No. 179 Abril - Junio DOI: 10.29236/sistemas ISSN 0120-5919

SISTEMAS

Confianza digital: ¿innovación confiable?

ACIS
ASOCIACIÓN COLOMBIANA DE INFORMÁTICA, SISTEMAS Y TECNOLOGÍAS AFINES
Calle 93 No. 13 - 32 of. 102
Bogotá, D.C.
www.acis.org.co

ACIS

ASOCIACIÓN COLOMBIANA DE INFORMÁTICA, SISTEMAS Y TECNOLOGÍAS AFINES

50 AÑOS

1976 - 2026

Celebrando **medio siglo** de **potenciar y transformar el talento TI** en Colombia

Con orgullo celebramos 50 años como el pilar de la comunidad tecnológica colombiana, liderando el desarrollo e impulsando la educación.

En esta edición

Publicación de la Asociación Colombiana de Informática, Sistemas y Tecnologías Afines (ACIS)
Resolución No. 003983 del Ministerio de Gobierno
Tarifa Postal Reducida Servicios Postales Nacional S.A. No. 2015-186 4-72
ISSN 0120-5919
Apartado Aéreo No. 94334
Bogotá D.C., Colombia

Abril - Junio 2026
Calle 93 No.13 - 32 Of. 102
Teléfonos 616 1407 - 616 1409
A.A. 94334
Bogotá D.C.
www.acis.org.co

XXVI JORNADA INTERNACIONAL DE SEGURIDAD INFORMÁTICA

JISI 2026

Asociación Colombiana de Informática, Sistemas y Tecnologías Afines (ACIS)

PERSONAS, PROCESOS, TECNOLOGÍAS, NORMATIVAS

Celebrando el hito de JISI, consolidada como el evento líder en Colombia y Latinoamérica. Un espacio para revisar el pasado, explorar nuevos horizontes y abordar los desafíos en personas, procesos, tecnologías y normativas.
Evolucionemos: de proteger a defender y anticipar.

1 y 2 de Octubre de 2026

ACIS
ASOCIACIÓN COLOMBIANA DE INFORMÁTICA, SISTEMAS Y TECNOLOGÍAS AFINES

La confianza digital: el activo invisible que sostiene el futuro

DOI: 10.29236/sistemas.n179a1

Jeimy J. Cano M.

Andres R. Almanza J.

Durante años, la conversación sobre el mundo digital estuvo dominada por la tecnología. Hablamos de transformación digital, automatización, computación en la nube, inteligencia artificial, plataformas, datos y ciberseguridad. Sin embargo, a medida que las organizaciones avanzan en su proceso de transformación, emerge una pregunta más retadora: ¿qué es realmente lo que permite que las personas, las empresas, las instituciones y entidades gubernamentales adopten, apropien, utilicen y asuman estos entornos digitales?

La respuesta parece sencilla, pero encierra un desafío particular: la confianza.

La confianza ha sido históricamente el fundamento de las relaciones humanas, comerciales e institucionales. Hoy, en un contexto donde gran parte de las relaciones ocurren a través de sistemas, algoritmos, plataformas y ecosistemas digitales, la confianza se ha convertido en un activo estratégico que determina la viabilidad misma de los modelos de negocio y de las sociedades digitales.

Ya no es suficiente con que un servicio y/o producto funcione de acuerdo con su especificación. Debe ser confiable y mostrar confiabilidad. Ya no es suficiente asegurar la confidencialidad, la integridad y la disponibilidad. Es necesario entender con claridad sobre cómo se comporta, quién lo administra y cuáles son sus principios de gobernanza. Tampoco es suficiente incorporar nuevas tecnologías; es preciso asegurar que estas sean distinguidas como transparentes, responsables y alineadas con las expectativas de las personas y los objetivos de negocio, más aún en contextos donde la inteligencia artificial ha transformado la realidad empresarial.

La confianza digital, de acuerdo con Saveljeva & Volkova (2025) se fundamenta en cuatro pilares esenciales: la capacidad, la confiabilidad, la integridad y la dimensión de la apertura.

La capacidad abarca la competencia técnica y la seguridad del sistema para operar con estabilidad y desempeño. La confiabilidad se centra en la intención positiva hacia el usuario y la calidad del soporte recibido durante su experiencia. Por su parte, la integridad asegura que la organización actúe bajo principios éticos y cumpla con las normativas de protección de datos vigentes. Finalmente, la apertura asegura la transparencia, auditabilidad y trazabilidad de los procesos, permitiendo la verificación externa de las operaciones.

La integración de estas dimensiones permite reducir los riesgos y fomentar interacciones bien fundadas en entornos tecnológicos complejos.

Este desafío adquiere una relevancia particular en América Latina y el Caribe. La región enfrenta simultáneamente importantes oportunidades de desarrollo digital y brechas estructurales. Organismos internacionales han señalado que millones de personas aún carecen de acceso significativo a Internet, mientras persisten desigualdades en conectividad, capacidades digitales y acceso a servicios tecnológicos (OECD, 2023; OECD 2025). A ello se suma una realidad histórica caracterizada por bajos niveles de confianza interpersonal e institucional, un factor que inevitablemente se proyecta sobre los entornos digitales.

En este contexto, la confianza digital deja de ser un asunto exclusiva-

mente tecnológico para convertirse en un desafío sistémico. Involucra liderazgo, gobernanza, regulación, cultura organizacional, diseño de servicios, ética, transparencia, resiliencia y responsabilidad.

Las organizaciones más interconectadas están comprendiendo que la confianza no es el resultado accidental de hacer bien las cosas. Es un concepto multidimensional que debe diseñarse, desarrollarse, medirse, gestionarse y gobernarse. Se diseña desde la estrategia, se asegura en los procesos, se despliega en los productos y servicios, y se valida permanentemente en cada interacción con clientes, ciudadanos, colaboradores y socios de negocio.

La innovación tecnológica acelera aún más esta necesidad. La inteligencia artificial, la automatización de decisiones, los ecosistemas interconectados y las nuevas formas de intercambio digital están redefiniendo la manera como se generan valor y se construyen relaciones. En este escenario, la pregunta ya no es únicamente cómo innovar más rápido, sino cómo hacerlo motivando, preservando y fortaleciendo la confianza.

De igual forma, la resiliencia también adquiere una nueva lectura. Durante mucho tiempo se pensó en llegar a organizaciones capaces de evitar cualquier falla (OECD, 2023; NIST, 2024). Hoy comprendemos que la complejidad creciente de los entornos digitales hace prácticamente imposible eliminar la incertidumbre. Por tanto, lo relevante es desarrollar capacidades para responder, adaptarse, aprender, reinventarse y permanecer aun frente a la inevitabilidad de la falla. La confianza se consolida no porque los incidentes nunca ocurran, sino porque las organizaciones demuestran que pueden gestionarlos con transparencia, responsabilidad, preparación y eficacia, porque pueden tomar decisiones responsables, y adicionalmente porque toda la organización desde su nivel directivo hasta su nivel operacional está involucrada en ello.

Esta edición de la Revista Sistemas invita precisamente a reflexionar sobre esta transición. Una transición que desplaza el foco desde la protección aislada hacia la construcción integral de confianza. Desde la tecnología como herramienta hacia la confianza como propósito, como una perspectiva multidimensional. Desde la gestión de riesgos como obligación hacia la generación de valor como resultado.

Quizás uno de los mayores aprendizajes de esta era digital sea reconocer que la tecnología puede conectar sistemas, automatizar procesos y acelerar decisiones, pero solamente la confianza (que no es ausencia de eventos adversos o errores) es capaz de sostener relaciones duraderas entre personas, organizaciones y sociedades cuando las cosas no ocurren como esta-

ban planeadas. En última instancia, el futuro digital no será definido por quienes desarrollen más tecnología, sino por quienes logren inspirar, demostrar y mantener mayores niveles de confianza en ella.

Referencias

Saveljeva, J., & Volkova, T. (2025). A Survey on Digital Trust: Towards a Validated Definition. *Digital, 5*(2), 14. https://doi.org/10.3390/digital5020014

OECD/CAF (2023), *Digital Government Review of Latin America and the Caribbean: Building Inclusive and Responsive Public Services*, OECD Digital Government Studies, OECD Publishing, Paris. https://doi.org/10.1787/29f32e64-en.

OECD (2025), *OECD Survey on Drivers of Trust in Public Institutions in Latin America and the Caribbean 2025 Results*, OECD Publishing, Paris. https://doi.org/10.1787/ea3385cf-en.

NIST (2024) National Institute of Standards and Technology. *Cybersecurity framework (CSF) 2.0*. U.S. Department of Commerce. https://www.nist.gov/cyberframework

Jeimy J. Cano M., Ph.D, CFE, CICA. Ingeniero y Magíster en Ingeniería de Sistemas y Computación, Universidad de los Andes. Especialista en Derecho Disciplinario, Universidad Externado de Colombia; Ph.D en Business Administration, Newport University, CA. USA. y Ph.D en Educación, Universidad Santo Tomás. Profesional certificado como Certified Fraud Examiner (CFE), por la Association of Certified Fraud Examiners y Certified Internal Control Auditor (CICA) por The Institute of Internal Controls. Profesor Distinguido de la Facultad de Derecho, Universidad de los Andes. Es director de la Revista SISTEMAS de la Asociación Colombiana de Ingenieros de Sistemas –ACIS–.

Andreés R. Almanza J., CISM, Ingeniero Ingeniería de Sistemas de la Universidad Católica de Colombia y Master en seguridad de la información de la, Universidad Oberta de Catalunia. Especialista en Seguridad de Redes de la Universidad Católica de Colombia, . Profesional certificado como Certified Information Security Manager (CISM), por la Information Systems Audit and Control Association (ISACA) .Catedrático de la Facultad de Administración de la Universidad Exetrnado de Colombia. Miembro del comité editorial de la Revista SISTEMAS de la Asociación Colombiana de Ingenieros de Sistemas –ACIS–.director general CISOS.CLUB y la Asociación de Profesionales de Seguridad y Ciberseguridad (APSIC)

Confianza digital: activo a desarrollar

DOI: 10.29236/sistemas.n179a2

La confianza digital es un activo estratégico fundamental para sociedades crecientemente dependientes de ecosistemas digitales hiperconectados que redefinen el riesgo.

Diego Andrés Zuluaga Urrea

La transformación digital ha dejado de ser únicamente un proceso de modernización tecnológica. Cada vez más desarrolla la economía, la industria, la operación de infraestructuras críticas y otros sistemas que controlan el mundo físico desde el mundo virtual (ciberfísicos), así como la interacción entre gobiernos, empresas y sociedad civil en entornos digitalizados.

Por ello, la confianza digital se ha convertido en un activo estratégico fundamental para nuestra sociedad, ya que vivimos no solo en el mundo físico sino en el virtual y permitimos cada vez más que nuestra vida se desarrolle en este mundo virtual y que la tecnología digital tome cada vez más decisiones y acciones que afectan el mundo físico que nos rodea y que presta los servicios esenciales que requerimos en la sociedad actual, a la vez que facilita el desarrollo de los productos y servicios que consumimos en nuestra vida diaria.

Sin embargo, a diferencia de otros activos tecnológicos o financieros, el desarrollo de esta confianza digi-

tal aún es limitado, difuso y, en muchos casos, subestimado o dejado de lado por las diferentes instituciones, gobiernos y empresas. En un contexto caracterizado por redes altamente interconectadas, automatización avanzada, virtualización de servicios e inteligencia creciente de los sistemas, la confianza ya no puede asumirse como una consecuencia natural de la tecnología, sino como una capacidad que debe diseñarse, construirse y mantenerse de forma explícita. Su ausencia no solo incrementa el riesgo, sino que compromete la estabilidad operativa, la resiliencia organizacional, la credibilidad organizacional e institucional y el uso futuro de los ecosistemas digitales en su conjunto.

En los últimos años ha comenzado a consolidarse una nueva visión sobre la evolución del riesgo digital y operacional. Por ejemplo, el estudio: *How can reimagining risk prepare you for an unpredictable world?* de EY introduce el concepto de entorno NAVI para describir un mundo crecientemente No lineal, Acelerado, Volátil e Interconectado [1]. Según este enfoque, los riesgos modernos ya no evolucionan de manera aislada ni predecible. Por el contrario, interactúan entre sí, se aceleran mutuamente y generan efectos en cascada capaces de impactar simultáneamente operaciones, cadenas de suministro, reputación, estabilidad financiera e infraestructura crítica, creando policrisis que deben ser gestionadas a tiempo para evitar impactos significativos en los servicios prestados por las instituciones y empresas o en ellas mismas.

En este contexto NAVI, la confianza digital deja de ser únicamente un atributo tecnológico o reputacional para convertirse en una capacidad estratégica de resiliencia organizacional. La capacidad de anticipar, adaptarse, responder y mantener continuidad operacional frente a escenarios impredecibles debe ser uno de los principales diferenciadores de las organizaciones modernas.

Ya que La sociedad está comenzando a delegar crecientemente decisiones críticas a sistemas digitales autónomos. Algoritmos y plataformas participan en procesos financieros, médicos, logísticos, industriales y de infraestructura crítica. Esto convierte la confianza digital no solo en un requisito tecnológico, sino en un elemento esencial de gobernanza y estabilidad social.

Emergen preguntas fundamentales como: ¿qué tan confiables son los ecosistemas digitales sobre los cuales estamos construyendo nuestra dependencia económica y social para estos entornos NAVI? Y ¿Cómo debemos desarrollar la confianza digital en nuestras organizaciones en esta nueva era donde la IA acelera los riesgos?

Ya no basta con digitalizar procesos, migrar servicios a la nube o

incorporar inteligencia artificial. El verdadero reto consiste en garantizar que los ecosistemas digitales sean seguros, resilientes, verificables y sostenibles frente a amenazas crecientemente sofisticadas y dinámicas, especialmente aceleradas por el uso creciente de la Inteligencia artificial agentica.

La confianza digital dejó de ser únicamente un atributo reputacional para convertirse en una propiedad esencial de la infraestructura tecnológica moderna que rige nuestra sociedad.

Actualmente ciudadanos, gobiernos, industrias y empresas dependen de plataformas digitales para operar procesos críticos: energía, salud, servicios financieros, telecomunicaciones, logística, producción industrial y servicios públicos. En consecuencia, cualquier afectación sobre la confianza del ecosistema digital puede traducirse en impactos económicos, sociales e incluso institucionales.

Este cambio de paradigma resulta particularmente relevante para países Latinoamericanos como Colombia, donde la acelerada digitalización convive con desafíos relacionados con ciberseguridad, madurez institucional, baja protección de la infraestructura crítica y débil fortalecimiento de capacidades especializadas.

Después de más de tres décadas observando la evolución de Internet, las telecomunicaciones, la infraestructura tecnológica y la ciberseguridad en Colombia y América Latina, resulta evidente que estamos entrando en una etapa distinta del riesgo digital.

Durante muchos años la ciberseguridad se enfocó principalmente en proteger información. Hoy el problema es considerablemente más complejo. La seguridad digital protege continuidad operacional, estabilidad económica, confianza institucional, la seguridad física y defensa nacional a la vez que protege de afectaciones a sistemas ciberfísicos que soportan la infraestructura crítica nacional, la industria, sistemas médicos avanzados, entre otros, es decir los servicios esenciales que la sociedad requiere para su funcionamiento, evitando no sólo que estos fallen sino que se descontrolen y puedan causar impactos sobre las vidas humanas y el medio ambiente [2].

La convergencia entre tecnologías de información IT y las de operación OT, la hiperconectividad, la virtualización de servicios, la dependencia de terceros y la creciente incorporación de inteligencia artificial han transformado radicalmente la superficie de exposición de las organizaciones.

“La ciberseguridad es la base de nuestro mundo digital. Está en el centro de la confianza y permitirá que la sociedad aproveche plenamente las transformaciones impul-

sadas por nuevas tecnologías como la inteligencia artificial y la computación cuántica…" -Michael Miebach, CEO de Mastercard [3]

El modelo NAVI acelera esa complejidad. Los ecosistemas digitales modernos funcionan como redes profundamente interdependientes donde convergen infraestructuras distribuidas, automatización avanzada, plataformas cloud, inteligencia artificial, dispositivos IoT, entornos industriales conectados, servicios virtualizados y cadenas de suministro digitales globales.

En este contexto, la confianza digital no puede entenderse únicamente como un asunto tecnológico. Se trata de una capacidad multidimensional que involucra seguridad, resiliencia, gobernanza, privacidad, integridad, trazabilidad y capacidad de recuperación.

Diversos organismos internacionales han desarrollado aproximaciones complementarias sobre este concepto. El World Economic Forum define relaciona la confianza digital como *" la expectativa de las personas de que las tecnologías y los servicios digitales, y las organizaciones que los proporcionan, protegerán los intereses de todas las partes interesadas y respetarán las expectativas y los valores de la sociedad.*" y las relaciona en su modelo con la capacidad de las organizaciones para garantizar seguridad de las personas y la operación(Safety), la ciberseguridad, transparencia, reparabilidad, auditabilidad, privacidad, interoperabilidad y equidad de la tecnología, agrupando estos conceptos en las dimensiones de Seguridad y confiabilidad, supervisión y rendición de cuentas, ética, inclusividad y uso responsable, [4].

La OECD ha insistido en que la confianza constituye un habilitador esencial para el crecimiento sostenible de la economía digital. Por su parte, marcos desarrollados por NIST como el CSF 2.0 han contribuido a operacionalizar conceptos asociados a gestión de riesgo, resiliencia y arquitecturas de confianza y pueden usarse como marco de referencia para la construcción de entornos seguros y resilientes [5].

La confianza digital adquiere entonces una dimensión estratégica nacional, Cuando una sociedad pierde confianza en sus sistemas digitales, el impacto trasciende lo tecnológico: disminuye la adopción de servicios, aumenta la incertidumbre, se debilita la legitimidad institucional y se afecta la estabilidad económica.

Sin embargo, la aparición de nuevas generaciones de inteligencia artificial está introduciendo un punto de inflexión aún más complejo en la evolución del riesgo digital. La IA ya no representa únicamente una herramienta de automatización o productividad. Está comenzando a convertirse en un multiplicador operacional capaz de alterar signi-

ficativamente la dinámica entre defensores y atacantes.

Las discusiones recientes alrededor de capacidades emergentes de modelos avanzados reflejan una preocupación creciente: la posibilidad de que agentes inteligentes reduzcan las barreras técnicas necesarias para ejecutar ataques complejos y amplifiquen la capacidad ofensiva de actores maliciosos o incluso actúen completamente de manera autónoma descubriendo y explotando los entornos de interés de los atacantes.

Más allá del debate mediático, el problema de fondo es estratégico. Las nuevas inteligencias artificiales incrementarán la capacidad operacional tanto de los defensores como de los atacantes.[6]

Solo entre abril y mayo de 2026 tres agentes de OpenAI, Anthropic y Microsoft sobrepasaron la barrera del 80% de efectividad en la generación de Pruebas de concepto funcionales de vulnerabilidades dentro del CyberGym de UC Berkeley [7] lo que muestra que las capacidades actuales superan equipos humanos especializados en velocidad y escala, ampliando significativamente el panorama de amenazas.

Actividades como reconocimiento de infraestructura, identificación y explotación de vulnerabilidades, phishing avanzado, evasión de controles o desarrollo de malware pueden ejecutarse con mayor automatización y escala. La velocidad de evolución de las amenazas también aumentará, dificultando los mecanismos tradicionales de detección.

Incluso esta evolución ya no pertenece únicamente al terreno teórico. Investigaciones recientes de Trend Micro documentaron campañas dirigidas contra sectores gubernamentales y financieros en América Latina donde actores maliciosos utilizaron capacidades de IA agentica para automatizar diferentes etapas del ciclo de ataque, desde reconocimiento inicial hasta despliegue de herramientas ofensivas y exfiltración de información [8].

Con lo cual se evidencia cómo la inteligencia artificial comienza a reducir las barreras técnicas necesarias para ejecutar operaciones avanzadas, aumentando la velocidad, escala y adaptabilidad de los ataques. Este tipo de escenarios marca un punto de inflexión particularmente relevante para la confianza digital en entornos NAVI, donde la hiperconectividad y la dependencia de cadenas de suministro digitales amplifican el impacto potencial de incidentes cibernéticos sobre múltiples organizaciones y sectores simultáneamente.

Además, la combinación de inteligencia artificial generativa, deepfakes y manipulación sintética afectará directamente la capacidad de verificar la legitimidad de personas,

transacciones y comunicaciones generando erosión de la autenticidad digital con lo cual, “ver” o “escuchar” dejará de ser suficiente como mecanismo de validación.

Por otro lado, desde la perspectiva de cadenas de suministro digitales, las organizaciones operan dentro de ecosistemas interconectados que incluyen proveedores cloud, integradores, plataformas SaaS, open source, OT, IoT y terceros. La confianza ya no depende únicamente de la seguridad interna, sino de toda la cadena de suministro.

La automatización ofensiva impulsada por IA permitirá explotar eslabones débiles y comprometer ecosistemas completos con velocidades sin precedentes. La experiencia internacional ya ha demostrado el impacto global de incidentes en cadena de suministro [9].

La industrialización progresiva del cibercrimen impulsada por IA, estamos entrando en una etapa donde campañas ofensivas completas podrán ejecutarse de manera crecientemente automatizada, incluyendo reconocimiento, explotación, evasión y monetización [10].

En este escenario, la confianza digital dependerá cada vez más de la capacidad de las organizaciones para gestionar riesgos sistémicos e interdependencias complejas dentro de ecosistemas digitales altamente distribuidos.

En infraestructuras críticas, donde convergen tecnologías IT y OT, el riesgo adquiere implicaciones operacionales y físicas. Por ello, marcos especializados como ISA/IEC 62443 resultan fundamentales para la seguridad industrial [11].

Como plantea EY, las organizaciones más resilientes evolucionan desde enfoques tradicionales de gestión de riesgo hacia modelos de *Risk Strategists*, donde el riesgo deja de gestionarse únicamente desde cumplimiento y pasa a integrarse directamente con estrategia, resiliencia, gobernanza y toma de decisiones [1].

Esta transición resulta especialmente relevante en ciberseguridad, donde los modelos tradicionales basados únicamente en prevención perimetral ya no son suficientes frente a amenazas hiperconectadas, automatizadas y potenciadas por inteligencia artificial.

Todo esto obliga a replantear profundamente las estrategias tradicionales de ciberseguridad frente a amenazas adaptativas. Las organizaciones deben fortalecer capacidades de ciberresiliencia, inteligencia de amenazas, monitoreo continuo, gestión de terceros, seguridad de cadena de suministro y gobernanza de inteligencia artificial.

La conversación ya no es sobre proteger sistemas, sino sobre preservar la confianza operacional de

ecosistemas digitales en entornos NAVI.

En un mundo donde los sistemas digitales controlan progresiva y determinadamente procesos físicos esenciales, desarrollar confianza digital en la sociedad dejará de ser opcional para convertirse en una condición de estabilidad económica y organizacional, resiliencia institucional y seguridad nacional.

Latinoamérica y en especial Colombia enfrentan aquí una oportunidad estratégica. La transformación digital del país debe evolucionar hacia un modelo de confianza digital y ciberresiliencia que permita afrontar las multicrisis a las que ya están acostumbradas en el mundo físico ahora en el mundo digital.

La confianza digital no es cumplimiento, es estrategia

Los países y organizaciones que logren integrar y consolidar ecosistemas digitales resilientes y confiables tendrán ventajas significativas en innovación, estabilidad y confianza pública que redundará en desarrollo económico para el largo plazo.

En un mundo hiperconectado, la confianza será uno de los principales diferenciadores estratégicos. Y la ciberresiliencia dejará de ser un tema técnico para convertirse en un pilar de la competitividad empresarial y nacional.

