la respuesta entre organizaciones durante eventos mayores.

810+

organizaciones miembro

110+

países en cinco continentes

≈50

CSIRT de gobierno en CSIRTAmericas (OEA-CICTE)

FIRST México

Punto de contacto local: un grupo de FIRST Liaisons acreditados que articula a los equipos de respuesta del país mediante colaboración institucional y entrenamiento especializado.

El servicio que México no está consumiendo

60 % de las empresas mexicanas reportó un incidente en el último año. Solo 27 % declara contar con servicios especializados de protección. Se compran herramientas; no se consume inteligencia de amenazas de forma estructurada.

QUÉ HACER

Tres cosas importantes

01

Formalizar la capacidad

Estructura mínima de un CSIRT: roles, guardias, canales y criterios de escalamiento documentados. Un plan que se ensaya, no un documento que pasa la auditoría.

02

Medir la madurez

Evaluarse contra SIM3 en lugar de contra la propia percepción. Sin una medición externa, la madurez que uno se atribuye es una opinión.

03

Integrarse a la comunidad

Consumir y aportar información de primera mano en una comunidad de confianza. Membresía full de FIRST como vía formal para acceder al modelo.

Convierte el incidente de una empresa en la advertencia temprana de otras diez.

CONCLUSIONES

Cinco mensajes para llevarse

1. La IA no creó atacantes más sofisticados. Creó más atacantes con alcance.
2. En México el cuello de botella es humano, y no se resuelve contratando.
3. La automatización del SOC es necesaria pero no suficiente y la mayoría la está haciendo mal.
4. Lo que falta no es tecnología para respuesta a incidentes. Es ensayo y comunidad.
5. Primero hay que saber de quién nos defendemos para poder elegir la estrategia.

Gracias.

Alberto Ávalos

CIO — Pryse Cyber CERT

FIRST México · Liaison

Conectemos

linkedin.com/in/**bettoavalos**