# Informe: Del riesgo de los agentes a los logros humanos

## Introducción

El panorama corporativo ha superado la etapa de experimentación teórica con inteligencia artificial y las organizaciones mexicanas ya están incorporando agentes autónomos y automatizaciones impulsadas por IA en sus operaciones diarias.

Sin embargo, mientras las organizaciones buscan capturar ganancias de eficiencia, sus posturas de seguridad comienzan a fragmentarse. El informe 'From Agentic Risk to Human Wins', de KnowBe4, destaca los puntos ciegos críticos, las vulnerabilidades psicológicas y las brechas de preparación que enfrentan actualmente las organizaciones del país.

## Sección 1: Shadow AI y el déficit de gobernanza

Mientras las organizaciones mexicanas incorporan IA en sus ecosistemas tecnológicos formales, una capa significativa y no gestionada de "Shadow AI" está emergiendo entre la fuerza laboral. Los actores maliciosos saben que los equipos de seguridad no pueden proteger activos que no pueden ver.

- El 56% de las organizaciones encuestadas en México implementan agentes autónomos de IA capaces de tomar decisiones y ejecutar acciones por sí mismos, mientras que el 55% de los líderes admite que el uso de IA dentro de su entorno carece de aprobación formal o de una gobernanza corporativa adecuada.
- En México, el 43% de los responsables de ciberseguridad afirma que el uso de software externo y herramientas de IA no aprobadas se encuentra entre los comportamientos que más han impactado la seguridad de sus organizaciones durante los últimos 12 meses
- Más de la mitad (53%) de los trabajadores mexicanos admite que no siempre utiliza las herramientas de IA proporcionadas por su organización, lo que refleja una tendencia a recurrir a soluciones externas cuando las opciones corporativas no satisfacen completamente sus necesidades laborales.

## Reflejos

- **Descubra la brecha de gobernanza:** Descubra el alcance de la IA no autorizada en las organizaciones y por qué este suele funcionar sin una gobernanza corporativa formal.
- **Evalúe el cambio hacia los agentes:** Descubra el porcentaje de organizaciones que implementan agentes de IA autónomos capaces de actuar por sí mismos, y las consecuencias que esto tiene para la seguridad.
- **Conozca el campo de batalla psicológico:** Descubra cómo el espectro de amenazas ha pasado a ser el de los deepfakes hiperrealistas creados por IA y por qué la mayoría de los empleados admiten que podrían ser engañados con éxito.
- **Identifique la amenaza real para los seres humanos:** Consulte los datos sobre cómo la sobrecarga cognitiva y la fricción operativa están obligando a los empleados a tomar atajos, lo que provoca un porcentaje significativo de incidentes de seguridad.
- **Evalúe su preparación:** Comprenda el estándar principal de la madurez en materia de seguridad y averigüe si los responsables de seguridad se sienten lo suficientemente resilientes como para sobrevivir a las amenazas emergentes impulsadas por la IA.

## Sección 2: Deepfakes y el campo de batalla psicológico

El vector de amenaza ha evolucionado drásticamente, pasando del phishing tradicional a manipulaciones hiperrealistas generadas por inteligencia artificial. Las defensas cognitivas humanas están teniendo dificultades para enfrentar contenidos creados por máquinas.

- Un contundente 85% de los empleados mexicanos afirma que los contenidos de voz y video generados mediante deepfakes son tan realistas que resulta cada vez más difícil saber qué es auténtico.
- El 50% de la fuerza laboral reconoce abiertamente que podría ser engañada por una estafa sofisticada basada en deepfakes que se haga pasar por un ejecutivo o una parte interesada dentro de la organización.
- El panorama de amenazas ha pasado de la velocidad humana a la velocidad de las máquinas. Las organizaciones mexicanas aún enfrentan desafíos para responder a ataques automatizados y multietapa que pueden dirigirse simultáneamente a los empleados a través del correo electrónico, SMS y plataformas de colaboración.

## Sección 3: Fricción operativa y la brecha del "Estándar de Oro"

La principal amenaza no radica únicamente en ataques avanzados, sino en el efecto acumulativo de la fatiga humana dentro de entornos corporativos de alta presión combinado con bajos niveles de madurez operativa en materia de seguridad.

- El 79% de los empleados mexicanos afirma sentirse confiado en su capacidad para reconocer amenazas impulsadas por inteligencia artificial, mientras que un 21% reconoce no sentirse preparado para identificarlas adecuadamente.
- El 50% de los líderes de ciberseguridad en México señala que los errores operativos cotidianos se encuentran entre los comportamientos humanos que más han impactado la seguridad de sus organizaciones durante el último año.
- A pesar de la creciente concienciación sobre ciberseguridad, solo el 12.5% de las organizaciones ha alcanzado un modelo de madurez considerado de 'estándar de oro', definido como una cultura completamente sincronizada capaz de gestionar simultáneamente riesgos humanos y riesgos asociados a la IA
- Además, el 88% de los líderes de seguridad en México afirma sentirse preparado para afrontar amenazas emergentes o inesperadas impulsadas por inteligencia artificial durante los próximos 12 meses.

## Conclusión

La fuerza laboral en México ha cambiado. Los empleados y los agentes de IA ahora operan como una única capa interconectada de defensa. Esto redefine por completo lo que significa una buena estrategia de seguridad.

Las organizaciones que alinean la concienciación, el comportamiento y la cultura en torno a una comprensión clara de sus riesgos hacen mucho más que gestionar la exposición. Se convierten en objetivos más difíciles de comprometer.

Las amenazas no desaparecerán, pero encontrarán organizaciones mejor preparadas para enfrentarlas.

KnowBe4, Inc.
R. Gomes de Carvalho, 911 | Sala 208 - Vila Olímpia, São Paulo - SP | CEP: 04547-003, São Paulo - Brasil
(0800) 7612668 | www.KnowBe4.com | Sales@KnowBe4.com