Referencias

[1] McCowan, S., Krumbmüller, F. & Jaggi, G. (2025). *How can reimagining risk prepare you for an unpredictable world?* EY Insights. https://www.ey.com/en_us/insights/consulting/how-can-reimagining-risk-prepare-you-for-an-unpredictable-world

[2] ZULUAGA, Diego, et al. *Escenarios de alto impacto por ciberseguridad en sistemas industriales del sector eléctrico*. Madrid: CCI-CIGRE, 2023. ISBN 978-84-126727-6-3. Disponible en: https://www.cci-es.org/activities/escenarios-de-alto-impacto-por-ciberseguridad-en-sistemas-industriales-sector-electrico/

[3] World Economic Forum. (2026). *Global Cybersecurity Outlook 2026*. https://reports.weforum.org/docs/WEF_Global_Cybersecurity_Outlook_2026.pdf

[4] World Economic Forum. (2022). *Earning digital trust: Decision-making for trustworthy technologies*. World Economic Forum. https://www3.weforum.org/docs/WEF_Earning_Digital_Trust_2022.pdf

[5] National Institute of Standards and Technology. (2024). *Cybersecurity framework (CSF) 2.0*. https://www.nist.gov/cyberframework

[6] World Economic Forum. (2025). *Artificial intelligence and cybersecurity: Balancing risks and rewards*. World Economic Forum. https://reports.weforum.org/docs/WEF_Artificial_Intelligence_and_Cybersecurity_Balancing_Risks_and_Rewards_2025.pdf

[7] International Conference on Learning Representations. Wang, Z., Shi, T., He,

J., Cai, M., Zhang, J., & Song, D. (2026). *CyberGym: Evaluating AI agents' real-world cybersecurity capabilities at scale*. *In Proceedings of the Fourteenth International Confe-rence on Learning Representations (ICLR 2026)*. https://openreview.net/forum?id=2YvbLQEdYt

[8] Trend Micro. (2026). *Vibe hacking: Two AI-augmented campaigns target government and financial sectors in Latin America*. https://www.trendmicro.com/en_us/research/26/e/vibe-hacking-two-ai-augmented-campaigns-target-government-and-financial-sectors-in-latin-america.html

[9] IBM Security. (2025). *Cost of a data breach report 2024*. https://www.ibm.com/security/data-breach

[10] Trend Micro. (2025). *The AI-fication of cyberthreats: Security predictions for 2026*. https://www.trendmicro.com/vinfo/us/security/research-and-analysis/predictions/the-ai-fication-of-cyberthreats-trend-micro-security-predictions-for-2026

[11] International Society of Automation. (2026). *ISA/IEC 62443 series of standards*. International Society of Automation. https://www.isa.org/standards-and-publications/isa-standards/isa-iec-62443-series-of-standards

Diego Andrés Zuluaga Urrea. *Ingeniero de Sistemas, Especialista en Gerencia y Executive MBA con más de 25 años de experiencia y certificaciones en ciberseguridad, gestión de riesgos, privacidad, seguridad de la información y protección de sistemas de control industrial. Actualmente es CSO para Latinoamérica y responsable del gobierno de seguridad para mercados internacionales del grupo AXA, desde Madrid, España. Lideró la construcción de la regulación y guías de ciberseguridad en el sector eléctrico colombiano, ha aportado en diferentes escenarios para la seguridad de las infraestructuras críticas nacionales. Es coordinador nacional del Centro de Ciberseguridad Industrial y líder del capítulo SC D2 de CIGRE. Ha sido, consultor internacional, conferencista y docente universitario en América Latina y Europa, además de autor y revisor de múltiples publicaciones especializadas. Ha recibido reconocimientos nacionales e internacionales, entre ellos en dos oportunidades el premio continental "Americas Information Security Leadership Awards" de ISC².*

Alfredo Amore Pardo

DOI: 10.29236/sistemas.n179a3

Sin duda, un líder que representa buena parte de los acontecimientos más relevantes del sector informático del país.

Desde su perspectiva y visión de la tecnología, así como su desarrollo en Colombia, ¿las empresas colombianas realmente están asumiendo una transformación digital?
Sí, sin duda. Las empresas a todo nivel, se han dado cuenta de la ventaja y necesidad para mantenerse y competir, están asumiendo la transformación digital como una obligación.

Como uno de los primeros ingenieros de sistemas del país, ¿qué transformaciones ha visto en el país por la incorporación de las tecnologías de información?
El gobierno nacional y los gobiernos locales, han asumido la tecnología de información, como un campo en el que tiene que estar presentes, y obtener ventajas en sus campos de acción.

Desde que recibió su título como Ingeniero de Sistemas, ¿qué podría comentarnos qué ha cambiado y cómo? ¿Qué temáticas ve a futuro en la formación de las nuevas generaciones de ingenieros?
Ha cambiado y continúa permanentemente el cambio. Todas las empresas y entidades tienen una división de sistemas de informa-

ción, que está en los más altos niveles de toma de decisiones, y le está destinado un presupuesto propio, un plan de desempeño y obtención de resultados, al más alto nivel de la organización.

¿Cómo ha visto el papel de la ACIS en los diferentes momentos y olas de tecnología que hemos tenido en el país? ¿Qué recuerda particularmente con especial atención? ¿Cómo ve a ACIS en los próximos 50 años? ¿Qué retos y expectativas ve para la Asociación?

ACIS está presente como entidad profesional a la cual recurren muchas empresas y entidades, para obtener orientación en el uso y desarrollo de sus sistemas de información.

Encuesta latinoamericana de seguridad de la información 2026.

Resultados y revelaciones estratégicas

DOI: 10.29236/sistemas.n179a4

Jeimy J. Cano M.

Andrés R. Almanza J.

Gabriela M. Saucedo M.

Resumen

El presente estudio analiza el estado de la seguridad de la información de 115 organizaciones participantes de Latinoamérica, con el objetivo de caracterizar el estado de su gobierno de ciberseguridad y de identificar las relaciones de mayor valor entre sector económico, capacidad presupuestal, exposición a incidentes y prioridades temáticas para 2026. Los resultados muestran que el 74.1% de las organizaciones dispone de rubro presupuestal para 2026, con una marcada asimetría sectorial: los sectores regulados presentan una probabilidad ocho veces mayor de contar con presupuesto. No obstante, la existencia de presupuesto se desacopla de su magnitud, y el costo de los incidentes se asocia más con la exposición sectorial y el tamaño organizacional, que con el nivel de inversión. La agenda 2026 está dominada por la inteligencia artificial, la fuga de información y las amenazas persistentes avanzadas.

Palabras clave

Seguridad de la información, ciberseguridad, gobierno de seguridad, gestión de riesgos, madurez organizacional

Introducción

La transformación digital ha situado a la seguridad de la información en el centro de la gestión del riesgo empresarial. La creciente sofisticación de las amenazas, impulsada por tecnologías emergentes y por un entorno geopolítico inestable, ha ampliado la superficie de ataque de las organizaciones y ha tensionado su capacidad de respuesta.

En este contexto, anticipar las amenazas cibernéticas dejó de ser una función exclusivamente técnica para convertirse en una decisión estratégica que compromete a la alta dirección, al presupuesto y a la cultura organizacional.

A pesar de la abundante evidencia internacional sobre costos de las brechas y tendencias de amenazas, persiste un vacío en la comprensión de cómo se distribuyen, a escala sectorial, las capacidades de gobierno, inversión y prevención dentro de un mismo ecosistema regional. Comprender estas asimetrías es indispensable para diseñar intervenciones que eleven la resiliencia cibernética colectiva.

Este estudio aborda dicho vacío mediante el análisis de los datos de la Encuesta Latinoamericana de Seguridad de la Información 2026 (ELSI 2026). El estudio prioriza, por su valor accionable, los cruces entre sector económico, capacidad presupuestal, costo de los incidentes y temas clave para 2026.

Metodología

Diseño y muestra

Se empleó un diseño cuantitativo, observacional y de corte transversal. La muestra estuvo conformada por 115 respuestas válidas obtenidas mediante un cuestionario autoadministrado de 22 ítems, distribuido entre responsables y profesionales de seguridad de la información de organizaciones de diversos sectores en Latinoamérica. El muestreo fue no probabilístico por autoselección, condición que delimita el alcance inferencial de los hallazgos a la población participante.

Técnicas estadísticas

El análisis descriptivo empleó distribuciones de frecuencia y, para las variables de respuesta múltiple, conjuntos de respuesta múltiple expresados como porcentaje de casos. Para las asociaciones nominales se utilizó la prueba chi-cuadrado de independencia de Pearson, acompañada del tamaño del efecto V de Cramér; cuando las frecuencias esperadas fueron reducidas se recurrió a la prueba exacta de Fisher. Las asociaciones entre variables ordinales se evaluaron con el coeficiente tau-b de Kendall, y la presencia de tendencias monótonas de prevalencia a través de niveles ordinales de presupuesto se contrastó con la prueba de Cochran-Armitage (Agresti, 2019). El procesamiento se realizó con el apoyo de inteligencia artificial gene-

rativa avanzada, validando cada una de sus respuestas contra los datos originales.

A continuación se detallan las técnicas estadísticas aplicadas: (Agresti, 2019).

- *Prueba chi-cuadrado de independencia de Pearson*: Compara las frecuencias observadas en cada celda de la tabla con las frecuencias que se esperarían si las variables fueran totalmente independientes. Un valor de X^2 elevado, con un p-valor pequeño, sugiere que la asociación observada es poco probable que ocurra por puro azar.

- *Prueba exacta de Fisher*: Es una prueba de independencia diseñada específicamente para tablas de contingencia (originalmente de 2×2) que no depende de aproximaciones para muestras grandes. Es la opción preferida cuando el tamaño de la muestra es pequeño o cuando los datos están muy desequilibrados.

- *Prueba de Cochran-Armitage*: En lugar de buscar cualquier tipo de asociación, esta prueba detecta específicamente si existe una tendencia lineal (aumento o disminución) en la proporción de éxitos a medida que aumenta el nivel de la variable ordinal.

- *Coeficiente tau-b de Kendall*: Evalúa la fuerza y la dirección de la asociación basándose en la concordancia y discordancia de los pares de observaciones.

- *Efecto V de Cramér*: Mientras que la prueba Chi-cuadrado solo indica si hay una asociación significativa (p-valor), la V de Cramér escala ese resultado entre 0 (independencia total) y 1 (asociación perfecta), permitiendo comparar la magnitud del efecto independientemente del tamaño de la muestra.

Resultados

Análisis descriptivo

La muestra refleja una fuerte presencia de los sectores de consultoría especializada, gobierno y servicios financieros, seguidos por educación (Figura 1). En conjunto, el 74.1% de las organizaciones declara contar con un rubro presupuestal para seguridad de la información en 2026, mientras que el 25.9% carece de él. El número medio de mecanismos de protección implementados es de 9.8. En el plano del gobierno, el 27% de las organizaciones no cuenta con un CISO y el 32% no realiza ejercicios de análisis de escenarios de riesgo, lo que evidencia una base de gestión todavía incipiente en una porción relevante de la muestra. Se destaca dentro de los resultados que 55 organizaciones reportan la gestión de incidentes a sus equipos ejecutivos o junta directiva, en tanto que 19 no presentan informe alguno.

Figura 1
Distribución de la muestra por sector económico

Nota. Frecuencia de organizaciones por sector tras la normalización de la variable (n = 115). Elaboración propia.

Análisis exploratorio: cruces de mayor valor

Sector y presupuesto. La existencia de rubro presupuestal varía significativamente entre grupos sectoriales (Tabla 1). La prueba chi-cuadrado resultó estadísticamente significativa ($\chi^2 = 11.137$; gl = 5; p = 0.0487) con un tamaño del efecto moderado (V de Cramér = 0.23). Dado que el 41.7% de las celdas presentó frecuencias esperadas inferiores a cinco, se complementó con una prueba exacta de Fisher sobre la dicotomía regulados frente a no regulados, que confirmó una asociación robusta: las organizaciones de sectores regulados (finanzas y gobierno) tienen una probabilidad cerca de ocho veces mayor de contar con presupuesto (OR = 8.3; p = 0.001).

Presupuesto, costo de incidentes y tamaño. La existencia de presupuesto se desacopla de su magnitud: sectores como gobierno presentan alta cobertura presupuestal, pero montos comparativamente bajos, mientras que manufactura y servicios financieros concentran los presupuestos totales más elevados. El costo de los incidentes, por su parte, no se asocia de forma significativa con el nivel de inversión (tau-b costo–presupuesto total = 0.204, p = 0.11), sino que sigue un

Tabla 1

Existencia de presupuesto de seguridad por grupo sectorial

Grupo sectorial	Sí	No	Total	% Sí
Consultoría	20	11	31	65%
Educación	7	5	12	58%
Finanzas	15	0	15	100%
Gobierno	17	2	19	89%
Otros	23	10	33	70%
Salud	4	2	6	67%

Nota. $\chi^2(5) = 11.137$, $p = 0.0487$, V de Cramér = 0.23. Prueba exacta de Fisher (regulados vs. no regulados): OR = 8.3, $p = 0.001$.

Nota. $\chi^2(5) = 11.137$, $p = 0.0487$, V de Cramér = 0.23. Prueba exacta de Fisher (regulados vs. no regulados): OR = 8.3, $p = 0.001$.

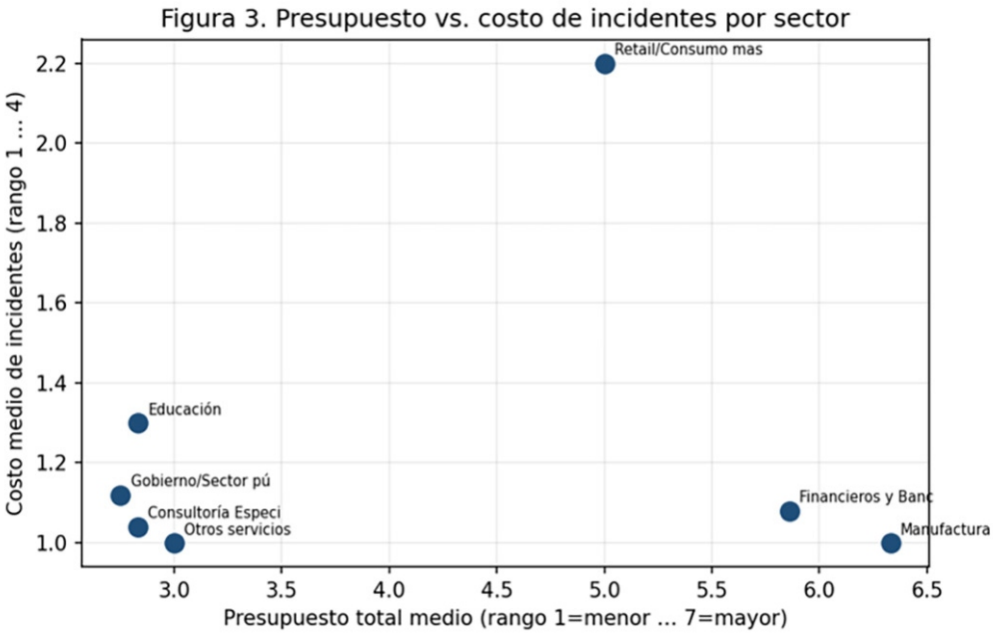
patrón de exposición sectorial: retail y telecomunicaciones reportan los costos medios más altos (Figura 2). El costo muestra además una tendencia ascendente con el tamaño organizacional (tau-b = 0.177, p = 0.0518), consistente con una mayor superficie de ataque y mayores activos en riesgo en las organizaciones grandes.

Sector, presupuesto y agenda 2026. La agenda temática para 2026 está dominada por la inteligencia artificial y las amenazas basadas en IA, seguidas por la fuga de información sensible y las amenazas persistentes avanzadas (Tabla 2 y Figura 3). El análisis de tendencia muestra matices según la capacidad presupuestal: las organizaciones con menor presupuesto enfatizan temas defensivos fundamentales, fuga de información e inteligencia de amenazas, mientras que las de mayor presupuesto incorporan con más frecuencia la protección de infraestructuras críticas y el internet de las cosas (Tabla 3).

Incidentes y obstáculos. Entre las organizaciones que experimentaron incidentes en 2025, predominan los originados en el factor humano: errores humanos (60.2%) y phishing (51.8%), seguidos de la ingeniería social (36.1%). De manera coherente, el obstáculo principal percibido para una adecuada postura de seguridad es la ausencia de una cultura de seguridad (44.7%), seguida por la falta de apoyo directivo (32.5%) y la complejidad tecnológica (30.7%). El factor humano y el gobierno, más que la tecnología, emergen como los determinantes críticos (Figura 5).

Figura 2

Relación entre presupuesto total medio y costo de incidentes por sector



Nota. Cada punto representa un sector; los ejes expresan rangos ordinales medios. La ausencia de una pendiente clara ilustra el desacople entre inversión y costo de incidentes. Elaboración propia.

Tabla 2

Temas clave priorizados para 2026 (diez principales)

Tema	n	% de casos
Inteligencia Artificial	88	76.5%
Amenazas basadas en IA	79	68.7%
Fuga de información sensible	78	67.8%
Amenazas persistentes avanzadas	70	60.9%
Ataques a infraestructuras críticas	64	55.7%
Seguridad y control en la computación en la nube	54	47%
Inteligencia de amenazas	53	46.1%
Talento Humano de Seguridad	45	39.1%
CyberRisk Quantification	44	38.3%
Ransomware de las Cosas	42	36.5%

Nota. Porcentaje calculado sobre las organizaciones que respondieron el ítem (base = 115). Pregunta de respuesta múltiple. Elaboración propia.

Nota. Porcentaje calculado sobre las organizaciones que respondieron el ítem (base = 115). Pregunta de respuesta múltiple. Elaboración propia.

Figura 3

Prevalencia de los temas clave priorizados para 2026

Nota. Porcentaje de organizaciones que seleccionó cada tema (respuesta múltiple). Elaboración propia.

Tabla 3

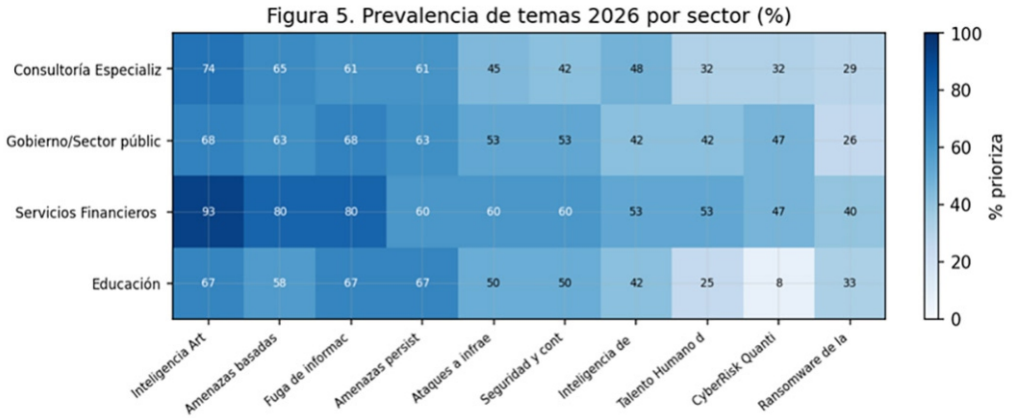
Tendencia de priorización temática según nivel de presupuesto

Tema	z	p	% PP bajo	% PP alto
Fuga de información sensible	-1.697	0.0897	80%	43%
Inteligencia de amenazas	-1.674	0.0941	70%	43%
Ataques a infraestructuras críticas	1.56	0.1187	45%	71%
Amenazas persistentes avanzadas	-1.536	0.1245	80%	57%
Internet de las cosas	1.351	0.1766	20%	43%
CyberRisk Quantification	1.219	0.2228	40%	71%

Nota. Prueba de tendencia de Cochran-Armitage. “% PP bajo” y “% PP alto” indican la prevalencia del tema en el nivel presupuestal más bajo y más alto, respectivamente. Ningún contraste alcanza $p < .05$. Elaboración propia.

Figura 4

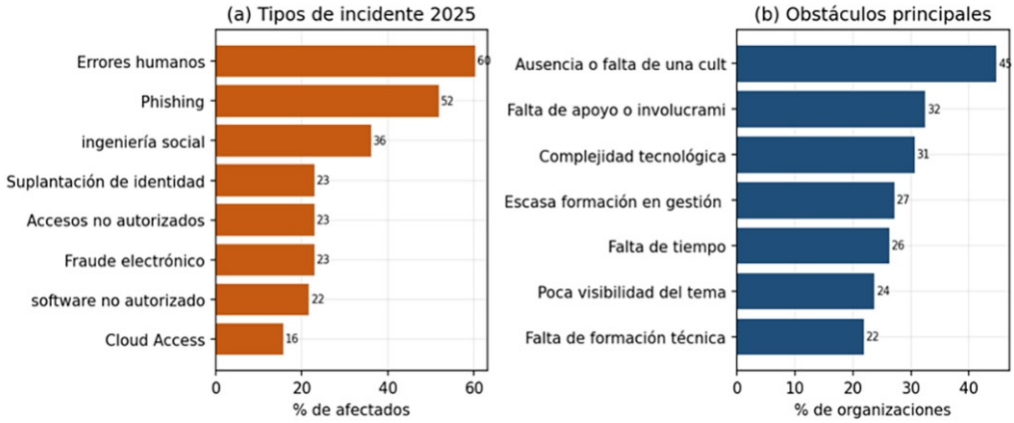
Prevalencia de temas 2026 por sector



Nota. Intensidad de color proporcional al porcentaje de organizaciones del sector que prioriza cada tema. Elaboración propia.

Figura 5

Tipos de incidente en 2025 y obstáculos percibidos



Nota. (a) Incidentes reportados por las organizaciones afectadas; (b) obstáculos principales percibidos. Respuestas múltiples. Elaboración propia.

Análisis de resultados

Convergencias. Los resultados convergen con la evidencia internacional más reciente en tres frentes. Primero, el protagonismo de la inteligencia artificial en la agenda 2026 coincide con el *Global Cybersecurity Outlook 2026*, en el que el 94% de los líderes encuestados identifica a la IA como el principal motor de cambio en ciberseguridad para el año y el 87% considera las vulnerabilidades asociadas a la IA como el riesgo de más rápido crecimiento (WEF, 2026). Esta convergencia se ve amplificada por Deloitte (2026), que pronostica que hasta el 75% de las empresas invertirá en IA agéntica hacia el cierre de 2026, intensificando el carácter dual de la IA como amenaza y como herramienta.

Segundo, el predominio del factor humano, errores humanos y phishing, como origen de los incidentes coincide con el *Cost of a Data Breach Report*, que atribuye al error humano una cuarta parte de las brechas y sitúa al phishing como vector inicial más frecuente (IBM, 2025).

Tercero, la fuerte asimetría presupuestal entre sectores regulados y no regulados refleja la inequidad cibernética descrita por el WEF (2026): mientras un 19% de las organizaciones ya supera los requisitos de resiliencia (frente al 9% del año anterior), las de menor tamaño tienen 2,5 veces más probabilidad de declarar una resiliencia insuficiente, profundizando una brecha estructural análoga a la observada entre los sectores regulados y no regulados.

Divergencias y matices. El estudio aporta matices que enriquecen la literatura. A diferencia de la narrativa habitual que vincula linealmente inversión y protección, los datos no evidencian un efecto protección significativo: el costo de los incidentes se explica mejor por la exposición sectorial y el tamaño que por el nivel de presupuesto. Esta divergencia, lejos de contradecir la importancia de invertir, sugiere que la eficacia depende de la calidad del gobierno y no solo del monto, en línea con el énfasis de NIST (2024) en la función Gobernar.

Asimismo, la posición rezagada del sector salud en madurez contrasta con su reconocido estatus como el sector de mayor costo de brecha a escala global (IBM, 2025), configurando un perfil de riesgo elevado que merece atención prioritaria. Finalmente, la baja conciencia directiva observada en parte de la muestra resuena con la divergencia de prioridades documentada por el WEF (2026), donde los directores ejecutivos ya sitúan el fraude habilitado por medios cibernéticos como su principal preocupación: el 73% reportó haber sido afectado por este fenómeno en 2025, mientras los CISO siguen centrados en el ransomware y la cadena de suministro.

Conclusiones

El estudio caracterizó la postura de seguridad de la información de 115 organizaciones y priorizó los cruces de mayor valor para la toma de decisiones. Tres conclusiones se destacan. Primera, la capacidad de gobierno y de inversión está distribuida de forma desigual: los sectores regulados lideran ampliamente la dotación presupuestal, lo que confirma una inequidad cibernética estructural. Segunda, el costo de los incidentes responde más a la exposición sectorial y al tamaño que al nivel de gasto, lo que desplaza el foco desde el cuánto se invierte hacia el cómo se gobierna la inversión. Tercera, la agenda 2026 está marcada por la inteligencia artificial y por amenazas que explotan el factor humano, lo que exige fortalecer simultáneamente la cultura, el gobierno y las capacidades analíticas en la gestión de riesgos.

Referencias

Agresti, A. (2019). *An introduction to categorical data analysis* (3rd ed.). Hoboken, NJ. USA: John Wiley & Sons.

Deloitte. (2026). *Technology, media & telecommunications predictions 2026: Narrowing the gap between the promise of AI and its reality.* Deloitte Center for Technology, Media & Telecommunications. https://www.deloitte.com/global/en/about/press-room/2026-tmt-predictions.html

International Business Machines Corporation (IBM). (2025). *Cost of a data breach report 2025.* IBM Security. https://www.ibm.com/reports/data-breach

National Institute of Standards and Technology (NIST). (2024). *The NIST Cybersecurity Framework (CSF) 2.0 (NIST CSWP 29).* U.S. Department of Commerce. https://doi.org/10.6028/NIST.CSWP.29

World Economic Forum (WEF). (2026). *Global cybersecurity outlook 2026.* World Economic Forum. https://www.weforum.org/publications/global-cybersecurity-outlook-2026/

Jeimy J. Cano M., Ph.D, CFE, CICA. *Ingeniero y Magíster en Ingeniería de Sistemas y Computación, Universidad de los Andes. Especialista en Derecho Disciplinario, Universidad Externado de Colombia; Ph.D en Business Administration, Newport University, CA. USA. y Ph.D en Educación, Universidad Santo Tomás. Profesional certificado como Certified Fraud Examiner (CFE), por la Association of Certified Fraud Examiners y Certified Internal Control Auditor (CICA) por The Institute of Internal Controls. Profesor Distinguido de la Facultad de Derecho, Universidad de los Andes. Es director de la Revista SISTEMAS de la Asociación Colombiana de Ingenieros de Sistemas –ACIS–.*

Andres R. Almanza J., Ms.C, CISM. *Chief Growth Officer en CISOS.CLUB, Investigador en Ciberseguridad SegInfo y Liderazgo. Executive Certificate in Cybersecurity Leadership & Strategy by FIU University. Certificado como ISO 27001 Lead Implementer and 27005 Lead Manager from PECB | CISM, ITILv3, LPI. Certificado como Coach Profesional Internacional, Master in Leadership and Organizational Development with Coaching, Executive Master's in Leadership Skills Developed in Harvard, & Coach Profesional avalado por International Coach Federation. Profesional en Ingeniería de Sistemas. Especialista en seguridad en redes y máster en seguridad de la información. Docente del programa de maestría de la Universidad Externado de Colombia y de la Universidad de las Américas en Ecuador. Creador de la Comunidad CISOS.CLUB, CISOS-COL y CISOS-LATAM (Linkedin) y Miembro del comité editorial de la revista sistemas de ACIS.*

Gabriela María Saucedo Meza, Ph.D. *Licenciado en Sistemas Computaciones y Maestría en Desarrollo Organizacional y Humano por la Universidad del Valle de Atemajac, México. Doctora en Educación por la Universidad Santo Tomás, Colombia. Certificada en Consultoría General por el Consejo Nacional de Normalización y Certificación de Competencia Laboral (CONOCER), México. Cuenta con más de 35 años de experiencia en gestión educativa, docencia e investigación en seguridad de la información, auditoría de TI, liderazgo educativo, cambio y cultura organizacional. Actualmente Coordinadora Académica de Posgrados de la Facultad de Contaduría Pública de la Universidad Externado de Colombia.*

Reduc@te 2026

5 AL 9 DE OCTUBRE DE 2026

REDUC@TE es un espacio creado para que educadores, ingenieros, diseñadores y entusiastas de las TIC descubran y compartan buenas prácticas, experiencias, tendencias y metodologías innovadoras en el punto donde la tecnología se encuentra con la educación y la formación. A continuación, presentamos las temáticas que abordará el evento.

TEMÁTICAS DEL EVENTO

- Neo Competencias y habilidades digitales
- Ética digital
- Evaluación alrededor de las nuevas tecnologías
- Casos de éxito

Confianza digital: ¿Innovación confiable?

Este número de la revista está dedicado a todos los asuntos relacionados al concepto base de la confianza digital.

DOI: 10.29236/sistemas.n179a5

Los directivos de la revista Sistemas Jeimy J. Cano M. y Andrés Almanza J. dieron la bienvenida a los invitados Rafael Gamboa Serrano de Data y TIC y Víctor Vásquez Mejía, director de IT Advisory en KPMG Colombia y formularon el primer asunto a tratar.

¿Cuáles son los elementos básicos que dan forma al concepto sobre confianza digital?

Rafael Gamboa

La confianza digital es la posibilidad que tiene una empresa, una persona, un estado digital, para usar la tecnología con una expectativa razonable de seguridad, transparencia y responsabilidad.

Se trata de un sistema confiable, es decir, saber quién está interactuando con la persona. Ahí hablamos de identidad, de autenticación, de controles de acceso. Un segundo elemento es, una vez yo identifico a la persona con la que estoy interactuando, es qué estoy recogiendo o para qué lo estoy recogiendo. Entonces, ahí aparecen normas, datos personales, la 1581 de 2012, pues que obviamente establece principios de finalidad, li-

bertad, transparencia, todos estos temas. Un tercer punto que tenemos es, si el sistema es seguro, y ahí ya obviamente entran el código penal, ley 1273-2009. Un cuarto asunto es si la operación es verificable, y ahí entra un tema mucho más técnico, y se trata de todos los temas de logos, trazabilidad, auditoría, evidencia, para poder determinar que lo que se dijo fue lo que ocurrió con la persona que dijo ser, y siempre decimos eso, y es que toda la parte jurídica siempre es en caso de alguna reclamación, entonces en caso de una reclamación, una queda, llámese penal o del titular, es donde cobra muchísima relevancia estos logs, estas trazabilidades para poder generar certeza al juez, a la fiscalía o a la misma contraparte, es que usted sí lo dijo, usted sí se conectó. Un quinto punto es qué pasa si el sistema falla, y ahí es donde se dispara todo lo que nosotros veníamos haciendo, y es ver quién va a responder por esa confianza. Y digamos que quién va a responder es el que puso el sistema a disposición de las personas y cómo va a demostrar que no incurrió en el error, sino que la otra persona fue el que incurrió en el error eventualmente, y qué herramientas se suple para poder demostrarlo. Y algo muy importante que no quiero dejar pasar muy jurídicamente aceptado, pero no es menor, y es que siempre que yo pongo a disposición un sistema informático, los jueces han reconocido que yo no puedo trasladarle el riesgo por su mal uso a los usuarios, valga la redundancia. sino que si yo pongo a disposición al sistema, yo tengo que responder por ese sistema, inclusive, como digo yo, contra la propia inversibilidad de los clientes. Entonces veíamos por eso es que cuando había una, toda la época del phishing, que decía, oh, es culpa suya, oh, usted se fue a engañar, oh, es que el sistema falló por allá en la porra, dicen, no es problema mío, usted me responde por ese sistema. Y eso es algo bien importante, sobre todo a la hora de poner a disposición sistemas informáticos. En otras palabras, un sistema confiable no es solo el que no se cae. Un sistema confiable es el que puede explicar lo que hizo. bien o mal, pero poderlo soportar.

Víctor Vásquez

Desde una visión de aseguramiento, todas las organizaciones se montaron en la ola de la transformación digital. Desde esa perspectiva ¿qué están generado? Servicios digitales, herramientas y eso al final, ¿qué es? Implementación de cualquier tipo de tecnología. Hoy el gran reto es cómo aseguro que todo lo que yo saque a producción para prestar un servicio o mejorar algo en cualquier organización lo puedo hacer de forma que sea confiable que sea segura y que tenga una serie de elementos. En el tema de la ética y transparencia se enmarcan muchos de los elementos que mencionaban anteriormente

que son como una especie de dominios que hay que trabajar para que definitivamente cualquier servicio aplicación o herramienta o tecnología que se implemente en una organización pues cumpla con eso. Que sea confiable para los diferentes stakeholders que la puedan usar utilizar o necesiten algo de esta tecnología en las organizaciones.

Andrés Almanza

Entonces ¿se podría utilizar confianza como un reemplazo sustituto de confianza seguridad y al revés? porque ambas posturas me llevarán a decir en algún momento que un sistema confiable podría transmitir la idea de qué es seguro y un sistema seguro podría transmitir la idea de qué es confiable y eso podría crear algún tipo de tensión en las organizaciones y podría también ser trampa.

Jeimy J. Cano M

¿La confianza digital es un concepto interdisciplinario? ¿Hay que convocar especialistas de diferentes áreas para darle sentido y profundidad al concepto aplicado en una organización?

Víctor Vásquez

En mi opinión hay que interactuar con diferentes personas en las organizaciones. Esto no es solo de seguridad, no es solo de tecnología, no es solo de auditoría, sino que ahí tienen que intervenir varias funciones dentro de una organización, dentro de un marco organizado, para poder implementar cualquier tecnología. Inclusive ISACA propone que un comité específico de Digital Truth, donde intervengan diferentes áreas de la organización. Y definitivamente sí es interdisciplinario.

Rafael Gamboa

La confianza digital no la construye una sola disciplina, no lo podría jamás constituir una sola disciplina, ni únicamente los ingenieros, ni solo los administradores, sino es una conjunción de elementos para donde tenemos. Entonces, procurando no extenderme mucho, digamos, ahorita que andamos mucho con tema financiero, uno dice, oiga, ¿cómo es el sistema? O sea, cuando uno dice el sistema, ¿qué es? Es que el sistema por quien está constituido. ¿Quién se inventó la arquitectura? ¿Quién la diseñó? Un in-

geniero, clarísimo. Pero el abogado es el que tiene que revisar la legitimidad, digamos, del tema de los datos, del tema contractual, de los impactos que tiene ese sistema o el cumplimiento normativo. Así no haya norma, ¿no? que eso es algo bien importante, no sé, solo con el tema de la inteligencia artificial, no está regulado, y yo soy de los que digo que jamás va a estar regulado la inteligencia artificial, lo cual no quiere decir que no vamos a utilizar esa herramienta, porque la vamos a seguir utilizando, sin lugar a dudas, en Colombia y en todos los países hay regulación de inteligencia artificial, y que un abogado diría, no la puedes utilizar, eso no va a pasar, o sea, claramente lo vamos a seguir utilizando, entonces el rol del abogado un poco es Cómo ser ese puente, cómo ser ese, cómo si lo hacemos con algo muy importante y es lo que se conoce como la mitigación del riesgo. Pero ese es otro elemento que se lo adjudica otra persona al gobierno. Entonces tenemos el ingeniero, el abogado que revisa y el experto en seguridad informática identifica las amenazas.

Porque por cuenta del halo del UX, como nos dicen a las abajadas, es que usted no le puede preguntar tantas cosas porque afecta al UX.

Entonces uno es como, ¿usted quiere engañar a la persona? Porque más allá de llevarlo a un tema va a generar una... desconfianza en el sistema en el producto en el proveedor y eso es un tema muy delicado que debe estar alineado el experto de el diseñador de experiencia el quinto importantísimo es el gerente es el que define el apetito del riesgo el callo que tenga ¿sabe qué? no fuimos me importa un carajo voy a decir que esto hace y deshace.

Y usted ingeniero, hágame caso, abogado me lo paso por la galleta, el experto en seguridad informática no me importa porque yo lo que quiero es esto. Y ese es el que en últimas va a decir qué tanto voy a hacerlo o va a ser el gerente que dice no, yo tengo que cumplir y hasta que no haya una ley específica de la República de Colombia o de Venezuela o de Panamá donde diga cómo puedo utilizar los sesgos algorítmicos de la inteligencia artificial, no voy a hacer nada.

Pero está el negocio por otro lado. Y finalmente, la alta dirección que decide si la innovación está alineada con la promesa de valor de la empresa. Oiga, nosotros sí vamos a hacer esto. Nosotros estábamos en un banco las semanas, 15 días, que decía, si mi cliente tiene la posibilidad de tener una mejor oferta, que se vaya para otro lado, porque yo soy un banco enfocado en social y si yo no tengo la capacidad ni las herramientas técnicas, los dejo libres, y uno dice, chévere, y otros bancos dicen, no, yo voy a retener a este cliente de patas y manos, en-

tonces, claro, un error es creer que la confianza digital solo es un tema tecnológico, no lo es, es un sistema, hace parte de todo lo que se tiene, y básicamente lo que uno dice es, no sé, en un tema de scoring, que en este tema de scoring el ingeniero es el que dice qué modelo predice bien, el abogado es el que dice si los datos fueron escogidos legítimamente, el de negocio dice cómo pudo mejorar sin aumentar el riesgo y el de experiencia dice si entendió por qué fue rechazado o no y finalmente cómo puede mostrar que no hubo una discriminación en un tema tan complejo donde hay una norma y una tendencia que se llama la transparencia algorítmica donde la Corte Constitucional ha dicho, usted tiene que decirme por qué llegó a ese resultado. Y cuando miramos ese resultado de un scoring de crédito, es súper complejo porque hay un know-how, hay una... Y lo decíamos, obviamente, sin nombres, es porque a mí no me gustan darle crédito a los abogados. ¿Por qué? Porque yo sé que los abogados son mala paga. ¿Es legítimo? Sí. ¿Es legal? Sí. En mi negocio, yo soy un profesional y tengo que proteger de la mejor manera recursos que no son nuevos, que no son míos. Entonces, pues la confianza digital no es solamente código, no se vive solamente el código, sino cómo ese código afecta derechos, decepciones y expectativas humanas. Es como lo que uno tiene ahí en su concepto, entendiéndolo como un todo.

Jeimy J. Cano M.

Entonces el tema de la transparencia algorítmica hasta cierto punto es válida, pero cuando hace elaboraciones muchísimo más profundas y mucho más detalladas, pues posiblemente no vamos a entender ni vamos a poder seguirle el paso a cómo se ejecutan esas redes neuronales de aprendizaje profundo no sabemos exactamente cómo lo van a hacer. Hay momentos en que se nos pierde, se nos pierde completamente, desde el punto de vista técnico.

Andrés Almanza

En mi opinion surge otra pregunta que queda ahí en el aire y es qué sería de las cosas, qué tendríamos que pensar de cara a un futuro. O sea, porque apenas escuchándolos empiezo a decir, ok, sí, estamos en el presente, pero el futuro nuestro, yo estoy de acuerdo

con los dos. La transparencia algorítmica se va a perder, sobre todo Jeimy que lo menciona como muy explícitamente. se va a perder y se va a seguir degradando en la medida en que sigan creciendo nuestras redes neuronales, a punto en donde en algún momento tendremos que pensar si vamos a prescindir de la transparencia inclusive para poder seguir hablando de confianza o si no lo vamos a tener que pensar. De hecho, una de las cosas en la misma línea de Rafael es que si uno no le hace mantenimiento al algoritmo, el algoritmo se degrada.

De hecho, eso es una de las recomendaciones cuando uno tiene inteligencia artificial por lo menos generativa funcionando. Si uno no le hace mantenimiento al algoritmo, el algoritmo cada vez más va a responder menos bien, con mayor precisión con menos precisión cada vez más se va degradando el algoritmo; como diría Stephen Covey si no se afila la sierra pues no se corta bien el hacha aquí es igual entonces eso es un tema que la transparencia pone realmente un reto a la confianza digital muy importante. Porque uno quisiera que realmente la inteligencia artificial operara en un acuario, para que uno pudiera ver cómo los algoritmos están funcionando completamente hasta el final. Pero no. Llega un momento en que se nos pierden, por el solo el hecho de ejecución. O sea, inclusive los señores de OpenAI y los de Antropic dicen que llega un momento en que nosotros no sabemos por qué responde la inteligencia artificial así. No sabemos; entonces eso es un tema realmente espinoso espinoso cuando estamos hablando de confianza digital bueno,

Jeimy J. Cano M.

¿Existen marcos de trabajo concretos para desarrollar el concepto de confianza digital? ¿Conocen ustedes alguno que se aplique en Colombia?

Rafael Gamboa

Así como aquí regula todo, no, no hay, no hay una única ley de confianza digital, como se puede decir, lo que sí hay son normas que regulan las distintas etapas de esta confianza, camas, llamémosla así, entonces, pues la primera, datos personales, sin duda, todos son datos, ley 1581 del 2012, decreto 1074, que es cómo puedo manejar, cómo debo manejar los datos, que en últimas es un generador de confianza para los titulares. Una segunda capa sería, sin duda, el tema del data financiero, un tema reputacional, que ya nos brincamos a la 12.66 de 2008 para temas de obligaciones dinerarias. Una tercera capa, el tema de la seguridad de la información, Ley 1273-2009, cómo penalmente se está protegiendo todo ese tema, cómo se protege el tema del acceso abusivo, del daño informático, interceptaciones, violación de datos personales, porque es que acordémonos que esta ley sur-

ge más allá del convenio UDAPES, surge es porque Cuando empezaron este tipo de reclamaciones, la respuesta del juez dijo, y el fiscal dijo, este es un ciberdelito. Eso no está tipificado, eso no es un delito. Y por eso es que nos tocó correr a sacar eso. De hecho, yo siempre utilizo esta ley para demostrar cómo los hechos generaron la norma del derecho. Los hechos generaron el derecho, a diferencia de la ley 527, donde el hecho de tener una ley de comercio electrónico no generó automáticamente ley de comercio electrónico. La explosión. La explosión, exactamente. Siempre digo, nosotros tuvimos ley de comercio electrónico antes que Estados Unidos, y eso quiere decir que tenemos mejor comercio electrónico. La respuesta es no. Entonces aquí la 1273 surge por una necesidad de tipificar actuaciones de phishing, de hacking, de todos esos temas que se venían desarrollando pero no estaban tipificados. Una cuarta etapa es el tema de las evidencias y transacciones digitales. Claro, ahí sí nos remontamos a nuestra antiquísima ley 527 del 99, que fijó un marco interesante. Esta ley, yo siempre digo que ya tiene más papás que quién sabe quién. Lo cierto es que aquí fusilamos esta norma. Nos trajimos una serie de conceptos que siguen teniendo vigencia y que nos siguen estableciendo unos principios bien importantes sumado a decisiones jurisprudenciales donde dice, oiga, en últimas a mí lo que me tiene que generar es convencimiento al juez o a la persona que está haciendo. O sea, más allá de lo que de lo que. La prueba tecnológica es generar convencimiento a la contraparte o generar convencimiento al juez, al juzgador, al fiscal. Y finalmente una quinta capa, que por supuesto siempre hay que mencionarla, y son los CONPES. Son los CONPES el 3995 sobre confianza jurídica. Y seguridad digital o el decreto 767 de política de gobierno digital, que muy seguramente va a venir a cambiar, pero estos compes dicen por dónde nos vamos a mover. Fíjense que este Conpes trae una cantidad de cosas que se vienen desarrollando inclusive hace un mes nos terminaron sacando un decreto de Open Finance aunque el compes hablaba de Open Data terminaron aterrizando en Open Finance porque eventualmente lo que podían hacer entonces para responder en forma concreta. Si hay marcos aplicables en Colombia, lo que más falta muchas veces es poderlo integrar. Articulación es la palabra.

Jeimy J. Cano M.

Ok. Sí, interesante. Pues aquí hizo Rafael una enumeración de, digamos, de los recursos jurídicos que tenemos hoy en Colombia, ¿no? Sí, desde protección de datos personales, avias data, la ley 1273, todo el tema de la 527 y claramente los CONPES donde, digamos, se han articulado algunas de estas iniciativas a nivel nacional. Intere-

sante. Como arquitectura, ¿no? Como arquitectura legal alrededor del tema.

Víctor Vásquez

ISACA tiene o diseñó un marco específico que nos habla de confianza digital, que es el DTEV. Nace más o menos como en el 22, y lo que nos habla es de realmente un ecosistema digital confiable considerando como cuatro vertices, que son las personas, los procesos, la tecnología, la organización, y que logra integrar atributos claves para la confianza que los hemos mencionado ahí, pero mire que aquí los integra, ¿no? De una forma como organizada, habla de seguridad, privacidad, integridad, resiliencia, calidad, confiabilidad.

Entonces es un marco que busca no reemplazar lo que hay, sino integrar, ¿no? Cómo integrar lo mejor de NEEDS, cómo integrar COVID, cómo integrar, no sé, la 27.000, o esta regulación que se mencionaba a nivel local, porque a nivel local no hay nada explícito diseñado específicamente para eso, ¿no? Hay muchas, como mencionaba, normatividad que en conjunto nos puede ayudar.

Pero un marco como estos fue diseñado para eso, ¿no? Para integrar y tener una visión más holística de lo que es confianza digital en un ecosistema digital para generar confianza, ¿no?

Jeimy J. Cano M.

Entonces, yo creo que la palabra clave aquí es ecosistema, es decir, muchos actores que interactúan entre ellos para generar de alguna manera valor para aquellos que están dentro del ecosistema. Entonces, eso es como conectando un poco con la arquitectura que proponía Rafael, casi que eso tiene que articularse ¿no? o sea por ejemplo el modelo que presenta Isaac que es el DTEF que tiene todo este montón de elementos y componentes pues de alguna manera es una visión mucho más amplia y que de alguna forma cuando Rafael integra y pone toda la arquitectura legal detrás pues de alguna forma cobra muchísimo más dinámica ¿sí? porque está detrás no solamente el desarrollo de la iniciativa digital sino la responsabilidad por esa iniciativa digital que es lo que eleva el nivel de confianza final-

mente en algunas organizaciones le preguntan a uno siempre con el tema de inteligencia artificial y quien responde. Si eso falla, ¿quién responde? Que es un poco lo que precisamente comenta y articula Rafael en su respuesta.

Entonces miremos como los dos lados que hemos venido como revisando. No sé, Andrés. Ahí lo único que se me ocurre, porque creo que la palabra clave es ecosistema, y de pronto lo que yo agregaría es que en este momento ese ecosistema, para que sea mucho más confiable, articulando todas las cosas escuchadas es no existe un elemento que pese más que otro para poder desarrollar esa confianza es lo que me traigo con todas estas cosas entonces estos ecosistemas confiables son elásticos entre todos sus componentes y de todos va a depender que esa confianza aumente para producir más valor y con que uno solo de sus elementos, que sería la otra cosa, el lado negativo, si lo quisiéramos llamar, con que uno solo de sus elementos no tenga el soporte suficiente, va a generar efectos, cascada, que va a hacer que ese ecosistema en total, puede que una de sus aristas del ecosistema esté funcionando uno a la otra, si no está funcionando adecuadamente, genera esa diatriba de no hay confianza del ecosistema y por ende va a tener algún problema. La pregunta ahí podrá ser, ¿eso será suficiente para la sostenibilidad de una empresa en un entorno digital como el que tenemos y vamos a tener? El tiempo lo dirá. Muy interesante esto que hemos venido conversando. Bueno, seguimos con la siguiente pregunta. Ahora sí, centrado en el tema de inteligencia artificial y con incorporación de la inteligencia artificial en prácticamente todos los escenarios de las organizaciones y la sociedad. ¿Cómo juega la confianza digital en este nuevo escenario?

Víctor Vásquez

Pues hoy en día se vuelve más relevante hablar de confianza digital y con esos dos vértices, la ética y la transparencia, sobre todo cuando hablamos de inteligencia artificial que si no está, como decía Jamie, probada periódicamente, evaluada, alineada, pues puede generar muchos resultados no deseados, porque lo pienso para... para algo, pero si no le implemento como sus límites, su marco, el tema, hoy se está hablando mucho del gobierno, cómo gobernar la IA, o sea, cómo llegar a implementarla de forma más confiable en las organizaciones para ayudar a que se obtengan los objetivos que quiere la organización.

Entonces definitivamente se vuelve más relevante hablar de confianza y hablar específicamente con la idea de transparencia y de ética, ¿no? Cómo usamos éticamente las diferentes herramientas y que vemos que esto va evolucionando

muy rápido y hablamos de algo... Y ya en ocho días ya el panorama evoluciona muy, muy, muy rápido.

En eso encaja muy bien el marco de ISACA y prácticamente hace dos años ya se está desarrollando y mostrando como este marco podía ser utilizado para empezar de forma organizada a trabajar con este tipo de tecnologías emergentes que ya hoy en día se volvió más masiva ok, solo una pregunta adicional Víctor, ahí en el marco de ISACA de alguna manera se detalla los componentes por cada uno de ellos y de alguna manera como las métricas, pues, muestra unos dominios que hay que abordar, básicamente, y como ya hay varios papers de cómo utilizar este marco para las implementaciones de IA de forma confiable, pero al final es hacerlo con un gobierno, con una organización, con un método.

Jeimy J. Cano M.
La ventaja de este marco es eso, que da como unos lineamientos, unos dominios, unas áreas, unas funciones que hay que contemplar para implementar la IA o cualquier tecnología.

Rafael Gamboa
Claro, ustedes saben, y Jamie que me conoce tanto, y Ricardo, yo soy ingeniero wannabe, pero uno dice, claro, una información jurídica, pero ¿cómo funciona? Entonces, claro, uno dice, pues, es que antes la tecnología almacenaba y transmitía información, y me acuerdo muchos años en que decía, oiga, me va a quitar el trabajo la firma digital.

Y me decía, no, pues no, internet me lo va a hacer el trabajo. Ahora no solo almacena ni transmite, sino que recomienda, previse, clasifica y hasta decide. Entonces, claro, uno coge al ingeniero y los ingenieros dicen, oiga, ¿y ya? Ya no, pues ya, pues fácil, va a sacar dependiendo de los datos, del modelo, de los parámetros, del entrenamiento, de la inferencia y el monitoreo. O sea, Uno más uno es dos.

O sea, más o menos va a salir ahí porque son elementos muy, entre comillas, tangibles y claros, aunque no lo son. En cambio, usted le pasa la misma pregunta al abogado y le dice, es que la respuesta de la IA tiene que tener en cuenta si los datos eran legítimos, si eran ciertos, lo que se llama la contaminación de los datos, ¿quién revisa eso? Jurídicamente lo revisaron, nadie lo revisa. Oiga, ¿la finalidad fue informada a las personas para que lo utiliza? Si hay un modelo que eventualmente pueda discriminar, así sea legítima esa discriminación, si lo había, si estaba identificado y las empresas contrataron con esos parámetros y quedaron así en el contrato. Si el resultado es explicable, lo que pretende la Corte Suprema, la Corte Constitucional con la transparencia algorítmica, usted tiene que decirme por qué me rechazaron el crédito, si hay intervención

humana o no, o si todo fue automático, porque lo que dice la Corte Constitucional es donde haya un byte de decisión automática, usted tiene que decirme por qué lo explicó. ¿O cómo llegó a esa conclusión? ¿Y quién va a responder en caso de error? A mí no me vaya a decir que es que, que jueque, que jueque, que no. Usted es el que tiene que responder. Que jueque, que jueque, que usted es el que tiene que responderme. Y finalmente, ¿cómo vamos a auditar la decisión? La transparencia está muy en línea de la auditabilidad de las decisiones a las que llega.

Y hay unas decisiones vía revisión constitucional que está diciendo que cualquier entidad jurídica pública que maneje recursos públicos, público privado que maneje recursos públicos, tienen que tener la capacidad de poder explicar y auditar y ser auditables las decisiones, lo cual se vuelve súper difícil. Entonces, en Colombia hay alguna regulación puntual. No, no la hay. Digamos que todos los proyectos están en línea de lo que dice Europa. La Unión Europea, lo que es el reglamento europeo de la IA. Y uno dice, ¿por qué? Porque allá son Dios y allá saben todo. Y uno dice, no, el efecto Bruselas es absoluto no, porque claro, en Latinoamérica consumimos mucha normatividad europea, digo yo, por el idioma, porque nos llega todo a través de España, pero lo cierto es que allá en Europa ya le levantaron la mano y ya hay una serie de proyectos conocidos como Omnibus Digital, donde para el tema de datos personales y para el tema de regulación de inteligencia artificial, dice, oiga, no me ponga ese... Un universo fantástico Disney, porque es que se vuelve imposible negociar.

Entonces, cuando uno mira el mercado, uno dice, ¿quiénes son los líderes? China y Estados Unidos. ¿Quiénes son los líderes en inteligencia artificial? China y Estados Unidos. ¿Quién es el que más regula? Europa, que ni siquiera es líder, que ni siquiera tiene la potencialidad de ser líder por temas presupuestales y por temas de desarrollo. Entonces, ahí es donde uno sí tiene que aterrizar desde el punto de vista latinoamericano, cómo es que vamos a manejar. Entonces, ¿cómo se maneja el tema de la confianza? Lo mencionábamos anteriormente con temas de protección de datos, avías data, seguridad digital... Los compes, etcétera, etcétera. Entonces ahí es donde uno dice el reto no es que el modelo responda. A mí no importa que el modelo funcione o no funcione, sino a mí como organización lo que más me importa es que pueda responder por el modelo, saber responder. ¿Por qué dijo uno o dos? ¿Por qué fue que lo dijo? Ese es, digamos, el gran reto y desafío y de ahí no nos van a sacar a los jueces y de ahí no nos van a sacar a los abogados. Yo a tener que explicar algo que eventualmente es inexpli-

cable, como decía ahorita Jamie, y es que yo ni siquiera sé por qué llegó, porque hay muchísimas concusiones de información y si a esa data original o real le sumamos todos los sistemas de alimentación de data sintética, pues va a ser imposible, no difícil, imposible.

Entonces, pero tenemos que tener claro que los jueces no se van a mover de ahí. Van a decir usted tiene que explicarme y el desafío es yo como organización, cómo voy a poder demostrar o convencer al juez o al supervisor regulador que sí sé lo que está haciendo mi sistema y por qué lo está haciendo.

Jeimy J. Cano M.
Aquí se hace evidente el tema de la confianza ese tema de que hay un tercero que tiene una identidad no humana que está ayudando a hacer cosas y que de alguna manera respondiendo a una programación respondiendo a unos datos propone elabora, diseña y como decía Rafael hasta decide Y ahí sí, claramente, cuando hablamos de decidir sobre un campo específico, pues claramente los efectos se tienen que colocar sobre la mesa con claridad, porque lo que está en juego es algún tercero o algún derecho que se vaya a vulnerar.

Andrés Almanza
Fundamental lo que yo creo que también rescato, porque me conectó mucho las palabras de Víctor. Y tomando tus palabras, Jeimy, el tema de gobernar, ¿se va a volver o se vuelve para fortalecer esa confianza? O sea, un instrumento que sí o sí no podemos, dado la IA, necesita hoy con tantas capacidades que puede desarrollar, inclusive decidir, necesitará, no digamos que un bozal freno, pero yo veo lo que menciona Víctor, como el instrumento que sí puede, sí o sí, han coincidido casi todos los expertos, es decir, la IA para que sea estructurada, eficiente, pueda tener unas implicaciones y no tenga tanto dolor como a veces lo vemos, necesita ser gobernada. como un instrumento fundamental que va a apalancar esa confianza dentro de todo el ecosistema entonces yo creería que habría que hacer un resalto de gobernar la IA va definitivamente a apalancar como esa confianza que va a tener que plasmarse en todo el ecosistema si de hecho hay una frase que dice recientemente dice el problema no es incorporar la inteligencia artificial sino como gobernarla y a que velocidad

Jeimy J. Cano M.
Y a qué velocidad, correcto. Y a qué velocidad. Y eso es realmente el gran reto que hay de aquí detrás, ¿no? O sea, y el marco que la debe cubrir precisamente para poder habilitar todas sus posibilidades que podemos tener de aquí en adelante. Entonces, es realmente interesante, digamos, las dos posturas, muy en assurance, muy en legal, pero tienen una línea de convergencia, ¿no?

Y vamos ya para ser juiciosos con el tiempo a la quinta y última pregunta. Dice, ¿qué recomendaciones darían a los ejecutivos de tecnología y gerentes de empresa para apropiarse, es una palabra bien importante, del concepto de confianza digital? ¿Cómo mostrar que es un concepto clave y relevante para la promesa de valor digital de la empresa.

Rafael Gamboa

Hay una serie de recomendaciones aunque suena pero sí temas a tener en cuenta lo mejor que obviamente lo digo desde el punto de vista de abogado con un algo de conocimiento técnico. Y en últimas, de los riesgos que tienen las organizaciones, porque insisto, un abogado sobre todo es el enfoque litigante y lo que hemos visto siempre en las organizaciones, los contratos, todo nuestro enfoque es cuál es el real riesgo. A mí donde me pongan en un contrato que es que si me incumple se va a ir a la cárcel la resta de la vida. Yo no me mato por eso porque es una cláusula claramente y absolutamente ilegal que en el escenario judicial pues nunca va a prosperar. Pero aquí sí es importante entender varios puntos. Uno, dejar de creer que la confianza digital es un asunto, está en el aspecto legal, es decir, que el contrato es lo único y que va a definir, o sea, que la confianza digital es un asunto de cierre, legal de cierre, es decir, que yo hago un contrato y que con eso ya fue suficiente sin fijarme en las otras condiciones del producto, cómo está construido, cómo está la tecnología, porque es que el papel aguanta todo y eso hay que tenerlo súper claro. Si a mí me ponen a escoger entre tecnología y papel, pues casi que prefiero es la tecnología que el papel yo como hago para evitar que cumplan las lo iba dudando lo iba dudando no no no yo soy super ayatol en eso digo si y con Jeimy hemos compartido muchos foros donde a uno le preguntan es que me están negando mi derecho a la información porque me tienen bloqueado internet yo digo bloqueelo es el mejor sistema para que le cumplan su política de internet o sea ¿Que eso es ilegal o es ilegal? No sé, es un tema donde yo, si yo no implemento herramientas tecnológicas y ocurre algo, lo primero que me van a decir es ¿por qué no implementó herramientas tecnológicas? Y me van a calificar de negligente por no haber utilizado esas herramientas tecnológicas. Y ese es un tema súper, súper sensible, pero real. No solo ver lo que dice la norma fría en un papel, sino ¿qué pasa si? Un segundo punto es documentar las disecciones técnicas. Eso es importantísimo, uno lo ve todos los días, es, oiga, ¿usted por qué decidió eso? ¿Quién lo decidió? ¿Para qué lo decidió? Nadie tiene ni idea y ese es un leak, esa es una falla que siempre las organizaciones tienen, es que no documentan todas las decisiones mediante logs, mediante evaluaciones de impacto de

privacidad, bitácoras de cambio, nada de eso, nunca queda documentado y ese es un tema que cuando ocurre se vuelve supremamente relevante. Tres, hacer privacidad, seguridad e inteligencia artificial desde el diseño. Trabaja desde el diseño, no, sino cuáles son los requerimientos funcionales, cómo vamos a minimizar el tema de los datos, el control de acceso, cómo vamos a manejar eso desde el momento cero, cómo lo vamos a implementar desde el momento mismo del diseño, como su mismo nombre lo dice. Un cuarto punto es traducir cumplimiento a controles técnicos si bien dijimos que no todo va a estar regulado hay normas y esas normas hay que cumplirlas gusten o no, hay normas entonces como decía un profesor de procesal en la universidad la ley tiene vacíos pero el derecho no, el derecho tiene que entrar a regular esa realidad de la inteligencia artificial tiene que entrar a regular esos vacíos o esas incertidumbres tecnológicas y ahí es donde entran elementos como debida diligencia, negligencia experiencia, profesionalidad que es algo que utilizamos doctor Jamie usted es un profesional usted tiene que saber así no se le ocurra usted es el profesional yo no sé y es usted el que tiene que saber entonces eso es algo bien importante y que toda esta semana hemos estado precisamente en ese tema diciéndole a unos profesionales que usted no me diga que no sabe porque usted entre los dos usted es el que sabe entonces es un tema bien importante de entender esa traducción de normas profesionalidad a cómo lo manejan lo manejan medir la confianza, entender cuál es el apetito del riesgo de la organización. Es importantísimo.

Aquí no trabajamos en escenarios dignos, sino cómo lo vamos a manejar. Y algo que suena casi que obvio, pero que nadie utiliza, y es creen un comité pequeño, tres personas, que hagan seguimiento y actualicen todos los temas. ¿Por qué? Porque es que las cosas se firman y nadie volvió a ver ese proyecto, nadie volvió a manejar esa herramienta, y lo que sí nos ha demostrado la experiencia es que cuando hay un seguimiento, no digo diario, semestral, quincenal, no, semestral es mucho, quincenal o mensual, sí podemos advertir los riesgos, porque están cambiando todo el tema y en inteligencia artificial, lo que era la locura, o sea, es que lo hablábamos en Open Finance, hace tres años nadie hablaba de Open Finance, perdón, nadie hablaba de inteligencia artificial, hoy por hoy nadie deja de hablar de inteligencia artificial.

Entonces, claramente los ingenieros son los que construyen los sistemas, pero son los abogados los que ayudamos a que estos sean sostenibles, defendibles y legítimos. Es la única manera en la que podemos procurar utilizar una herramienta que no va a estar regula-

da, pero que no podemos hacerle el quite. Uy, yo ahorita, Rafael, yo creo que esas se las voy a robar de esas tres palabras que acabo de decir, que son claves, sostenibles, defendibles y legítimos. La inteligencia artificial tiene que cumplir esos tres.

Y en ese sentido, los abogados tendrán que trabajar de manera interdisciplinaria, claramente ingenieros y todos, alrededor de esas tres palabras para darle carnecita. Es decir, qué significa defendible, qué significa sostenible y qué significa legítimo.

Hay muchas disciplinas conjugadas para darle... digamos marco a ese ejercicio final de de cómo se llama de esas recomendaciones casi que que es el las recomendaciones para el ejecutivo oiga ponga en su comité a trabajar alrededor de estas tres palabras y de la forma para que no para que nosotros podamos decir mire cuando hablamos que nuestra inteligencia artificial es sostenible, defendible, legítima hablamos de esto eso es como un marco casi que de debido cuidado guardada proporciones o no Rafael, tal cual.

Y no son tomas jurídicas y no son tecnológicas, son transversales. Las dos le aplican en distinta óptica, pero le aplica definitivamente en la necesidad técnica y en la necesidad jurídica. Interesante, muy interesante.

Víctor Vásquez

Darle como un nivel, llevarlo a nivel estratégico, ¿no? Pensando que eso habilita crecimiento, temas de reputación, sostenibilidad, entonces hay que involucrar eso en el plan estratégico, en el gobierno corporativo, en las agendas de las juntas directivas, del comité de auditoría, o sea, tiene que subir de nivel, ¿no? Y ser top down, ¿no? Desde arriba hacia abajo.

Lo otro es que hay que ponerlo también, porque nos ponemos muy técnicos, y a esos niveles se pierden, ¿no? Entonces hay que hablarle en términos de negocio, ingresos.

Si los clientes confían, pues van a pedir más o a cobrar. me va a ayudar pues en temas de negocio, ¿no? Me compran más y pueden que con esa confianza pues sigan con mi organización.

En temas de riesgo, pues sí, si tengo menos riesgos, pues tengo menos pérdidas, sanciones. En temas de cumplimiento, pues voy a tener menos multas. En temas de reputación, pues me va a mejorar la ventaja competitiva.

En innovación, pues voy a tener más confianza dentro de mi organización para implementar nuevas tecnologías como la IA. Lo otro tiene que ser un tema integral, porque a veces todos estos temas de tecnología se piensa como en TI,

cuando se piensa en aseguramiento, se piensa en auditoría, y se piensa como aisladamente, ¿no?

Entonces, esto es un tema integral, ¿no? Que contempla varios... varios dominios y que no puede ser no puede ser aislado solo la responsabilidad para TI y finalmente pues digo los marcos pues son para utilizarlos entonces hay que lo llamaría a los ejecutivos y hay una documentación muy sencilla, muy fácil para ejecutivos de lo que es este marco de Digital Truth, entonces también pues leerla y no inventarse la rueda, porque queremos inventarnos la rueda y ya varias personas a nivel global se sentaron, pensaron, hicieron los planteamientos, entonces hay que utilizar eso, ¿no? No inventarnos la rueda, sino tomar eso y mirar cómo lo adapto a mi organización con algo lógico, ¿no? Un método de cómo hacerlo más fácil sin inventarme algo nuevo, ¿no? Ya está definido, pues utilicémoslo, ¿no? Y leamos más el tema, ¿no?

Jeimy J. Cano M.
Entonces, mire, yo creo que aquí, y quiero retomar y cerrar con esta pregunta, porque es un concepto que lo introdujo Rafael, que era el tema del apetito de riesgo. Y ese apetito de riesgo creo que podría ser como el elemento base para poder llegar a la confianza digital.

Es decir, en la medida en que yo tengo claro cuál es mi apetito de riesgo, y tengo claro qué capacidades tengo para responder frente a ese apetito de riesgo, pues así voy a construir mejor mi confianza digital.

Rafael Gamboa
Nosotros lo hablamos formal e informalmente con los clientes, decimos, oiga, ¿cuál es su apetito del riesgo? ¿Cuál es su ADN de la organización? Que también es cierto, hay organizaciones, hay empresas que son muy jurídicas, y si nuestra ley no está, hay unas que son absolutamente comerciales, diría que la mayoría... Y hay otras que son muy tecnológicas y dicen, pues yo lo hago, pero lo cierto, y siempre hablamos en la oficina del trilingüismo, de entender de dónde saca la plata el negocio, cuál es el valor de la tecnología, entendida que la tecnología solo es útil si aumenta utilidad y reduce costos, y cuál es el marco jurídico, es decir, dónde nos vamos a mover. Y una consideración final que dicen, no, yo soy muy respetuoso de la ley, todos somos respetuosos de la ley, pero desde el punto de vista jurídico, legal, nadie debería tener ni nube ni redes sociales. ¿Por qué? Porque lo que se establece en los términos de uso de la nube o de las redes sociales, cuando estamos hablando de los grandes proveedores AWS, Azure, Huawei, lo que quiera, dice que cualquier disputa, legislación extranjera. Y es lo mismo que dicen las redes sociales. Entonces, desde el punto de vista

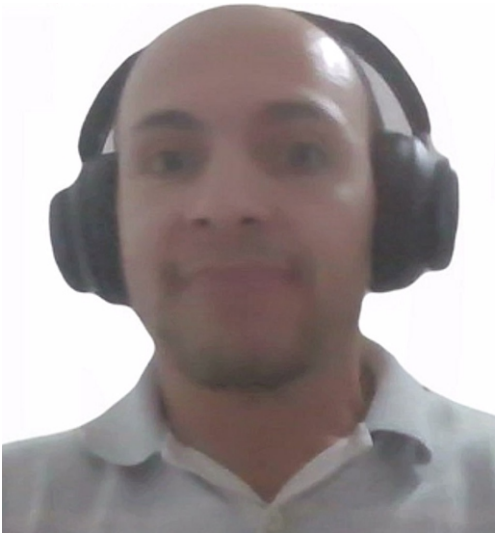
legal, exegético conservador, uno no debería tener ni nube ni redes sociales. ¿Quién no va a tener nube ni redes sociales? Nadie. El mismo Estado dice, ni loco yo voy a hacer eso, voy a dejar de utilizarlo porque es una necesidad. Entonces, ¿cuál es mi riesgo? Cuando hablamos de entidades públicas, decimos mi objeto misional, ese es mi riesgo, mi objeto misional se potencia con redes sociales y con nube y cuando hablamos de entidades privadas mi objeto misional es generar utilidades. Entonces yo no voy a tener todo on-premise, súper inseguro, cuando puedo tener muy buenas alternativas comerciales y lo que hago es mitigar el riesgo. Yo como compenso el apetito, lo compenso con la mitigación del riesgo, mitigación técnica, mitigación jurídica, información a los eventuales afectados y a las autoridades, y fue básicamente lo que hizo la superfinanciera que tuvo la oportunidad de trabajar hace varios años cuando expidió la circular de tercerización de servicios dijo señores llegaron a regular una situación que ya existía y era que todo el mundo ya estaba subido en la nube pero la superfinanciera dijo oiga vigilados supervisados usted quiere tercerizar servicios pero dentro de las muchas cosas que debe tener de seguridad claridad arquitectura es que en caso de toma de control

Víctor Vásquez
Yo como supervisor, usted va a firmar Banco X con un proveedor, va a firmar un otro, donde me permita a mi supervisor acceder a su datacenter, a su información, inclusive en su contra. Entonces hoy por hoy todos tienen que firmar cuando contratan con un proveedor de nube, tienen que firmar y negociar con el proveedor de nube que en caso de toma de control le deben dar acceso a la superfinanciera.

Y eso es simplemente la mitigación de un riesgo de tercerizar a una empresa que no tiene presencia ni está sometida a autoridad colombiana a ver yo cómo esta autoridad pudiera cumplir su objeto misional que es mantener seguridad y confianza en el sector asegurador.

Jeimy J. Cano M.
Muy interesante. Víctor. Yo creo que desde un planteamiento de riesgos es como ese acercamiento que decía yo a términos de negocio, ¿no? De cómo se mide el riesgo en cada organización y que los impactos positivos y negativos de confiar o no confiar en la tecnología se vuelve un tema que podrías llegar desde ese punto de vista a acercarlos al negocio y ser más conscientes de que hay que trabajar en estos temas desde el punto de vista estratégico, teniendo ese entendimiento de cuál es el riesgo que quiere cubrir la organización y sus temas de tolerancia y apetito para efectivamente implementar o no, o trabajar con con esta tendencia, este término de confianza digital.

Andrés Almanza



Sí, el apetito de riesgo, porque a mí me gustó mucho lo que decía Rafael del tema de el solo objeto social ya define el criterio del apetito, que fue como yo lo entendí. Entonces, eso me parece como una vista muy interesante. Porque también llamará a que las empresas entiendan cuál es el objeto social, por lo menos en las privadas, que es donde más está hacia el lado comercial, si ese puede ser el marcador definitivo. Porque al final lo que tú estás planteando es, si la empresa conoce su apetito y lo tiene claro, pues va a modelar mejor ese factor de confianza.

Entonces, si vamos más atrás, entonces es lo que está diciendo Rafael, Para acercarlo con lo de Víctor, acercarlo al nivel ejecutivo, se está conociendo bien la empresa, su objeto social en los términos más profundos posibles para que ese apetito de verdad sea con claridad y no sea simplemente como decir, bueno, sí, yo tengo un alto apetito al riesgo, pero tú no sabes qué es alto apetito al riesgo en el momento de la crisis. El apetito al riesgo yo lo dejo escrito en la ley de papel, pues sigue funcionando. Alto apetito al riesgo, pero eso no lo sabemos sin entender mejor el concepto del objeto social, que creo que es un interesante indicador que podría dar una pauta verdaderamente real de cuánto puede ser ese apetito real.

Jeimy J. Cano M.

Bueno, yo creo que vamos cerrando este ejercicio y quisiera que cada uno hiciera como su mensaje de cierre, de síntesis de lo que hemos conversado y con eso vamos cerrando nuestra sesión. Entonces no sé quién quiere empezar con su reflexión final sobre esto que hemos conversado.

Víctor Vásquez

Sí, creo que definitivamente toca buscar más espacios como estos y subirle el nivel para que generar más conciencia. Realmente se está hablando poco, ¿no? Sabemos que es importante, pero no está teniendo la relevancia que debería tener hablar más y en diferentes niveles de este tema de confianza digital y realmente sí es un tema relevante para los negocios, ¿no? Para generar negocios hay que tener confianza, ¿no? entonces creo

que eso es mi mensaje que hay que trabajar más en espacios como este o espacios académicos para generar más iniciativas y que esto coja mayor tracción y se utilicen los marcos que existen actualmente como el DT de Isaac

Rafael Gamboa

Siempre he pensado, he dicho que la confianza es el activo más importante de cualquier organización. O sea, la gente compra, hace o deja de hacer con determinada organización por la confianza que le tiene. La gran oportunidad de mejora o el gran riesgo que tiene cualquier entidad es la parte tecnológica. A mí el sistema me falla, pierdo confianza, se pierden clientes, entonces ya no se trata solamente con que la tecnología funcione, sino que adicionalmente pueda estar complementada, soportada por el tema de la por el tema jurídico, porque necesariamente vivimos en un entorno jurídico y debo cumplirlo.

Andrés Almanza

Coincido con las dos visiones. Me gusta lo de Víctor de buscar más espacios y yo ampliaría en que los espacios con múltiples actores. Creo que una de las conclusiones fundamentales de hoy es definitivamente ecosistemas, multidisciplinaridad. Son dos aspectos claves de cara a lo que hoy estamos viviendo con la acelerada realidad que vivimos. Gobernar un ambiente denso, con la IA de por medio, con lo que se viene con Quantum. Entonces la confianza definitivamente es algo esencial, no para la seguridad, sino para la sostenibilidad de un negocio. Yo creo que lo de Rafael y sus tres pilares hacen o le ponen un marco importante a la confianza. La confianza hay que hacerla sostenible y legible.

¿La otra cuál es? Sostenible, defendible y legítima. Entonces la confianza hay que hacerla legítima, defendible y sostenible. Y yo creo que ahí hay un interesante framework que tendremos que explorar, desarrollar a través de un marco cualquiera que este sea, hoy por hoy, pues el de Isaaca, que es uno de los que más se conoce, cualquiera que sea, pero tenemos ahí un punto clave que de cara al futuro de los negocios tendrá que ser tenido en cuenta en todos los escenarios. Y como lo dices tú, Jeimy, sobre todo en esos ambientes ejecutivos donde este tema debe ser más una conversación continua. ¿Cómo vamos a hacer sostenible la confianza de nuestros clientes?

Transformación de la postura

Sobre ciberseguridad ejecutiva en las juntas directivas: de "a prueba de fallas" a "resistente ante las fallas"

DOI: 10.29236/sistemas.n179a6

Resumen
Las juntas directivas enfrentan una brecha estructural entre la manera en que conciben el riesgo cibernético y la naturaleza real de los sistemas sociotécnicos modernos. La perspectiva dominante, denominada "*fail-safe*" ("a prueba de fallas"), busca prevenir todo falla mediante la implementación de controles formales, generando una falsa sensación de seguridad que expone a las organizaciones a efectos en cascada cuando se materializa la inevitabilidad de la falla. Este artículo propone una transformación conceptual y práctica hacia una postura "*safe-to-fail*" ("resistente ante la falla"), fundamentada en los principios de la ingeniería del caos en seguridad (*Security Chaos Engineering*, SCE) y la resiliencia de sistemas complejos, que permita a los equipos directivos aprender, desaprender y reaprender de las brechas y los incidentes de seguridad, para innovar y moverse rápidamente a pesar de la presencia de eventos cibernéticos adversos pasados, presentes y futuros.

Palabras clave
Ciberseguridad ejecutiva, junta directiva, resiliencia sistémica, a prueba de fallas, ingeniería del caos en seguridad,

Introducción

La digitalización acelerada de las organizaciones ha transformado el riesgo cibernético en uno de los riesgos empresariales más críticos de la década. Sin embargo, la manera en que las juntas directivas comprenden, gobiernan y responden frente a esta realidad permanece inmersa en paradigmas diseñados para un mundo lineal y predecible, que resulta disonante para la complejidad creciente de los sistemas de información modernos.

Según el Informe Global de Riesgos 2024 del Foro Económico Mundial, los ataques cibernéticos y la inseguridad de la información figuran entre los diez riesgos globales más importantes para los próximos dos años, siendo el segundo riesgo corporativo con mayor consenso entre líderes empresariales a nivel global (World Economic Forum, 2024). IBM reporta que el costo promedio global de una violación de datos en 2023 alcanzó los USD 4.45 millones, el valor más alto registrado en los 18 años de historia del informe, con un incremento del 15% respecto a 2020 (IBM, 2023). En América Latina, la región experimentó más de 137.000 millones de intentos de ciberataques en 2022, según Fortinet (2023), lo que representa un aumento significativo frente a años anteriores.

A pesar de estas cifras, la Encuesta Global de Directores 2023 de PwC reveló que solo el 39% de los miembros de juntas directivas encuestados declararon tener alta confianza en la capacidad de su organización para recuperarse de un ciberataque significativo, y sólo el 23% afirmó recibir información suficiente sobre los riesgos cibernétios para ejercer una supervisión efectiva (PwC, 2023). Gartner proyecta que para 2026, el 70% de las juntas directivas incluirán un miembro con experiencia en ciberseguridad, reconociendo la brecha de competencia actual (Gartner, 2023).

Esta asimetría entre la magnitud del riesgo y la capacidad de gobernanza directiva; es el resultado de un paradigma conceptual basado exclusivamente en la aplicación exclusiva de estándares y buenas prácticas que busca lograr una organización “a prueba de fallas”.

Esta perspectiva “a prueba de fallas”, se centra en prevenir todo incidente mediante controles acumulativos, lo que Shortridge y Rinehart (2023) lleva a un “teatro de la seguridad”: contexto que crea la percepción de protección sin lograr una resiliencia real. Cuando la falla, inevitable en sistemas complejos, finalmente ocurre, las organizaciones presentan limitaciones en su capacidad adaptativa para

responder con agilidad a estos eventos.

La propuesta que se detalla en este artículo no recaba en aspectos técnicos de implementación, sino en una reconfiguración de los marcos mentales con los que los miembros de junta conciben el riesgo, la responsabilidad y el éxito en materia de la gestión del riesgo cibernético.

Adoptando los principios de la ingeniería del caos en seguridad, se argumenta que las juntas directivas deben transitar de la búsqueda de la prevención hacia la construcción de capacidades organizacionales y cibernéticas que permitan absorber, adaptarse y aprender de la inevitabilidad de la falla.

La junta directiva y riesgo cibernético: cuatro tensiones para analizar

Las juntas directivas operan sobre la base de marcos de gobernanza diseñados históricamente para riesgos financieros, operacionales y regulatorios de naturaleza relativamente lineal. El riesgo cibernético, por su naturaleza sistémica, interdependiente y adaptativa, desafía estos marcos en sus fundamentos. Cuatro tensiones estructurales caracterizan esta problemática, que incomodan el saber previo de los miembros de junta sobre la gestión de riesgos empresariales.

Tensión No.1: *Complejidad técnica y la capacidad de comprensión directiva*. La mayoría de los miembros de juntas cuentan con limitada formación técnica en ciberseguridad, lo que los hace dependientes de representaciones simplificadas que a menudo distorsionan la naturaleza real del riesgo. Como señala el *National Association of Corporate Directors* (NACD, 2023), el 65% de los directores encuestados indicaron sentirse poco preparados para supervisar el riesgo cibernético de manera efectiva. Esta brecha conduce a la aceptación ciega de métricas basadas en volúmenes, número de vulnerabilidades detectadas, porcentaje de cobertura de controles; en lugar de métricas de desempeño y capacidad orientadas a la resiliencia efectiva, que muestre qué tanto aprende la empresa de sus eventos adversos y cómo se preparar no sólo para sobrevivir, sino para permanecer.

Tensión No.2: *Demanda de certezas y la naturaleza incierta del riesgo cibernético*. Los directivos, formados en culturas de toma de decisiones que privilegian las certezas y el control, demandan garantías de seguridad que el entorno de sistemas complejos no puede ofrecer. Shortridge y Rinehart (2023) señalan que las metodologías cuantitativas (como FAIR - *Factor Analysis of Information Risk*), ampliamente utilizadas en programas de seguridad de la información, requieren el conocimiento sobre la frecuencia y magnitud de los eventos cibernéticos adversos que son prácticamen-

te imposibles de calcular con precisión en sistemas complejos y con efectos en cascada, generando una ilusión cuantitativa de control.

Tensión No.3: Presión productiva y la inversión en resiliencia. En el ámbito del riesgo cibernético, existe una tensión fundamental: la presión por salir primero en mercados competitivos prioriza la eficiencia inmediata, limitando la inversión en resiliencia y resistencia a los ataques (Rasmussen, 1997). Esta presión lleva al debilitamiento de las defensas cibernéticas, aumentando el apetito de riesgo cibernético empresarial, fuera de la zona definida inicialmente, generando variaciones menores que desencadenan fallos inesperados, erosionando progresivamente la resiliencia incluso cuando los indicadores financieros son positivos.

Tensión No.4: Respuesta reactiva y la capacidad adaptativa proactiva. Durante una crisis, el impulso de actuar para retomar el control suele derivar en decisiones impulsivas que ignoran los costos de oportunidad, sacrificando a menudo evidencia forense crucial o la continuidad operativa. Frente a esta reacción, la proactividad exige una preparación deliberada mediante planes probados y el fomento del pensamiento lógico, permitiendo alternativas estratégicas como la "espera vigilante". Equilibrar esta tensión es vital para evitar que la urgencia emocional comprometa la resiliencia organizacional (Dykstra et al., 2022).

Estas cuatro tensiones generan consecuencias institucionales relevantes. Las juntas directivas tienden a aprobar presupuestos de seguridad estructurados alrededor de inversiones en herramientas preventivas y detectivas, sin asignar recursos equivalentes al desarrollo de capacidades de respuesta, recuperación y aprendizaje.

Los modelos de reporte sobre el riesgo de ciberseguridad al directorio hacen énfasis en métricas de estado que no capturan la dimensión dinámica de la resiliencia. Y los marcos de responsabilidad ejecutiva sancionan la falla mediante investigaciones de "causa raíz" que inevitablemente terminan atribuyendo responsabilidad a individuos, en lugar de identificar los factores sistémicos que Reason (1990) denomina "condiciones latentes", aquellas debilidades ocultas, fallas de diseño o decisiones organizacionales deficientes que permanecen inactivas en un sistema durante mucho tiempo. No causan accidentes inmediatos, pero crean brechas en las defensas, facilitando que errores humanos (fallas activas) desencadenen eventos desafortunados.

En síntesis, se presenta en la tabla 1 el resumen de las tensiones entre la junta directiva y el riesgo cibernético.

Tabla 1. *Síntesis de la problemática: Junta directiva y riesgo cibernético*

	Manifestación en la junta directiva	Consecuencia organizacional
Competencia técnica	Dependencia de representaciones simplificadas del riesgo.	Métricas de volumen en lugar de métricas de desempeño.
Marco conceptual	Paradigma dominante: "a prueba de fallas"	Falsa sensación de seguridad
Estructura de reporte	Indicadores de estado, no de capacidad adaptativa.	Bajo niveles de inversión en resiliencia.
Cultura de responsabilidad	Búsqueda de causa raíz individual.	Limitación del aprendizaje organizacional.
Presiones productivas	Eficiencia sobre resiliencia.	Erosión progresiva de las defensas cibernéticas
Respuesta al incidente	Sesgo hacia la acción inmediata.	Decisiones emocionales y aumento de costos.

Nota: Elaboración propia.

El paradigma "resistente ante fallas" y la ingeniería del caos de la seguridad

Ahern (2011) introduce el concepto "*safe-to-fail*" ("resistente ante fallas") en el contexto del diseño urbano resiliente, argumentando que las estrategias efectivas de gestión del riesgo no buscan eliminar la falla, sino diseñar sistemas donde ésta sea controlable, observable y recuperable. Kim et al. (2017) extienden este marco al dominio de sistemas complejos bajo incertidumbre, identificando que la transición de "*fail-safe"* ("a prueba de fallas") a "safe-to-fail" requiere un cambio en el foco: de prevenir la falla de componentes individuales a mantener las funciones críticas del sistema bajo condiciones adversas.

El paradigma "*safe-to-fail*" representa una transformación conceptual y filosófica en la gestión del riesgo cibernético, desplazando el enfoque de la prevención hacia la resiliencia operativa. A diferencia del modelo tradicional "*fail-safe*", que intenta cerrar toda vulnerabilidad bajo una falsa sensación de seguridad, este enfoque acepta la inevitabilidad de la falla en sistemas sociotécnicos complejos. En lugar de considerar las sorpresas como inaceptables, el diseño "*safe-to-fail*" prioriza la continuidad de las funciones críticas y la reducción de las consecuencias del impacto por encima de la reducción de la probabilidad de daño (Shortridge & Rinehart, 2023).

Este paradigma fomenta una cultura de aprendizaje y flexibilidad mediante la experimentación continua y el uso de la ingeniería del caos, habilitando que el sistema se adapte y recupere su estado operativo con agilidad. Esto es, reemplazar el control administrativo tradicional, propio de la vista exclusiva de las listas de chequeo y verifica-

ción centralizadas, por agentes autónomos y descentralizados, donde las decisiones se toman a nivel local por quienes realizan el trabajo, permitiendo al sistema que se adapte rápidamente a contextos específicos sin esperar una aprobación centralizada. De esta forma, las organizaciones fortalecen su resistencia frente a los ataques, asegurando que el sistema evolucione de forma rápida ante la adversidad (Shortridge & Rinehart, 2023).

La ingeniería de caos en seguridad (SCE – *Security Chaos Engineering*) es una disciplina sociotécnica diseñada para fortalecer la resiliencia mediante la experimentación continua y empírica. Este paradigma aplica el método científico introduciendo proactivamente fallos controlados y condiciones adversas, como errores de configuración o escenarios de ataque, para observar cómo el sistema responde y se adapta en la realidad. Utilizando el enfoque de Evaluación y Experimentación (E&E), la SCE permite descubrir debilidades sistémicas y refinar modelos mentales antes de que ocurran incidentes reales. Su uso sistemático optimiza la capacidad adaptativa, asegurando la continuidad de las funciones críticas frente a entornos digitales complejos (Shortridge & Rinehart, 2023).

En resumen, mientras que el diseño "*safe-to-fail*" busca expandir los umbrales de operación de la seguridad para otorgar un margen de maniobra al sistema frente a eventos inciertos, la SCE utiliza el enfoque de Evaluación y Experimentación (E&E) para demarcar o mapear esos límites y asegurar que las interacciones complejas a través del espacio-tiempo no resulten en fallos en cascada incontrolables. En conjunto, estos dos conceptos permiten que la seguridad/ciberseguridad sea algo que el sistema "hace" activamente (y no sólo "tiene") (Shortridge & Rinehart, 2023), permitiendo que la organización prospere incluso bajo escenarios de ataques continuos.

Transformación de la perspectiva de la junta directiva: de "a prueba de fallas" a "resistente ante las fallas"

La transformación de la perspectiva directiva en ciberseguridad requiere intervenir tres dimensiones simultáneas: conceptual (cómo los directivos entienden el riesgo cibernético), estructural (cómo se organiza la gobernanza y el reporte), y conductual (cómo actúan los directivos ante un incidente o una decisión de inversión).

El principio rector de esta transformación es que "la resiliencia es algo que un sistema hace, no algo que un sistema tiene" (Shortridge & Rinehart, 2023, p. 20). Para las juntas directivas, esto significa transitar de la pregunta "¿Estamos seguros?" hacia la pregunta "¿Qué tan capaces somos de absorber, adap-

tarnos y aprender cuando algo falla?".

En este contexto, se detallan a continuación cinco (5) estrategias básicas que habiliten la transformación de la mentalidad o perspectiva de la junta directiva de "a prueba de fallas" a "resistente ante las fallas" frente la gestión y gobernanza del riesgo cibernético, con algunas recomendaciones para la aplicación de las mismas.

Estrategia No.1 Reencuadre del lenguaje y los modelos mentales. El primer paso de la transformación es modificar el vocabulario directivo sobre ciberseguridad. Los términos "prevención" y "protección" evocan una lógica de fortaleza estática que es conceptualmente incompatible con la naturaleza sistémica de los ecosistemas digitales actuales. Deben introducirse progresivamente términos como "resiliencia", "capacidad adaptativa", "tiempo de recuperación" y "aprendizaje de las fallas", "pedagogía del error".

Esta estrategia debe implementarse mediante sesiones educativas periódicas, en las que el CISO (*Chief Information Security Officer*) ilustre las diferencias entre robustez y resiliencia con ejemplos concretos en el contexto de su sector de negocio. Esto es, entender que la robustez constituye la capacidad estática de resistir perturbaciones para retornar al estado original, mientras la resiliencia representa una capacidad adaptativa dinámica para anticipar, responder y aprender de fallos inevitables.

Recomendación: El reencuadre conceptual puede generar resistencia si los directivos perciben que se está minimizando la gravedad del riesgo. La comunicación debe ser gradual y consistente, enfatizando que el paradigma "*safe-to-fail*" no reduce la urgencia de la inversión, sino que la orienta de manera más efectiva.

Estrategia No.2 Rediseño del marco de reporte directivo. El marco de reporte del CISO a la Junta Directiva debe ser rediseñado para capturar capacidades de resiliencia en lugar de estados de control. Forsgren et al. (2018) ofrece las métricas DORA[1] (*Digital Operational Resilience Act*) como punto de partida: frecuencia de despliegue (velocidad de respuesta adaptativa), tiempo de entrega de cambios (agilidad organizacional), tasa de fallas luego de cambios (calidad del proceso), y tiempo de restauración del servicio (capacidad de recuperación).

1 Digital Operational Resilience Act (DORA) - La Ley de Resiliencia Operativa Digital (DORA) es una normativa introducida por la Unión Europea para reforzar la resiliencia digital de las entidades financieras. Entró en vigor el 17 de enero de 2025 y asegura que los bancos, las compañías de seguros, las empresas de inversión y otras entidades financieras puedan resistir, responder y recuperarse de las interrupciones de las TIC (tecnologías de la información y la comunicación), como los ciberataques o los fallos del sistema. Fuente: https://www.eiopa.europa.eu/digital-operational-resilience-act-dora_en

Complementariamente, el marco de reporte debe incluir evidencia experimental proveniente de ejercicios de ingeniería de caos en seguridad. Shortridge y Rinehart (2023) proponen el concepto de "confianza basada en evidencia": cada experimento genera datos concretos sobre cómo el sistema responde a condiciones adversas.

Recomendación: La transición de métricas de volumen a desempeño puede generar incomodidad inicial. Es necesario preparar al directorio para interpretar el incremento de la visibilidad de las fallas o brechas de seguridad como un indicador de madurez, no de deterioro de la postura de ciberseguridad.

Estrategia No.3 Incorporación de la tolerancia a la falla en el apetito de riesgo. Las declaraciones de apetito de riesgo deben incorporar explícitamente dimensiones de tolerancia a la falla: ¿Qué magnitud de interrupción es aceptable? ¿Durante cuánto tiempo? ¿Afectando qué funciones críticas? Shortridge y Rinehart (2023) proponen enmarcar la resiliencia en términos de funcionalidad crítica, obligando a los directivos a especificar qué puede ser sacrificado temporalmente para proteger la promesa de valor de la empresa.

Redefinir el apetito, no como la ausencia de incidentes, sino como la capacidad adaptativa para absorber y recuperarse de eventos adversos preservando las funciones críticas, permite construir "capacidad de amortiguación" y emplear la ingeniería de caos para validar los umbrales de operación empresarial, transformando el riesgo en una competencia operativa verificable.

Recomendación: La definición de lo "aceptable" requiere alineación entre junta, alta gerencia y equipos operativos alrededor de la promesa de valor empresarial. Las declaraciones desconectadas de la realidad operativa generan expectativas directivas que erosionan la confianza del cliente y de los accionistas.

Estrategia No.4 Institucionalizar el aprendizaje post-incidente. La junta directiva debe institucionalizar revisiones post-incidente "sin buscar culpables" como práctica estándar de gobernanza. Shortridge y Rinehart (2023) sugieren que las revisiones deben privilegiar los factores sistémicos de los eventos analizados como pueden ser presiones productivas, propiedades del diseño, brechas de observabilidad, en lugar de errores individuales.

La implementación de una cultura que "no busca culpables" en la gestión del riesgo cibernético exige una transformación estructural que priorice la seguridad psicológica sobre el castigo individual, permitiendo que el personal reporte fallas sin temor a represalias. Este paradigma desestima el "error humano"

como causa raíz, tratándolo en cambio como un síntoma de deficiencias sistémicas y de diseño en los sistemas.

Recomendación: La cultura que "no busca culpables" puede ser percibida como ausencia de responsabilidad por las acciones ejecutadas. Es fundamental distinguir entre responsabilidad sistémica, la organización aprende y mejora, y la culpa individual. La primera fortalece la resiliencia; la segunda la deteriora (Shortridge & Rinehart, 2023).

Estrategia No.5 Supervisión directiva de la capacidad experimental. La junta directiva debe incluir en su agenda la revisión de resultados de los ejercicios de la ingeniería de caos en seguridad. Esta supervisión no implica que los directivos diseñen o ejecuten experimentos; implica que demanden evidencia experimental y práctica de los resultados de las pruebas realizadas como estándar de gobernanza y forma de retar la postura de ciberseguridad actual. La pregunta directiva fundamental debe ser:

Tabla 2. *Síntesis de estrategias de transformación directiva*

Estrategia	Objetivo	Ventajas	Retos	Métricas de seguimiento
1. Reencuadre de modelos mentales	Modificar el vocabulario directivo sobre riesgo cibernético	Mayor comprensión sistémica; decisiones más contextualizadas	Resistencia al cambio conceptual; percepción de minimización del riesgo	Calidad de las preguntas directivas en sesiones de reporte
2. Rediseño del marco de reporte	Capturar capacidades de resiliencia, no estados de control	Decisiones de inversión más efectivas; visibilidad real del riesgo	Incomodidad con métricas que revelan más vulnerabilidades	Adopción de métricas DORA; cobertura de experimentos
3. Apetito de riesgo con tolerancia a la falla	Especificar funcionalidades críticas y niveles aceptables de interrupción	Priorización estratégica de la inversión	Dificultad para acordar qué es "aceptable"	Tiempo de recuperación de funciones críticas en ejercicios
4. Institucionalizar el aprendizaje post-incidente	Extraer conocimiento sistémico de cada incidente	Cultura de mejora continua; supresión reducida de información	Confusión entre cultura "sin buscar culpables" y ausencia de responsabilidad	Número de mejoras sistémicas implementadas post-incidente
5. Supervisión de capacidad experimental	Incluir evidencia de ingeniería del caos en seguridad en la agenda directiva	Decisiones de inversión basadas en comportamiento observado	Interpretación negativa de los hallazgos del SCE.	Frecuencia y cobertura de ejercicios; acciones tomadas en los sistemas

Nota: Elaboración propia.

"¿Qué hemos aprendido sobre nuestros sistemas este trimestre que antes no sabíamos?"

Recomendación: Los primeros ejercicios frecuentemente revelan comportamientos adversos con efectos sistémicos (en cascada) inesperados. Los directivos deben estar preparados para recibirlos sin interpretarlos como un fallo del equipo de seguridad, sino como el propósito mismo del ejercicio y como un "foro de la verdad" sobre la realidad de la ciberseguridad de la empresa.

En la tabla 2, se presenta una síntesis de las cinco estrategias planteadas.

Apropiación del paradigma "resistente ante las fallas": recomendaciones para los miembros de la junta directiva

Las siguientes recomendaciones, detalladas en la tabla 3, están orientadas para que los miembros individuales de la junta directiva adopten conductas concretas que permitan una mejor apropiación de la perspectiva "*safe-to-fail*" en su práctica cotidiana de gobernanza.

Conclusiones

La gestión moderna del riesgo cibernético ha llegado a un punto de inflexión donde las estrategias tradicionales de defensa conocidas resultan insuficientes ante la naturaleza dinámica de las amenazas digitales y el avance de la inteligencia artificial como herramienta base de los atacantes. La transformación necesaria para las organizaciones no es sólo tecnológica, sino que exige un cambio de paradigma en la toma de decisiones estratégicas que debe ser liderado desde la junta directiva. Esta transformación implica transitar de una mentalidad de prevención de riesgos conocidos hacia una de resiliencia operativa y aprendizaje continuo en un entorno de amenazas incierto y asimétrico (Smeets, 2022).

El paradigma "*fail-safe*" ("a prueba de fallas") ha sido el pilar de la ciberseguridad tradicional, basándose en la premisa de que los riesgos pueden predecirse y controlarse con precisión, y bloquearse mediante las herramientas tecnológicas disponibles. Este enfoque, enraizado en una visión determinista de la ciencia del siglo XX, busca eliminar vulnerabilidades y amenazas antes de que ocurran, operando bajo un factor de seguridad formalmente diseñada (Shortridge & Rinehart, 2023).

En contraste, el paradigma "*safe-to-fail*" ("resistente a las fallas") representa una evolución necesaria hacia la resiliencia operativa, aceptando la inevitabilidad de la falla en entornos sociotécnicos complejos y evolución permanente. Este modelo de diseño estratégico se enfoca en permitir que la infraestructura falle de manera contenida, priorizando la minimización de las consecuencias del impacto por encima de la simple reducción de la

Tabla 3. *Recomendaciones para apropiar el paradigma "resistente ante las fallas" en las juntas directivas*

Estrategia	Objetivo principal	Preguntas clave para la Junta
Reformulación del Cuestionamiento	Pasar de un enfoque de protección pasiva a uno de aprendizaje y respuesta.	• ¿Qué aprendimos sobre nuestras capacidades de respuesta este trimestre? • ¿Cuánto tiempo tardamos en recuperar nuestras funciones críticas? • ¿Qué hipótesis sobre nuestros sistemas resultaron incorrectas?
Reportes basados en evidencia	Incentivar la inversión en capacidades de aprendizaje organizacional.	• ¿Qué hipótesis de la aplicación de la ingeniería del caos en seguridad se probaron recientemente? • ¿Qué descubrimientos se hicieron y qué mejoras concretas se implementaron luego de las pruebas?
Definición de prioridades críticas	Establecer una estrategia "*safe-to-fail*" ("resistente ante fallas").	• ¿Qué funciones del negocio son verdaderamente vitales? • ¿Qué procesos pueden suspenderse para proteger la promesa de valor de la empresa?
Cultura de revisión "sin buscar culpables"	Fomentar el aprendizaje sistémico tras un incidente.	• ¿Nuestro marco actual de responsabilidad incentiva la transparencia o el ocultamiento de errores? • ¿Cómo estamos transformando las fallas identificadas en mejoras del sistema?
Calibración del apetito de riesgo	Alinear la estrategia con las capacidades reales de la empresa.	• Si nuestra meta de recuperación es de 24 horas pero los ejercicios muestran 72, ¿debemos ajustar la declaración o aumentar la inversión?
Diversidad de competencias	Fortalecer la supervisión experta del riesgo cibernético.	• ¿Contamos con directores con experiencia técnica suficiente para cuestionar la estrategia de ciberseguridad? • ¿Qué mecanismos de asesoría experta tenemos hoy?
Práctica de "espera vigilante"	Evitar decisiones emocionales por sesgo de acción durante crisis.	• ¿Cuál es el costo real de esperar 30 minutos más para recopilar información antes de aprobar una respuesta drástica?

Nota: Elaboración propia.

probabilidad de daño. Un diseño "*safe-to-fail*" privilegia la modularidad, la diversidad funcional y el empoderamiento de agentes autónomos descentralizados que pueden responder ágilmente a contextos locales de crisis (Shortridge & Rinehart, 2023).

De otra parte, la Ingeniería del caos en seguridad constituye la disciplina práctica para operacionalizar

esta filosofía "*safe-to-fail*" mediante el uso sistemático del método científico. Su objetivo primordial es generar "memoria muscular" tanto en los sistemas técnicos como en los equipos humanos, dosificando proactivamente fallos controlados y escenarios de ataque para observar la respuesta real del sistema. A diferencia de las pruebas de penetración tradicionales, que validan resultados conocidos, la SCE busca descubrir los "desconocidos desconocidos" y debilidades sistémicas ocultas antes de que se conviertan en incidentes que afecten al cliente, a la reputación corporativa y la promesa de valor de la empresa (Shortridge & Rinehart, 2023).

Finalmente, esta transformación impone retos críticos para los miembros de la junta directiva, quienes deben liderar el cambio cultural hacia una organización resiliente y luego antifrágil. Por tanto, el desafío es mitigar el "sesgo de acción" reactivo tras un incidente, evitando la implementación de sanciones o medidas tecnológicas costosas que privilegian la falsa sensación de seguridad (Dykstra et al., 2022).

Solo mediante la aceptación de la falla como una condición normal de los sistemas complejos, y un enfoque en basado en la flexibilidad de la respuesta y la capacidad adaptativa, podrá la junta directiva asegurar la viabilidad y permanencia de la organización en un entorno digital agresivo y competitivo.

La ciberseguridad del futuro no será construida sobre la ilusión del control, sino sobre la sabiduría de los sistemas que saben cómo fallar bien, esto es, *aprender* de lo que sabe y conoce en la actualidad sobre las vulnerabilidades para asegurar el resultado esperado, *desaprender* aquello que no suma o aporta, o ha quedado obsoleto de la práctica de seguridad y control vigente, para incorporar los nuevos retos, contextos y escenarios no lineales, acelerados, volátiles e interconectados donde se pueden perseguir objetivos valiosos para la empresas, y así *reaprender*, esto es, conectar los puntos inconexos y puntos ciegos identificados hasta ahora para avanzar en las asimetrías que propone el adversario (Gundu, 2024).

Referencias

Ahern, J. (2011). From fail-safe to safe-to-fail: Sustainability and resilience in the new urban world. *Landscape and Urban Planning*, *100*(4), 341–343. https://doi.org/10.1016/j.landurbplan.2011.02.021

Dykstra, J., Stevens, R., & Olson, L. (2022). Opportunity cost of action bias in cybersecurity incident response. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *66*(1), 1116–1120. https://doi.org/10.1177/1071181322661368

Forsgren, N., Humble, J. & Kim, G. (2018). *Accelerate: The science of lean software and DevOps: Building and scaling high performing technology organizations*. IT Revolution Press.

Fortinet. (2023). 2023 global threat landscape report. https://www.fortinet.com/blog/threat-research/2023-global-threat-landscape-report

Gartner. (2023). Gartner predicts 70% of boards will have a dedicated cybersecurity committee by 2026. *Gartner Research.* https://www.gartner.com/en/newsroom/press-releases/2023-03-28-gartner-predicts-70-percent-of-boards-will-have-a-dedicated-cybersecurity-committee-by-2026

Gundu, T. (2024). Learn, Unlearn and Relearn: Adaptive Cybersecurity Culture Model. *Proceedings of The 19th International Conference on Cyber Warfare and Security*, *19*(1), https://doi.org/10.34190/iccws.19.1.2177

IBM. (2023). Cost of a data breach report 2023. https://www.ibm.com/reports/data-breach

Kim, Y., Newman, G. & Güneralp, B. (2017). Fail-safe and safe-to-fail adaptation: Decision-making for urban flooding under climate change. *Climatic Change, 145*(3), 397–412. https://doi.org/10.1007/s10584-017-2100-5

National Association of Corporate Directors - NACD. (2023). *2023* NACD director survey: Cybersecurity oversight. *NACD*. https://www.nacdonline.org/globalassets/public-pdfs/nacd_cyber-risk-oversight-handbook_pages_web-compressed.pdf

PwC. (2023). 2023 global digital trust insights survey. *PricewaterhouseCoopers*. https://www.pwc.com/gx/en/issues/cybersecurity/global-digital-trust-insights.html

Rasmussen, J. (1997). Risk management in a dynamic society: A modelling problem. *Safety Science,* 27(2–3), 183–213. https://doi.org/10.1016/S0925-7535(97)00052-0

Reason, J. (1990). *Human error.* Cambridge University Press. https://doi.org/10.1017/CBO9781139062367

Shortridge, K. & Rinehart, A. (2023). *Security chaos engineering: Sustaining resilience in software and systems*. O'Reilly Media.

Smeets, M. (2022). *No shortcuts. Why states struggle to develop a military cyber-force*. New York, NY. USA: Oxford University Press.

World Economic Forum. (2024). Global risks report 2024. *WEF*. https://www.weforum.org/reports/global-risks-report-2024/

Jeimy J. Cano M., Ph.D, CFE, CICA. *Ingeniero y Magíster en Ingeniería de Sistemas y Computación, Universidad de los Andes. Especialista en Derecho Disciplinario, Universidad Externado de Colombia; Ph.D en Business Administration, Newport University, CA. USA. y Ph.D en Educación, Universidad Santo Tomás. Profesional certificado como Certified Fraud Examiner (CFE), por la Association of Certified Fraud Examiners y Certified Internal Control Auditor (CICA) por The Institute of Internal Controls. Profesor Distinguido de la Facultad de Derecho, Universidad de los Andes. Es director de la Revista SISTEMAS de la Asociación Colombiana de Ingenieros de Sistemas –ACIS–.*

Gobernanza de la inteligencia artificial generativa

Una introducción a la confianza por diseño.

DOI: 10.29236/sistemas.n179a7

Resumen

Este artículo propone el marco de "Confianza por Diseño" (CpD) como estrategia de gobernanza para la IA Generativa (IA Gen). Ante los vacíos de los enfoques tradicionales, la CpD integra la seguridad, la privacidad y la ética. El modelo busca armonizar los intereses del Estado (garante), las empresas (innovación) y los individuos (autonomía cognitiva). Se enfatiza que la sostenibilidad tecnológica depende de salvaguardas proactivas que aseguren la integridad algorítmica y mitiguen riesgos como la "*hipersuación*"[1] y la "infodemia". La reflexión concluye que la CpD permite una innovación responsable, transformando la IA en un habilitador de bienestar social y rentabilidad económica mediante una colaboración pluralista entre los diferentes actores del ecosistema digital soporte de la IA.

Palabras clave

Confianza por diseño, seguridad por diseño, privacidad por diseño, ética, IA Generativa

1 La "hipersuasión" (en inglés: *hypersuasion*) es un concepto acuñado por Luciano Floridi para describir formas de persuasión altamente sofisticadas y personalizadas habilitadas por sistemas de IA. Fuente: Floridi, L. (2024). Hypersuasion – On AI's persuasive power and how to deal with it. *Philosophy & Technology*, *37*(64). https://doi.org/10.1007/s13347-024-00756-6

Introducción

La inteligencia artificial (IA) y el aprendizaje automático han transformado radicalmente el panorama de los negocios y la dinámica del mundo actual. La tasa de adopción y penetración ha venido en un crecimiento sostenido desde el lanzamiento de ChatGPT en 2022, cuando alcanzó los 100 millones de usuarios activos mensuales en solo dos meses, convirtiéndose en la aplicación de consumo de más rápido crecimiento en la historia hasta ese momento (Floridi, 2023).

De otra parte, se registró un aumento masivo del 425% en las inversiones de capital de riesgo destinadas a la IA Gen desde el año 2020 hasta mediados de 2023. De igual forma, se estima que el mercado de la IA Gen se convertirá en una industria de 1,3 billones de dólares para el año 2032, lo que sugiere que las organizaciones continuarán desarrollando iniciativas para incorporar esta tecnología procurando mayor automatización y transformación de los diferentes sectores productivos a nivel internacional (Habuka & Socol de la Osa, 2025).

Esta carrera tecnológica necesariamente genera tensiones entre diferentes intereses en la dinámica de la sociedad actual mediada por tecnologías disruptivas y emergentes, creando vacíos de responsabilidad y riesgos emergentes, que no se pueden explicar o atender exclusivamente desde las perspectivas tradicionales como la seguridad y la privacidad, las cuales están diseñadas para contextos específicos de los retos y exigencias, tanto de los reguladores como de los individuos, asociados con estándares y buenas prácticas reconocidas.

Por tanto, este documento propone un modelo de “confianza por diseño” como una estrategia para la gobernanza corporativa de la IA Generativa (IA Gen). A través de un análisis de los intereses divergentes y convergentes de las empresas, los reguladores y los ciudadanos, se argumenta que la sostenibilidad a largo plazo de la IA depende de la integración proactiva de controles éticos y técnicos desde la fase de diseño, con el fin de limitar los daños por fallas propias del funcionamiento del ecosistema digital donde opera y así, asegurar su aceptación social.

Para ello, este trabajo inicia con una breve revisión de literatura de los conceptos de privacidad por diseño y seguridad por diseño como referentes básicos disponibles en la actualidad. Luego se plan-

tea la paradoja de la IA Gen en el contexto actual, a reglón seguido se detallan las tensiones entre los tres participantes claves de la dinámica social alrededor de la IA Gen como son el mercado, el estado y los individuos para establecer las bases de la construcción del modelo de confianza por diseño. Seguidamente se detalla la propuesta de la confianza por diseño y se finaliza con algunas conclusiones y retos para este modelo ahora y en el futuro.

La paradoja de la IA Gen: más humanidad que tecnología

La paradoja de la IA Gen establece que, a medida que la tecnología de procesamiento de lenguaje se vuelve más sofisticada y capaz de operar a escala, el éxito de su implementación depende más críticamente de factores humanos que la máquina no puede resolver: la capacidad de comprender, cuestionar e integrar sus resultados (Sieber, 2026).

Aunque la IA Gen se clasifica como una IA estrecha, diseñada para tareas específicas mediante algoritmos estadísticos (Abaimov & Martellini, 2022), su adopción masiva ha generado una zona de la desilusión productiva, donde las empresas invierten grandes sumas sin transformar realmente sus modelos de negocio.

Desde la perspectiva del riesgo cibernético, la paradoja se agrava por el fenómeno de la "adopción en la sombra" (*Shadow AI*), donde los empleados utilizan herramientas personales, exponiendo datos corporativos y comprometiendo el cumplimiento regulatorio (Sieber, 2026). Esto es particularmente peligroso debido a la naturaleza de "caja negra" de los modelos de aprendizaje profundo (*Deep Learning*), cuya opacidad dificulta anticipar decisiones o detectar manipulaciones externas como el envenenamiento de datos (*data poisoning*) (Abaimov & Martellini, 2022).

Por tanto, la transición de "operadores" a "orquestadores" requiere una humildad tecnológica que reconozca que la verificación humana debe ser una parte estructural del flujo de trabajo, no un control final. Mientras el operador tradicional gestiona herramientas de forma rutinaria, el orquestador diseña la interacción estratégica con sistemas autónomos. Esto es, se delega en la máquina la capacidad analítica y ejecución a escala, reservando para el juicio humano la ética, la empatía y la gestión de la ambigüedad. Así, el valor real surge de capturar el "dividendo de la aumentación" (valor creado de maneras que ni el ser humano ni las máquinas podrían lograr en solitario) mediante una supervisión humana estructural (Sieber, 2026).

En la Tabla 1 se presenta un cuadro resumen de los retos de la paradoja de la IA Gen.

Tabla 1. *Retos de la paradoja de la IA Gen*

Fiabilidad Técnica	Las "alucinaciones" son fallos estructurales; la IA Gen genera información plausible pero falsa, requiriendo validación constante (Sieber, 2026).
Ciberseguridad	La IA es de doble uso: las mismas herramientas que defienden pueden usarse para ataques adversarios y generación de malware elusivo (Abaimov & Martellini, 2022).
Psicológico	Resistencia al cambio ante el miedo a la obsolescencia, la pérdida de control y el fracaso público (Sieber, 2026).
Gobernanza	Necesidad de pasar de un enfoque basado solo en eficiencia a uno de "dividendo de aumentación", donde la IA amplifica el juicio ético humano.
Explicabilidad	Superar la falta de transparencia de los modelos no lineales para construir sistemas confiables y auditables (Abaimov & Martellini, 2022)

Nota: Elaboración propia basado en Sieber, 2026 y Abaimov & Martellini, 2022.

Seguridad y privacidad por diseño. Un resumen de las posturas tradicionales

Los pilares teóricos de la Seguridad por Diseño (SpD) se remontan a 1975, cuando Jerome Saltzer y Michael Schroeder (1975) publicaron los principios de diseño que aún hoy constituyen el fundamento de los sistemas confiables. Su visión subrayaba que la robustez de un sistema no debe depender del secreto de su arquitectura (seguridad por oscuridad, no es seguridad), sino de la solidez de sus mecanismos intrínsecos de protección. A continuación se detalla en la tabla No.2 el resumen de las consideraciones de Saltzer y Schroeder (1975).

La implementación efectiva de la SpD en entornos modernos implica desplazar las actividades de seguridad hacia las etapas más tempranas de planificación y diseño de las soluciones informáticas. Esto incluye el modelado de amenazas proactivo, el análisis estático y dinámico de código, y la adopción de metodologías DevSecOps para automatizar la verificación de seguridad en flujos de trabajo ágiles. Desde una perspectiva económica, la literatura indica que resolver fallas de seguridad en la etapa de diseño puede reducir los costos de remediación post-producción, al tiempo que fortalece la confianza de los interesados y la resiliencia operativa frente a amenazas persistentes (NIST, 2022; Valdés-Rodríguez et al., 2023).

De otra parte, la Privacidad por Diseño (PpD) es un marco concep-

Tabla 2. *Principios de diseño seguro*

Principio	Definición	Impacto
Mínimo privilegio	Restricción de acceso a los permisos básicos necesarios para una función.	Minimiza el radio de impacto ante un posible compromiso de cuenta.
Mediación completa	Verificación sistemática de cada intento de acceso a cada objeto.	Previene la existencia de rutas de acceso no supervisadas.
Diseño abierto	La seguridad no debe depender del desconocimiento del atacante sobre el diseño.	Facilita la auditoría externa y la identificación colectiva de fallos.
Economía del mecanismo	Mantener el diseño del sistema de protección lo más simple y pequeño posible.	Reduce la probabilidad de errores de implementación y facilita la verificación.
Valores predeterminados seguros	Denegación de acceso por defecto; el permiso debe ser una excepción explícita.	Asegura que un error de configuración no resulte en una apertura no deseada.
Separación de privilegios	Requiere más de una condición o llave para acceder a recursos críticos.	Evita que la vulnerabilidad de un solo control comprometa todo el sistema.
Aceptabilidad psicológica	El diseño debe ser intuitivo para que el usuario no eluda los controles.	Fomenta el cumplimiento natural de las normas de seguridad.

Nota: Elaboración propia basada en Saltzer & Schroeder, 1975

tual y metodológico que propone la integración de la protección de la privacidad como un requisito funcional y estructural desde la fase de concepción de cualquier sistema, tecnología o práctica organizacional. Desarrollado originalmente por la Dra. Ann Cavoukian en la década de 1990, el PpD traslada la responsabilidad de la privacidad de un modelo basado exclusivamente en el cumplimiento legal reactivo a un paradigma de ingeniería proactivo.

Este enfoque se articula a través de siete principios fundamentales que rigen el ciclo de vida del desarrollo: (Cavoukian, 2009)

- *Proactivo, no reactivo*: Anticipa riesgos antes de que ocurran fallos de privacidad.
- *Privacidad como configuración predeterminada*: Los sistemas deben proteger los datos automáticamente, sin requerir acción del usuario (Privacy by Default).
- *Privacidad integrada en el diseño*: La protección no es un "añadido", sino un componente esencial de la arquitectura.
- *Funcionalidad total* (Suma positiva): Rechaza falsas dicotomías

entre seguridad y funcionalidad, buscando soluciones donde ambos objetivos coexistan.
- *Seguridad de extremo a extremo*: Asegura la protección de los datos desde su recolección hasta su eliminación segura.
- *Visibilidad y transparencia*: Las operaciones deben ser verificables y abiertas al escrutinio independiente.
- *Respeto por la privacidad del usuario*: Centra el diseño en las necesidades y el empoderamiento del individuo.

En el contexto contemporáneo, la PpD ha trascendido el ámbito teórico para convertirse en un estándar legal global, siendo el pilar central del Artículo 25 del Reglamento General de Protección de Datos (GDPR) de la Unión Europea bajo el término "Protección de datos desde el diseño y por defecto". La literatura reciente destaca que el éxito del PpD en entornos de inteligencia artificial y desarrollo ágil depende de la identificación de vulnerabilidades de privacidad en las etapas más tempranas de los requisitos de software, reduciendo costos de reparación post-producción hasta en un 90% (Del-Real et al., 2025).

La investigación técnica y jurídica demuestra de forma particular que la seguridad y la privacidad por diseño no son obstáculos a la innovación, sino cimientos fundamentales para las tecnologías inteligentes. En un mundo donde la IA Gen media en casi toda interacción informativa, la responsabilidad de los arquitectos de sistemas es equiparable a la de los legisladores. Por tanto, la defensa del futuro digital descansa sobre tres pilares: (Paseri & Durante, 2025)
- Preservación del capital semántico: La defensa de la capacidad humana para generar significado frente a la imitación sintáctica.
- Salvaguarda de la autonomía humana: asegurar la soberanía cognitiva frente a la manipulación computacional que perfila y persuade para cambiar comportamientos.
- Equidad sistémica: Asegurar que los beneficios de la IA Gen no intensifiquen la asimetría de poder o la exclusión social.

La dependencia tecnológica excesiva, en ausencia de estos tres pilares, conlleva una subordinación a poderes técnicos opacos que pueden desestabilizar la soberanía estatal y la autonomía individual. Como advierte Zou et al. (2025), el diseño es hoy una forma de legislación por código.

Iniciativas de inteligencia artificial: tensiones entre el Estado, el mercado y el individuo

La irrupción de la Inteligencia Artificial Generativa (IA Gen) en el tejido socio-técnico contemporáneo no representa sólo una optimización de la capacidad de cómputo o un refinamiento incremental de los modelos de aprendizaje automáti-

co. Constituye, en esencia, un desafío existencial[2] para los marcos legales y prácticos tradicionales de seguridad y privacidad, forzando una reevaluación de la soberanía cognitiva y la autonomía individual.

En este sentido, los intereses de las empresas por la incorporación de mayor innovación y rentabilidad, establecen nuevas propuestas y retos en el uso de la IA Gen que llevan a las organizaciones a tomar mayores riesgos para los cuales generalmente no se encuentran preparadas, obligando a las áreas jurídicas a buscar estrategias que permitan minimizar los riesgos legales y reputacionales, en una perspectiva de cumplimiento basado en el pasado: luego de que ya pasaron los hechos y buscar asumir la menor responsabilidad posible (Coles-Kemp & Burdon, 2025).

Por otro lado, está el Estado en su función de garante de los derechos individuales, de la seguridad nacional, el orden público y el cumplimiento de la ley. En este contexto, busca establecer marcos de trabajo lo suficientemente amplios y exigentes para limitar riesgos sistémicos que puedan afectar la dinámica de las infraestructuras críticas y la gobernabilidad de un país, así como evitar abusos de poder por parte de las grandes empresas de tecnología, procurando un uso limitado y vigilado de las tecnologías que permita a los individuos sentirse tranquilos por las implementaciones de iniciativas basadas en IA Gen (Coles-Kemp & Burdon, 2025; Taddeo et al., 2019).

Finalmente y no menos importante, los individuos como parte esencial de la sociedad, demandan transparencia sobre cómo se usan sus datos y si están interactuando con una máquina o un humano, con el fin de limitar los posibles efectos de una manipulación dirigida, o el uso no autorizado de sus datos personales para alimentar un modelo de IA Gen que pueda, no sólo comprometer su privacidad o buen nombre, sino crear escenarios de suplantación, o ser utilizados para distorsionar el comportamiento de un grupo de personas con fines contrarios a la Constitución y la ley (Coles-Kemp & Burdon, 2025).

Por tanto, la gobernanza de la IA Gen exige un cambio de paradigma hacia una responsabilidad digital colectiva e híbrida, que distribuya las obligaciones entre los creadores de los modelos y quienes los despliegan en el mercado. Solo a través de esta convergencia se podrá atender la “brecha de responsabilidad” que surge cuando los sistemas autónomos actúan sin una supervisión humana significativa (Coles-Kemp & Burdon, 2025). Al integrar controles éticos y técnicos desde el inicio, las organizaciones no solo aseguran el cumpli-

2 En el ámbito de la IA, esto implica la creación de una superinteligencia que supere el control humano o el despliegue de armas autónomas letales amenazando la existencia de la humanidad (Zou et al., 2025)

miento normativo, sino que construyen el capital semántico de confianza (intervención humana), que da sentido y contexto, necesario para que la IA Gen sea aceptada como una herramienta de bienestar común y no como un riesgo existencial para la humanidad.

Confianza por diseño: hacia un marco de trabajo de intereses convergentes.

La intersección donde convergen los tres intereses previamente comentados es la Confianza por Diseño (CpD). En este punto, la tecnología deja de ser un factor de incertidumbre para convertirse en un habilitador de ventajas competitivas: es lícita al cumplir con la regulación estatal, ética al respetar la autonomía del individuo y estratégica al generar valor sostenible para la empresa. Por tanto, la CpD se configura como una práctica para integrar gobernanza, ética y rendición de cuentas en la base misma de las tecnologías, asegurando que el comportamiento confiable sea el valor predeterminado y no una excepción.

Las empresas deben transitar de una innovación "sin límites" a una "responsable". La gobernanza de la IA ahora requiere directivos competentes en tecnología que puedan supervisar la "integridad de la IA" como parte de sus deberes fiduciarios (Coles-Kemp & Burdon, 2025). Cuerpos de gobierno que reconozcan la inevitabilidad de la falla de las iniciativas basadas en inteligencia artificial, la opacidad algorítmica, y por lo tanto, exijan la mediación humana en las respuestas que puedan afectar negativamente a los diferentes grupos de interés, así como protocolos de actuación definidos y practicados, cuando la IA Gen genere acciones adversas por fallas estructurales que perjudiquen directamente a sus clientes (Abaimov & Martellini, 2022).

De otro lado, las regulaciones como el Reglamento de IA de la Unión Europea clasifican los sistemas según su impacto en los derechos fundamentales. Clasifica los sistemas en cuatro categorías: inaceptable, alto, limitado y mínimo. Los de riesgo *inaceptable* están prohibidos, abarcando la afectación social y la manipulación conductual. Los sistemas de *alto* riesgo, aplicados en infraestructuras críticas, salud, educación y justicia, requieren una evaluación rigurosa, transparencia y supervisión humana. El riesgo *limitado* exige transparencia (como en chatbots), mientras que el riesgo *mínimo* (filtros de spam) no se regula. El Estado actúa como garante para evitar riesgos sistémicos y la manipulación de la opinión pública, que limite la erosión de la autonomía individual y el Estado de derecho (EUPC, 2024).

Y finalmente el individuo, como "consumidor algorítmico", puede estar perdiendo el control por la emergencia del fenómeno de la "hipersuasión", la capacidad de la

IA para manipular el comportamiento humano de forma sutil y personalizada, que amenaza la autodeterminación mental de los ciudadanos, lo que exige un diseño técnico que respete la dignidad humana y los derechos fundamentales. Un ejercicio de cuidado y responsabilidad que las organizaciones deben atender de forma proactiva para asegurar una implementación confiable y socialmente aceptada por los diferentes grupos de interés (Poncibò, 2025).

Para lograr una ventaja competitiva sostenible con IA, las organizaciones deben integrar capacidades operativas únicas. Según Waltzman et al. (2020), el éxito no reside solo en la investigación, sino en la transición efectiva de la IA a aplicaciones específicas y en procesos robustos de verificación y validación. Abaimov y Martellini (2022) destacan que competitividad de la IA surge de procesar grandes volúmenes de datos para personalizar servicios y ganar eficiencia operativa. Además, priorizar la calidad del dato permite a empresas pequeñas reducir costos frente a grandes plataformas (Papadopoulos, 2025). Finalmente, la sostenibilidad de esta ventaja requiere una gobernanza que asegure la responsabilidad digital y la seguridad, transformando la confianza en un activo estratégico (Coles-Kemp & Burdon, 2025).

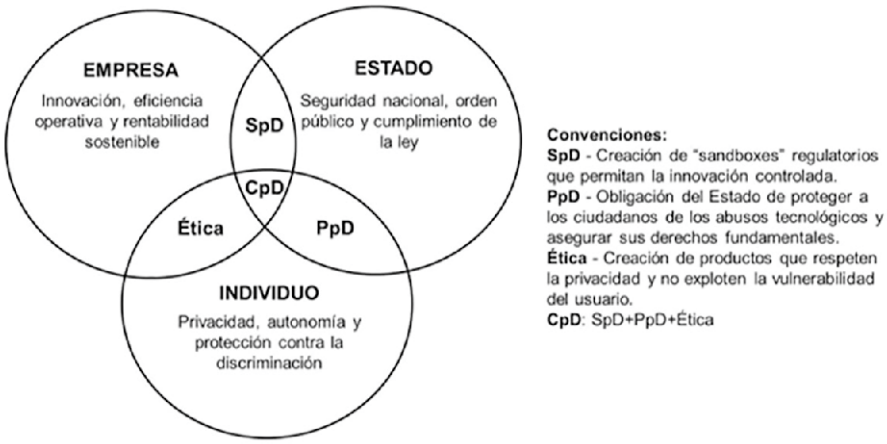
En resumen podríamos definir la confianza por diseño como un constructo relacional sociotécnico que armoniza las dimensiones de seguridad, ética y legalidad en el ecosistema de la IA, un ejercicio de equilibrio dinámico donde el mercado innova, el Estado protege y la sociedad participa, disminuyendo las lagunas de responsabilidad mediante una colaboración pluralista que trasciende la sola eficiencia técnica. En términos prácticos la CpD es la suma de la SpD+PpD+Ética, donde los requisitos de confiabilidad, explicabilidad, equidad, privacidad, seguridad y ética deben ser consideraciones de primer orden en el diseño arquitectónico de los sistemas inteligentes. La figura 1 detalla y resume el marco de trabajo detallado en esta sección.

Conclusiones

Hacia el futuro, la gobernanza de la IA Gen enfrenta desafíos claves. El primero es cerrar la “brecha de responsabilidad”, donde la complejidad de la cadena de suministro dificulta la atribución de daños causados por sistemas autónomos (Coles-Kemp & Burdon, 2025). A medida que los modelos desarrollan comportamientos emergentes impredecibles, los marcos legales tradicionales de negligencia y causalidad se ven tensionados, exigiendo nuevas formas de responsabilidad híbrida entre humanos y máquinas.

Un segundo reto crítico es la estandarización global. El reto de una

Figura 1. *Confianza por diseño*



Nota: Elaboración propia

visión homogénea de la IA enfrenta obstáculos críticos por la fragmentación de criterios nacionales e internacionales, generando un "mosaico global" de regulaciones que entorpece la innovación. El acelerado avance tecnológico supera los tiempos legislativos, impidiendo alcanzar la madurez necesaria para consensuar normas técnicas universales (Del-Real et al., 2025). Además, la opacidad algorítmica y la variabilidad del aprendizaje automático dificultan establecer métricas de seguridad y fiabilidad que sean aceptadas internacionalmente.

El tercer reto relevante es el riesgo de una "infodemia", ese flujo abundante de información que satura el ecosistema digital con desinformación automatizada y contenidos sintéticos, que impide distinguir hechos reales de datos fabricados, erosionando la confianza social en la tecnología y el conocimiento científico, exige mecanismos robustos de autenticación de contenido y marcas de agua digitales, los cuales deben asumirse como nuevas prácticas y estándares que enfrenten el reto de la integridad de la información, para configurar un estándar de debido cuidado de las empresas frente al despliegue de aplicaciones basadas en inteligencia artificial generativa (Abaimov & Martellini, 2022).

De esta forma, la Confianza por Diseño se debe configurar como un constructo jurídico de construcción colectiva que armoniza los intereses del mercado, el Estado y los individuos. Esta convergencia exige transitar hacia marcos constitutivos basados en la colaboración y el consenso entre las partes (Coles-Kemp & Burdon, 2025). Para

las empresas, implica incluir la integridad algorítmica en sus deberes fiduciarios, alineándose con la protección de derechos que el Estado asegura mediante la reglamentación vigente para la IA, y la apropiación de la ciudadanía, basada en la transparencia y la preservación de la autonomía frente a la "*hipersuación*".

Finalmente, la confianza por diseño es una propuesta para que la IA sea socialmente aceptable, tecnológicamente confiable y económicamente rentable. Solo mediante la incorporación proactiva de prácticas éticas y defensas técnicas desde la fase de diseño, junto con una conversación tripartita entre el Estado, el mercado y los ciudadanos, las organizaciones podrán navegar un futuro marcado por la incertidumbre, transformando los riesgos emergentes de la IA Gen, en un habilitador de progreso humano y bienestar para todos.

Referencias

Abaimov, S., & Martellini, M. (2022). *Machine learning for cyber agents: Attack and defence*. Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-030-91585-8

Cavoukian, A. (2009). Privacy by Design: The 7 Foundational Principles. Information and Privacy Commissioner of Ontario. https://www.ipc.on.ca/sites/default/files/legacy/2018/01/pbd-1.pdf

Coles-Kemp, L., & Burdon, M. (2025). *Understanding digital responsibilities*. Bristol University Press. https://doi.org/10.51952/9781529249798

Del-Real, C., De Busser, E., & van den Berg, B. (2025). A systematic literature review of security and privacy by design principles, norms, and strategies for digital technologies. *International Review of Law Computers & Technology*, *39*(3), 374–405. https://doi.org/10.1080/13600869.2025.2457227

European Parliament & Council of the European Union - EUPC (2024). *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. Official Journal of the European Union. http://data.europa.eu/eli/reg/2024/1689/oj

Floridi, L. (2023). AI as agency without intelligence: On ChatGPT, large language models, and other generative models. *Philosophy & Technology*, *36*(1), 1-15. https://doi.org/10.1007/s13347-023-00621-y

Habuca, H. & Socol de la Osa, D. (2025). Shaping Global AI Governance. A Path for the G7 to Foster Rule of Law in a World of Uncertainty. En Zou, M., Poncibò, C.,Ebers, M. & Calo, R. (Eds.), *The Cambridge Handbook of Generative AI and the Law*. Cambridge University Press.

National Institute of Standards and Technology (NIST). (2022). Engineering Trustworthy Secure Systems (NIST Special Publication 800-160, Vol. 1, Rev. 1). *U.S. Department of Commerce*. https://doi.org/10.6028/NIST.SP.800-160v1r1

Papadopoulos, S. (2025). Redefining rivalry: Generative AI and the evolving

landscape of competition law. En M. Zou, C. Poncibò, M. Ebers, & R. Calo (Eds.), *The Cambridge Handbook of Generative AI and the Law*. Cambridge University Press.

Paseri, L., & Durante, M. (2025). Normative and ethical dimensions of generative AI: From epistemological considerations to societal implications. En M. Zou, C. Poncibò, M. Ebers, & R. Calo (Eds.), *The Cambridge Handbook of Generative AI and the Law*. Cambridge University Press.

Poncibò, C. (2025). Regulating hypersuasion. En M. Zou, C. Poncibò, M. Ebers, & R. Calo (Eds.), *The Cambridge Handbook of Generative AI and the Law*. Cambridge University Press.

Saltzer, J. H., & Schroeder, M. D. (1975). The protection of information in computer systems. *Proceedings of the IEEE. Institute of Electrical and Electronics Engineers*, *63*(9), 1278–1308. https://doi.org/10.1109/proc.1975.9939

Sieber, S. (2026). La paradoja de la IA generativa: por qué más tecnología requiere más humanidad. *Harvard Deusto*, (358), 26-41. https://www.harvard-deusto.com/la-paradoja-de-la-ia-generativa-por-que-mas-tecnologia-requiere-mas-humanidad

Taddeo, M., McCutcheon, T., & Floridi, L. (2019). Trusting artificial intelligence in cybersecurity is a double-edged sword. *Nature Machine Intelligence*, *1*(12), 557–560. https://doi.org/10.1038/s42256-019-0109-1

Valdés-Rodríguez, Y., Hochstetter-Diez, J., Díaz-Arancibia, J., & Cadena-Martínez, R. (2023). Towards the integration of security practices in agile software development: A systematic mapping review. *Applied Sciences*, *13*(7), 4578. https://doi.org/10.3390/app13074578

Waltzman, R., Ablon, L., Curriden, C., Hartnett, G. S., Holliday, M. A., Ma, L., Nichiporuk, B., Scobell, A., & Tarraf, D. C. (2020). Maintaining the competitive advantage in artificial intelligence and machine learning. *RAND Corporation*. https://www.rand.org/pubs/research_reports/RRA200-1.html

Zou, M., Poncibò, C., Ebers, M., & Calo, R. (Eds.). (2025). *The Cambridge handbook of generative AI and the law*. Cambridge University Press. https://doi.org/10.1017/9781009492553

Jeimy J. Cano M., Ph.D, CFE, CICA. *Ingeniero y Magíster en Ingeniería de Sistemas y Computación, Universidad de los Andes. Especialista en Derecho Disciplinario, Universidad Externado de Colombia; Ph.D en Business Administration, Newport University, CA. USA. y Ph.D en Educación, Universidad Santo Tomás. Profesional certificado como Certified Fraud Examiner (CFE), por la Association of Certified Fraud Examiners y Certified Internal Control Auditor (CICA) por The Institute of Internal Controls. Profesor Distinguido de la Facultad de Derecho, Universidad de los Andes. Es director de la Revista SISTEMAS de la Asociación Colombiana de Ingenieros de Sistemas –ACIS–.*

El CISO como arquitecto de la confianza

Liderazgo, gobernanza y resiliencia digital en entornos NAVI latinoamericanos

DOI: 10.29236/sistemas.n179a8

Resumen

La evolución del riesgo digital ha transformado profundamente el rol del Chief Information Security Officer (CISO). En contextos caracterizados por aceleración tecnológica, interdependencia sistémica, volatilidad regulatoria y expansión de amenazas híbridas, la seguridad dejó de ser una función estrictamente técnica para convertirse en un componente estructural de la gobernanza organizacional. Este artículo examina el papel del CISO como articulador de confianza digital en organizaciones que operan en entornos NAVI—No lineales, Acelerados, Volátiles e Interconectados— con especial énfasis en América Latina. A partir de literatura académica, marcos de gobernanza y reportes especializados, se analiza cómo la confianza digital emerge de la interacción entre capacidades técnicas, liderazgo estratégico, comunicación organizacional y legitimidad institucional. El artículo propone el concepto de "arquitectura de confianza" como marco integrador para comprender la función contemporánea del CISO, entendida como la articulación de capacidades técnicas, organizacionales, comunicacionales y éticas que permiten sostener confianza digital en escenarios complejos. Asimismo, se reconocen los límites, tensiones y riesgos de sobredimensionar dicho rol. Se concluye que la confianza no puede entenderse únicamente como consecuencia de controles tecnológicos eficaces, sino como resultado de dinámicas organizacionales sostenidas que integran transparencia, resiliencia y capacidad adaptativa.

Palabras clave
CISO, confianza digital, gobernanza, liderazgo adaptativo, resiliencia organizacional, riesgo sistémico, ciberseguridad estratégica, América Latina.

Andrés R. Almanza Junco.

1. Introducción

Durante la última década, la ciberseguridad dejó de ocupar un lugar periférico en las organizaciones para convertirse en una preocupación de nivel estratégico. Los incidentes recientes demuestran que las brechas digitales ya no representan únicamente interrupciones técnicas: afectan reputación, continuidad operacional, legitimidad institucional y estabilidad económica.

El ataque contra instituciones del Estado costarricense en 2022 evidenció cómo un incidente cibernético puede escalar hacia una crisis nacional, afectando servicios públicos esenciales y obligando a declarar estado de emergencia (Mandiant, 2023; Porrúa et al., 2025). Casos similares en entidades financieras, compañías tecnológicas y organizaciones de salud en América Latina han demostrado que la fragilidad digital tiene consecuencias sistémicas.

En este contexto, el rol del Chief Information Security Officer (CISO) ha experimentado una transformación significativa. Tradicionalmente asociado con controles técnicos, gestión de vulnerabilidades y supervisión operativa, el CISO contemporáneo enfrenta expectativas crecientemente vinculadas con liderazgo, gobernanza y comunicación estratégica.

Sin embargo, esta evolución plantea una pregunta crítica: ¿hasta qué punto la seguridad puede entenderse únicamente como una capacidad tecnológica? La evidencia sugiere que los incidentes más disruptivos rara vez se explican exclusivamente por fallas técnicas. Frecuentemente involucran decisiones organizacionales, debilidades culturales, fragmentación de responsabilidades y deterioro de la confianza institucional.

A partir de esta premisa, este artículo sostiene que el valor estratégico del CISO no reside únicamente en proteger infraestructuras digitales, sino en contribuir a la construcción y sostenimiento de confianza digital dentro y fuera de la organización. No obstante, esta afirmación requiere matices importantes. La confianza no depende exclusivamente del CISO ni puede ser producida unilateralmente desde la función de seguridad. Más bien, emerge de la interacción entre estructuras de gobernanza,

capacidades organizacionales, comunicación transparente y comportamiento institucional consistente.

Para analizar esta problemática, el artículo adopta el concepto de entornos NAVI (EY, 2025) —No lineales, Acelerados, Volátiles e Interconectados— desarrollado por la firma de consultoría EY, como evolución conceptual de marcos previos como VUCA y aproximaciones posteriores como BANI.

El concepto NAVI busca enfatizar la intensidad de las interdependencias digitales y la velocidad con que las crisis pueden propagarse entre organizaciones, industrias y sociedades. Aunque NAVI representa una aproximación emergente proveniente del ámbito de consultoría estratégica, resulta útil como lente interpretativo para comprender dinámicas contemporáneas de interdependencia digital. Desde esta perspectiva, el artículo explora tres preguntas principales:

1. ¿Por qué la confianza digital se ha convertido en un componente central de la gobernanza contemporánea?
2. ¿Qué papel puede desempeñar el CISO en la construcción de dicha confianza?
3. ¿Cuáles son los límites y tensiones de atribuirle al CISO un rol de “arquitecto de confianza”?

La intención no es idealizar la función del CISO ni convertirla en solución universal frente al riesgo digital. Por el contrario, el objetivo es comprender cómo la seguridad, el liderazgo y la confianza convergen en organizaciones cada vez más dependientes de sistemas digitales complejos.

2. Entornos NAVI y transformación del riesgo organizacional

Las organizaciones contemporáneas operan en escenarios crecientemente complejos. La digitalización acelerada, la hiperconectividad y la dependencia de ecosistemas tecnológicos externos han modificado profundamente la naturaleza del riesgo corporativo.

El modelo VUCA —Volatility, Uncertainty, Complexity and Ambiguity— surgió inicialmente en contextos militares y posteriormente fue adoptado por la literatura de management para describir entornos caracterizados por incertidumbre y cambio acelerado (Bennett & Lemoine, 2014). Posteriormente, surgieron aproximaciones como BANI, que intentaron describir la fragilidad emocional y sistémica de los entornos contemporáneos. Más recientemente, EY propuso el modelo NAVI como evolución conceptual orientada a enfatizar dinámicas de no linealidad, aceleración, volatilidad e interconexión digital (Bax & Jaggi, 2025).

El concepto NAVI utilizado en este artículo busca enfatizar cuatro características estructurales:

- **No linealidad:** pequeñas vulnerabilidades pueden desencadenar impactos desproporcionados.
- **Aceleración:** las amenazas evolucionan más rápido que muchos procesos organizacionales.
- **Volatilidad:** las condiciones regulatorias, tecnológicas y reputacionales cambian constantemente.
- **Interconexión:** las organizaciones dependen de redes complejas de proveedores, plataformas y terceros.

El caso SolarWinds ilustra claramente esta lógica sistémica. La manipulación de actualizaciones de software permitió comprometer miles de organizaciones mediante una relación de confianza preexistente con un proveedor tecnológico (CISA, 2021; Mandiant, 2021). El incidente mostró que la superficie de riesgo ya no puede entenderse únicamente desde límites organizacionales internos.

Esta transformación tiene implicaciones profundas para la gobernanza corporativa. Los riesgos digitales dejaron de ser problemas aislados de departamentos tecnológicos para convertirse en riesgos estratégicos capaces de afectar continuidad operativa, reputación, cumplimiento regulatorio y estabilidad financiera.

3. Liderazgo adaptativo, transformacional y estratégico

Heifetz y Linsky (2002) acuñaron el concepto de liderazgo adaptativo para describir la capacidad de movilizar a las personas hacia la resolución de problemas que no tienen soluciones técnicas predefinidas. En este sentido, el CISO enfrenta un desafío típicamente adaptativo: la ciberseguridad no es un problema que se resuelve instalando una herramienta; es un proceso continuo de aprendizaje, adaptación y negociación cultural dentro de la organización.

Burns (1978) y Bass (1985) desarrollaron el concepto de liderazgo transformacional, que distingue entre líderes que transaccionan —dan y reciben dentro de reglas establecidas— y líderes que transforman —cambian las reglas, los valores y las motivaciones de sus seguidores. El CISO arquitecto de confianza opera en el espacio transformacional: no solo administra riesgos, sino que cambia la cultura de seguridad de la organización, eleva la madurez digital y redefine cómo la junta directiva y los equipos operativos perciben su función.

El liderazgo estratégico, por su parte, se define como la capacidad de anticipar, visionar, mantener flexibilidad y delegar autoridad para crear cambio estratégico cuando sea necesario (Ireland & Hitt, 1999). Para el CISO, esto significa operar simultáneamente en el presente operacional —respondiendo a incidentes— y en el futuro estratégico

—diseñando arquitecturas de seguridad alineadas con los objetivos del negocio.

4. Confianza digital y gobernanza organizacional

La confianza ha sido ampliamente estudiada en sociología, psicología organizacional y teoría institucional. Rousseau et al. (1998) la definen como la disposición a aceptar vulnerabilidad basada en expectativas positivas sobre la conducta de otro. Mayer, Davis y Schoorman (1995) complementan esta visión identificando tres factores fundamentales: competencia, integridad y benevolencia.

En contextos digitales, la confianza adquiere características particulares. Los individuos dependen cotidianamente de sistemas que no comprenden completamente, pero cuya confiabilidad deben asumir para operar social y económicamente. Luhmann (1979) y Giddens (1990) ya habían advertido que la modernidad depende crecientemente de la confianza en sistemas abstractos e infraestructuras complejas.

La confianza digital puede entenderse entonces como la expectativa razonable de que una organización gestionará de forma segura, ética y consistente los sistemas y datos que administra (WEF, 2022; WEFb, 2022). Según el WEF (2024), el 80% del valor de mercado de las empresas del S&P 500 corresponde a activos intangibles. La protección de estos activos no es responsabilidad exclusiva del CISO, pero su compromiso catastrófico —a través de una brecha de datos, un ataque de ransomware o una exposición regulatoria— puede destruir en horas lo que tomó años construir, esa confianza que requiere para desarrollarse en un ecosistema digital de gran tamaño.

Esta definición tiene implicaciones relevantes. La confianza digital no depende exclusivamente de la ausencia de incidentes. Ninguna organización puede garantizar inmunidad absoluta frente a amenazas sofisticadas. La confianza depende más bien de cómo la organización: gestiona sus riesgos, responde ante crisis, comunica incidentes, protege a sus partes interesadas, y demuestra coherencia institucional.

Finalmente, hay una razón de competitividad. En mercados donde la diferenciación es difícil y los consumidores son cada vez más conscientes del valor de su privacidad, la confianza digital se convierte en ventaja competitiva. Según Edelman (2023), el 71% de los consumidores globales considera que la confianza en una empresa es un factor determinante en sus decisiones de compra.

En América Latina, donde la desconfianza institucional es históricamente alta, las organizaciones que logran construir reputaciones sóli-

das de seguridad y transparencia digital obtienen una ventaja relacional significativa.

Desde esta perspectiva, la ciberseguridad deja de ser únicamente una disciplina de control para convertirse en un componente de legitimidad organizacional.

5. El CISO y la evolución hacia un rol estratégico

La literatura temprana sobre seguridad de la información estuvo dominada por enfoques económicos y técnicos centrados en optimización de inversiones y controles (Gordon & Loeb, 2002; Bodin et al., 2008).

En estos modelos, el rol del CISO aparecía principalmente como gestor operativo.

Posteriormente, investigaciones sobre comportamiento organizacional comenzaron a demostrar que muchos incidentes relevantes no derivaban exclusivamente de fallas tecnológicas, sino de problemas culturales, humanos y organizacionales (Crossler et al., 2013; Bada, M., & Sasse, A., 2014).

Este cambio amplió progresivamente las expectativas sobre el CISO. Organismos como ISACA (2023), NACD (2023) y Ribot (2025) han enfatizado la necesidad de que los líderes de seguridad desarrollen competencias vinculadas con: comunicación ejecutiva, influencia organizacional, pensamiento sistémico, liderazgo adaptativo, y gestión estratégica del riesgo.

La gobernanza corporativa contemporánea también ha contribuido a esta transformación. La ISO 37000:2021 y los principios de gobierno corporativo de la OCDE reconocen que los riesgos digitales deben formar parte de la supervisión estratégica de las organizaciones (OCDE, 2023).

En consecuencia, el CISO ya no opera exclusivamente como especialista técnico. En muchas organizaciones, se ha convertido en traductor entre complejidad tecnológica y toma de decisiones estratégicas. Así mismo, el Banco Interamericano de Desarrollo (BID & OEA, 2020) en su estudio sobre el impacto económico de los ciberataques en América Latina estimó pérdidas anuales superiores a 90.000 millones de dólares en la región, con efectos desproporcionadamente severos en economías con menor capacidad de respuesta institucional. En este escenario, el CISO que no opera como arquitecto de confianza no solo falla en su función técnica; falla en su responsabilidad estratégica con la organización y con el ecosistema al que pertenece.

Sin embargo, esta transición también introduce riesgos. Existe una tendencia creciente a sobredimensionar el rol del CISO, atribuyéndole responsabilidades que exceden sus capacidades reales o su autori-

dad institucional. La confianza organizacional no puede recaer únicamente en un individuo o función específica. Depende de estructuras más amplias de liderazgo y cultura corporativa.

6. Hacia una arquitectura de confianza

El concepto de arquitectura de confianza utilizado en este artículo se apoya en aproximaciones contemporáneas sobre *digital trust, governance by design* y resiliencia institucional promovidas por organismos como el WEF (2022), que plantean que la confianza digital debe incorporarse desde el diseño organizacional y no únicamente como capacidad reactiva de seguridad.

En este contexto, la arquitectura de confianza puede entenderse como la articulación estructurada de capacidades técnicas, organizacionales, comunicacionales y éticas que permiten generar, sostener y recuperar confianza digital en entornos complejos.

Esta aproximación no debe entenderse como un modelo cerrado ni como una metodología formalmente validada. Más bien, funciona como un marco integrador para comprender cómo distintas capacidades organizacionales contribuyen al sostenimiento de confianza digital.

Esta arquitectura involucra al menos cuatro dimensiones interdependientes.

6.1. Dimensión técnica

Incluye controles de seguridad, resiliencia tecnológica, monitoreo, respuesta a incidentes y protección de infraestructuras críticas. La competencia técnica sigue siendo condición necesaria para la confianza. Un liderazgo carismático sin capacidad operacional difícilmente puede sostener legitimidad frente a incidentes reales.

6.2. Dimensión organizacional

La confianza también depende de estructuras de gobernanza claras, roles definidos y alineación entre seguridad y estrategia corporativa. La posición del CISO dentro de la estructura organizacional resulta particularmente relevante. Estudios de ISACA (2023) sugieren que organizaciones donde el CISO posee acceso directo a niveles ejecutivos muestran mayores niveles de madurez y capacidad de respuesta. No obstante, la relación no es automática. Reportar al CEO o al directorio no garantiza influencia real ni transformación cultural.

6.3. Dimensión conversacional

La confianza posee un componente profundamente interpretativo. Las partes interesadas evalúan no solo lo que una organización hace, sino cómo comunica sus acciones. En escenarios de crisis, la transparencia se convierte en variable crítica. La comunicación en ciberseguridad requiere equilibrio entre honestidad, prudencia estratégica y manejo reputacional. Así las cosas, la confianza no depende

exclusivamente de mensajes correctos, sino de trayectorias organizacionales consistentes.

6.4. Dimensión ética y relacional

La confianza también involucra percepciones de integridad y coherencia institucional. El caso de Rappi en Colombia resulta ilustrativo porque muestra que la erosión de confianza puede originarse no solo en ciberataques, sino también en problemas de gobernanza del dato y cumplimiento regulatorio (Quinchía, 2021; Bloomberg Línea, 2021). Esto amplía el alcance de la discusión. La confianza digital no depende únicamente de proteger sistemas, sino de gestionar responsablemente la información y las relaciones con las partes interesadas.

Para pasar a través de todas estás dimensiones y poder operacionalizar estos elementos, una arquitectura de confianza podría observarse en organizaciones que integran al menos cinco capacidades: (1) supervisión ejecutiva del riesgo digital, (2) resiliencia operacional medible, (3) comunicación transparente de incidentes, (4) gobernanza de terceros y cadena de suministro, y (5) existencia de comités de revisión y rendición de cuentas, sobre decisiones tecnológicas (*accountability*).

Aunque estas capacidades no constituyen un modelo universal ni una metodología cerrada, pueden servir como indicadores observables de madurez organizacional en la construcción de confianza digital. Su relevancia no radica únicamente en la existencia formal de controles, sino en la capacidad de la organización para integrarlos dentro de procesos sostenibles de gobernanza, resiliencia y toma de decisiones. Para ello se ilustran estos elementos en la tabla 1.

7. Tensiones y límites del modelo

La idea del CISO como arquitecto de confianza posee utilidad estratégica, pero también enfrenta limitaciones importantes.

7.1. La paradoja de la transparencia

Una comunicación excesivamente abierta sobre vulnerabilidades puede producir efectos contraproducentes: deterioro reputacional, pérdida de confianza de inversionistas, incremento de presión regulatoria, o exposición frente a actores maliciosos.

La transparencia en ciberseguridad no puede entenderse como valor absoluto. Requiere criterio estratégico y comprensión contextual.

7.2. El riesgo de hipercentralización del CISO

Existe una tendencia creciente a convertir al CISO en símbolo organizacional de confianza digital. Sin embargo, esto puede generar expectativas imposibles de sostener. La seguridad depende de: cultura organizacional, recursos, decisiones ejecutivas, arquitectura tec-

Tabla 1. *Capacidades de una arquitectura de confianza*

Capacidad	Propósito organizacional	Indicadores o métricas observables
Supervisión ejecutiva del riesgo digital	Integrar el riesgo digital dentro de la gobernanza corporativa	Frecuencia de revisión del riesgo en junta directiva, participación del CISO en comités ejecutivos, existencia de indicadores de riesgo digital (KRIs)
Resiliencia operacional medible	Evaluar capacidad de continuidad recuperación y aprendizaje en las crisis cibernéticas	MTTR (*Mean Time to Recovery*), RTO/RPO, tiempo promedio de contención de incidentes, pruebas de continuidad ejecutadas
Comunicación transparente de incidentes	Preservar legitimidad y coordinación durante crisis	Tiempo de notificación, consistencia comunicacional, existencia de protocolos de crisis, cumplimiento regulatorio de divulgación
Gobernanza de terceros y cadena de suministro	Reducir exposición sistémica derivada de interdependencias	Evaluaciones de terceros, cobertura contractual de seguridad, monitoreo continuo de proveedores críticos. Mapeo y monitoreo de terceros críticos que soportan procesos esenciales del negocio que soportan las unidades claves de negocio.
Accountability sobre decisiones tecnológicas	Asegurar trazabilidad y responsabilidad organizacional	Registro de decisiones críticas, auditorías, mecanismos de supervisión ética y validación de riesgos tecnológicos

Nota: *Elaboración propia basada en principios de gobernanza digital, resiliencia organizacional y gestión de riesgo cibernético inspirados en marcos de NIST CSF 2.0 (2024), ISO 22301(2019) y World Economic Forum WEF (2021).*

nológica, cumplimiento regulatorio, y comportamiento humano colectivo.

Un CISO competente en una organización estructuralmente disfuncional difícilmente podrá construir confianza sostenible.

7.3. Fatiga y sostenibilidad del rol

La investigación reciente muestra altos niveles de agotamiento entre CISOs. Heidrick & Struggles (2023) reporta una permanencia promedio cercana a dos años en muchas organizaciones.

Esto introduce una contradicción estructural: la confianza requiere continuidad, mientras que el rol frecuentemente opera bajo presión extrema, alta exposición y expectativas ambiguas.

La sostenibilidad del liderazgo en ciberseguridad emerge, así como un problema organizacional y no exclusivamente individual.

7.4. Limitaciones regionales latinoamericanas

En América Latina, las dinámicas de confianza digital se desarrollan en un entorno marcado por profundas asimetrías institucionales, económicas y tecnológicas. Estas diferencias impactan directamente la capacidad de gobiernos y organizaciones para construir modelos sostenibles de resiliencia y gobernanza digital. El *2025 Cybersecurity Report* del Banco Interamericano de Desarrollo (BID) advierte que la región mantiene niveles heterogéneos de madurez en capacidades nacionales de ciberseguridad, preparación institucional, resiliencia operacional y coordinación estratégica frente a incidentes cibernéticos (Porrúa et al., 2025). Esta disparidad no solo ocurre entre países, sino también entre sectores económicos y organizaciones dentro de un mismo ecosistema nacional.

Adicionalmente, el informe conjunto de la Organización de los Estados Americanos (OEA) y el BID sobre ciberseguridad en América Latina y el Caribe señala que persisten desafíos estructurales relacionados con restricciones presupuestarias, fragmentación regulatoria, baja integración estratégica de la ciberseguridad y dependencia tecnológica externa (BID & OEA, 2020). Estas condiciones generan contextos donde la confianza digital no puede depender únicamente de capacidades tecnológicas, sino también de factores organizacionales, regulatorios e institucionales más amplios.

La región también enfrenta brechas significativas de digitalización e inclusión tecnológica. El Programa de las Naciones Unidas para el Desarrollo (PNUD) ha señalado que la digitalización en América Latina avanza de manera desigual y que las limitaciones de acceso, capacidades digitales y apropiación tecnológica continúan profundizando brechas económicas y sociales (Programa de las Naciones Unidas para el Desarrollo [PNUD], 2023). En una línea similar, el Banco Mundial advierte que el acceso desigual a infraestructura digital y conectividad limita la competitividad regional y ralentiza la capacidad de transformación digital sostenible (Banco Mundial, 2022).

Desde una perspectiva institucional, la Comisión Económica para América Latina y el Caribe (CEPAL) ha destacado que los procesos de transformación digital en la región siguen condicionados por niveles

heterogéneos de desarrollo estatal, gobernanza digital y articulación de políticas públicas, generando capacidades desiguales para responder a riesgos emergentes asociados a entornos digitales (CEPAL, 2024). Complementariamente, la Organización para la Cooperación y el Desarrollo Económicos (OCDE) identifica que muchos países latinoamericanos aún enfrentan desafíos relevantes en integración de gobierno digital, interoperabilidad institucional y fortalecimiento de capacidades públicas para la gestión de ecosistemas digitales complejos (OCDE, 2023).

En este contexto, modelos de gobernanza y confianza digital desarrollados en entornos norteamericanos o europeos no siempre resultan directamente transferibles a realidades latinoamericanas caracterizadas por volatilidad institucional, restricciones operativas y madurez desigual entre sectores económicos. A ello se suma un elemento particularmente crítico: la confianza, entendida como componente esencial de cohesión social y crecimiento económico, continúa siendo un desafío estructural para la región. El BID ha señalado que los bajos niveles de confianza interpersonal e institucional afectan la cooperación, la legitimidad organizacional y la capacidad de articular respuestas sostenibles frente a desafíos colectivos (BID, 2022).

Finalmente, la creciente presión reputacional alrededor de conceptos como resiliencia, transformación digital y confianza también introduce el riesgo de que algunas organizaciones privilegien narrativas de legitimidad sobre transformaciones estructurales reales de seguridad y gobernanza. En consecuencia, la confianza digital no puede reducirse a cumplimiento normativo o comunicación corporativa, sino que requiere capacidades sostenibles, verificables y articuladas dentro de modelos consistentes de resiliencia organizacional.

8. Discusión: confianza, resiliencia y legitimidad institucional

La discusión contemporánea sobre ciberseguridad frecuentemente oscila entre dos extremos problemáticos.

El primero es el reduccionismo técnico: asumir que la seguridad depende fundamentalmente de herramientas, controles y automatización. El segundo es el reduccionismo narrativo: asumir que liderazgo, comunicación o cultura pueden compensar debilidades estructurales de seguridad. Ambas posiciones son insuficientes.

La evidencia sugiere que la resiliencia organizacional emerge precisamente de la interacción entre capacidades técnicas y legitimidad institucional. Las organizaciones más resilientes no son necesariamente aquellas que evitan todos los incidentes, sino aquellas capaces de: absorber impactos, adap-

tarse rápidamente, preservar coordinación interna, y sostener confianza externa durante crisis.

En este contexto, el CISO puede desempeñar un papel relevante como articulador entre dominios tradicionalmente separados: tecnología, negocio, riesgo, cumplimiento, comunicaciones, y gobernanza.

Sin embargo, su efectividad dependerá menos de atributos heroicos individuales y más de la capacidad organizacional para integrar la seguridad dentro de su modelo de decisión estratégica.

Conclusiones

La transformación digital ha convertido la confianza en un componente central de la competitividad y legitimidad organizacional. En entornos caracterizados por interdependencia tecnológica y riesgo sistémico, la seguridad ya no puede entenderse exclusivamente como problema operativo.

El rol del CISO ha evolucionado en respuesta a esta realidad. Cada vez más organizaciones esperan que sus líderes de seguridad participen en conversaciones estratégicas, traduzcan complejidad técnica y contribuyan a la resiliencia institucional. No obstante, esta evolución requiere evitar simplificaciones.

El CISO no puede ser concebido como único responsable de la confianza digital. La confianza emerge de sistemas organizacionales más amplios donde intervienen gobernanza, cultura, liderazgo ejecutivo, ética institucional y capacidad adaptativa.

El concepto de arquitectura de confianza propuesto en este artículo busca precisamente enfatizar esa naturaleza sistémica. Su valor no reside en convertir al CISO en figura centralizadora, sino en reconocer que la seguridad contemporánea depende de relaciones organizacionales complejas que exceden la dimensión tecnológica.

En América Latina, este desafío adquiere particular relevancia debido a contextos marcados por volatilidad institucional, madurez desigual y aceleración digital asimétrica.

Las organizaciones que comprendan esta complejidad probablemente estarán mejor preparadas no solo para enfrentar incidentes, sino para sostener legitimidad y resiliencia en escenarios crecientemente inciertos.

En última instancia, la confianza digital no representa un estado permanente ni una garantía absoluta. Es una construcción dinámica, vulnerable y continuamente negociada entre organizaciones, tecnologías y sociedades. El valor estratégico del CISO contemporáneo reside, quizás, no en prometer control total sobre dicha incertidumbre, sino en ayudar a las organizacio-

nes a gestionarla con mayor transparencia, coherencia y capacidad adaptativa.

Referencias

Bada, M., & Sasse, A. (2014). Cyber security awareness campaigns: Why do they fail to change behaviour? *Proceedings of the International Conference on Cyber Security for Sustainable Society*. https://discovery.ucl.ac.uk/id/eprint/1468954/

Bass, B. M. (1985). *Leadership and performance beyond expectations*. Free Press.

Bax, H. J., & Jaggi, G. (2025). *What if disruption isn't the challenge, but the chance?* EY. https://www.ey.com/en_gl/megatrends/what-if-disruption-is-not-the-challenge-but-the-chance

Bennett, N., & Lemoine, G. J. (2014). What a difference a word makes: Understanding threats to performance in a VUCA world. *Business Horizons, 57*(3), 311–317. https://www.sciencedirect.com/science/article/abs/pii/S0007681314000020

Banco Interamericano de Desarrollo (BID), & Organización de los Estados Americanos (OEA). (2020). *Reporte Ciberseguridad 2020: Riesgos, avances y el camino a seguir en América Latina y el Caribe*. https://publications.iadb.org/es/publications/spanish/viewer/Reporte-Ciberseguridad-2020-riesgos-avances-y-el-camino-a-seguir-en-America-Latina-y-el-Caribe.pdf

Banco Interamericano de Desarrollo (BID). (2022). *Confianza: la clave de la cohesión social y el crecimiento en América Latina y el Caribe* (Resumen ejecutivo). https://publications.iadb.org/es/publications/spanish/viewer/Confianza-la-clave-de-la-cohesion-social-y-el-crecimiento-en-America-Latina-y-el-Caribe-Resumen-ejecutivo.pdf

Banco Mundial. (2022). *El escaso acceso digital frena a América Latina y el Caribe: cómo solucionarlo*. https://blogs.worldbank.org/es/latinamerica/el-escaso-acceso-digital-frena-america-latina-y-el-caribe-como-solucionar-este

Bloomberg Línea. (2021). Multan a Rappi en Colombia por violaciones al régimen de protección de datos. https://www.bloomberglinea.com/2021/10/29/multan-a-rappi-en-colombia-por-violaciones-al-regimen-de-proteccion-de-datos/

Bodin, L., Gordon, L., & Loeb, M. (2008). Information security and risk management. *Communications of the ACM, 51*(4), 64–68. https://doi.org/10.1145/1330311.1330325

Burns, J. M. (1978). *Leadership*. Harper & Row.

Comisión Económica para América Latina y el Caribe (CEPAL). (2024). *América Latina y el Caribe en la segunda mitad de la década digital*. https://repositorio.cepal.org/server/api/core/bitstreams/e4ca636c-2b8a-4138-8c62-b685540d9b99/content

Crossler, R. E., Johnston, A. C., Lowry, P. B., Hu, Q., Warkentin, M., & Baskerville, R. (2013). Future directions for behavioral information security research. *Computers & Security, 32*, 90–101. https://doi.org/10.1016/j.cose.2012.09.010

Cybersecurity and Infrastructure Security Agency. (2021). *Advanced persistent*

threat compromise of government agencies, critical infrastructure, and private sector organizations. CISA Advisory AR21-112A

Edelman. (2023). *Edelman trust barometer 2023: Global report*. Edelman Trust Barometer 2023

Giddens, A. (1990). *The consequences of modernity*. Stanford University Press.

Gordon, L. A., & Loeb, M. P. (2002). The economics of information security investment. *ACM Transactions on Information and System Security, 5*(4), 438–457. https://doi.org/10.1145/581271.581274

Heidrick & Struggles. (2023). *Global CISO survey 2023: The evolving role of the chief information security officer*. https://www.heidrick.com/en/insights/cybersecurity/2023-global-chief-information-security-officer-survey

Heifetz, R., & Linsky, M. (2002). *Leadership on the line: Staying alive through the dangers of leading*. Harvard Business Review Press.

Ireland, R. D., & Hitt, M. A. (1999). Achieving and maintaining strategic competitiveness in the 21st century: The role of strategic leadership. *Academy of Management Perspectives, 13*(1), 43–57. https://doi.org/10.5465/ame.1999.1567311

ISACA. (2023). *State of cybersecurity 2023: Global update on workforce efforts, resources and cyberoperations*. https://www.isaca.org/resources/reports/state-of-cybersecurity-2023

International Organization for Standardization. (2021). *ISO 37000:2021 — Governance of organizations: Guidance*. ISO.

ISO. (2019). *Security and resilience — Business continuity management systems — Requirements*. ISO. https://www.iso.org/standard/75106.html

Luhmann, N. (1979). *Trust and power*. John Wiley & Sons.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review, 20*(3), 709–734. https://doi.org/10.5465/amr.1995.9508080335

Mandiant. (2021). *M-Trends 2021 Special Report*. https://services.google.com/fh/files/misc/rpt-mtrends-2021-en.pdf

Mandiant. (2023). *M-Trends 2023 Special Report*. https://www.mandiant.com/resources/reports/m-trends-2023-special-report?auHash=iTAkoIVQOJBJJ8XvjFW34_KB6WJNeNAZ1HV2I3AEXdE&ref=guptadeepak.com

National Association of Corporate Directors. (2023). *Director's handbook on cyber-risk oversight* (4th ed.). https://isalliance.org/wp-content/uploads/2023/03/Cyber-Risk-Oversight-Handbook_WEB.pdf

NIST. (2024). The NIST cybersecurity framework (CSF) 2.0. *The NIST Cybersecurity Framework (CSF) 2.0*, *2.0*(29). https://doi.org/10.6028/nist.cswp.29

Organisation for Economic Co-operation and Development (OECD). (2023). *Digital government review of Latin America and the Caribbean*. https://www.oecd.org/content/dam/oecd/es/publications/reports/2023/09/digit

al-government-review-of-latin-america-and-the-caribbean_75a4be05/7a127615-es.pdf

Programa de las Naciones Unidas para el Desarrollo (PNUD). (2023). *La digitalización: motor de inclusión y crecimiento en América Latina*. https://www.undp.org/es/peru/noticias/la-digitalizacion-motor-de-inclusion-y-crecimiento-en-america-latina

Porrúa, M., Moncayo, G., Paz, S., Nowersztern, A., Bejarano, J. F., Baudino, M. F., Bordese, M. P., Barret, K., Baena, C. E., Jaramillo, M., Garces, O., & Isidro, A. (2025). *2025 Cybersecurity Report: Vulnerability and Maturity Challenges to Bridging the Gaps in Latin America and the Caribbean*. https://doi.org/10.18235/0013872

Quinchía, A. Z. (2021). Superindustria multa a Rappi por violación a protección de datos. *El Colombiano*. https://www.elcolombiano.com/negocios/multan-a-rappi-por-violacion-al-regimen-de-proteccion-de-datos-personales-JI15954739

Ribot, S. (2025). *The CISO Dilemma*. Kornferry.com. https://www.kornferry.com/institute/the-ciso-dilemma

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review, 23*(3), 393–404. https://doi.org/10.5465/amr.1998.926617

WEF. (2021). *Principles for Board Governance of Cyber Risk*. World Economic Forum. https://www.weforum.org/publications/principles-for-board-governance-of-cyber-risk/

WEF. (2022). *Earning Digital Trust: Decision-Making for Trustworthy Technologies*. World Economic Forum. https://www.weforum.org/publications/earning-digital-trust-decision-making-for-trustworthy-technologies/

WEF. (2022b). *Principles for board governance of cyber risk*. https://www.weforum.org/publications/principles-for-board-governance-of-cyber-risk/

World Economic Forum. (2024). *Global risks report 2024*. https://www.weforum.org/publications/global-risks-report-2024/

Andrés Ricardo Almanza Junco CISM, Ingeniero de Sistemas y Especialista de Seguridad de la Univesidad Católica de Colombia, Master en Seguridad de la Información de la Universidad Oberta de Catalunia, certificado como CISM por ISACA Internacional, Certificado como ISO 27001 Senior Lead Implementer and 27005 Lead Manager from PECB, Formación Ejecutiva Líderes Globales, Business Administration and Management por la Universidad de los Andes, Executive Certificate in Cybersecurity Leadership & Strategy por Florida Internation University, Certificado como Coach Profesional Internacional by INILID | Master in Leadership and Organizational Development with Coaching & Executive Master's in Leadership Skills Developed in Harvard & Coach Profesional avalado por International Coach Federation by EIDHI International University – USA. Profesor de la Universidad Externado de Colombia, director general de la Asociación de Profesionales de Seguridad y Ciberseguridad APSIC y CISOS.CLUB.

ACIS

ASOCIACIÓN COLOMBIANA DE INFORMÁTICA, SISTEMAS Y TECNOLOGÍAS AFINES

XL MARATÓN NACIONAL DE PROGRAMACIÓN ACIS/REDIS 2026

3 de octubre de 2026

LA COMPETENCIA DE EXCELENCIA

Clasificatoria a la Regional Latinoamericana y la Final Mundial del ICPC.

DETALLES DEL EVENTO

- Modalidad: 100% presencial
- Equipos: 3 estudiantes y 1 coach

SEDES NACIONALES

- BOGOTÁ, BUCARAMANGA
- CALI, MEDELLÍN
- MANIZALES, CARIBE

INSCRIPCIÓN: maraton@acis.org.co

ACIS

ASOCIACIÓN COLOMBIANA DE INFORMÁTICA, SISTEMAS Y TECNOLOGÍAS AFINES

ACIS Conecta

*Desde **ACIS** queremos invitarte a ser parte de ACIS Conecta, una nueva iniciativa creada especialmente para nuestros asociados.*

Actualiza tu perfil, destaca tus competencias y prepárate para conectarte con nuevas oportunidades profesionales. Gracias a nuestras alianzas con el Gobierno, el sector empresarial y organismos internacionales, podrás acceder a proyectos, vacantes especializadas y programas de formación que impulsarán tu crecimiento.

¡Aplica a nuestra encuesta!

acis.org.co/acis-conecta

¡AFILIATE YA!

Afilicación a ACIS + Estudio de Formulario
$ 342.000 + 63.000 COP

Para más información visita

https://www.acis.org.co/afiliate

301 5530540 administrador@acis.org.co www.acis.org.co