

BANNER DISPONIBLE
27 UF SEMESTRE

Inicio » Tecnología

Tecnología

Noticias

Auditar antes de acelerar: el control que necesitan los asesores, algoritmos y agentes de IA

Por Portal InnoVA | Grupo Prensa Digital | I.V - septiembre 21, 2026



X Post



José Lagos Docente UEjecutivos Facultad de Economía y Negocios Universidad de Chile

El desarrollo de inteligencia artificial ha entrado en una etapa en la que avanzar más rápido ya no significa necesariamente avanzar mejor. En su artículo We Must Pace the Frontier, Dario Amodei, director ejecutivo de Anthropic, sostiene que las capacidades de los modelos están progresando a una velocidad que amenaza con superar nuestra capacidad para comprenderlos, alinearlos y controlarlos. Su propuesta no consiste en detener la innovación, sino en acompañar el aumento de capacidades con mecanismos verificables de seguridad, evaluación independiente y supervisión.

Este planteamiento también resulta aplicable a cualquier organización que diseñe asesores de IA generativa, Gems, Skills, algoritmos de machine learning o agentes capaces de ejecutar acciones. En este contexto, la organización debe demostrar que funciona dentro de límites aceptables y la función de auditoría se transforma en una condición para avanzar.

Las soluciones de IA no presentan la misma exposición. Un asistente que redacta borradores tiene un impacto diferente al de un modelo que rechaza créditos o un agente que modifica registros. Sin embargo, todos requieren evidencia sobre su propósito, datos, instrucciones, comportamiento, controles y responsables.

En este caso, la auditoría debe comenzar antes de producción. Esto implica definir el caso de uso, los usuarios autorizados, las decisiones permitidas, los datos procesados y las acciones prohibidas. También exige métricas, pruebas de aceptación, trazabilidad, monitoreo y criterios de suspensión. Acelerar sin estas definiciones genera deuda técnica, ética, regulatoria y operacional.

Los asesores de IA generativa suelen percibirse como aplicaciones simples porque operan mediante instrucciones y lenguaje natural. No obstante, pueden revelar información personal o confidencial, entregar respuestas incorrectas, reproducir sesgos, citar fuentes inexistentes o inducir decisiones erróneas. Su auditoría debe revisar el propósito y alcance; la calidad y seguridad de las instrucciones; las fuentes; los permisos; la retención de conversaciones; la resistencia frente a prompt injection; y las salvaguardas para contenidos sensibles. El usuario debe saber que interactúa con una IA, conocer sus limitaciones y poder escalar el caso a una persona.

En este aspecto, no basta con revisar el prompt principal. Es necesario probar el comportamiento mediante escenarios normales, extremos y adversariales. Esto incluye las solicitudes ambiguas, intentos de extraer instrucciones internas, carga de datos personales, peticiones contrarias a la ética y preguntas fuera del dominio autorizado. El resultado debería documentarse en una ficha que identifique riesgos, controles, pruebas, limitaciones, propietario y versión.

Por el contrario, la auditoría de algoritmos de machine learning debe cubrir obtención de datos, preparación, entrenamiento, validación, aprobación, implementación y monitoreo. Un modelo puede mostrar buena exactitud general y, simultáneamente, perjudicar a determinados grupos o fallar en los casos de mayor impacto. Por eso se debe evaluar la calidad y representatividad de los datos, la legitimidad de las variables, la fuga de información y el uso de proxies de características sensibles. Las métricas deben responder al objetivo y costo del error: precisión, recall, F1, AUC, falsos positivos y negativos, además de indicadores de equidad. La auditoría también debe verificar reproducibilidad, versionamiento, explicabilidad, aprobación independiente y monitoreo de drift. Cuando cambian el negocio o la población, un modelo válido puede degradarse silenciosamente. Deben existir umbrales de alerta, recalibración y retiro, junto con revisión humana para decisiones de alto impacto.



Finalmente, el desarrollo e implementación de agentes de IA introducen un riesgo adicional: planifican, utilizan herramientas, consultan sistemas y ejecutan acciones. Un error puede propagarse rápidamente entre correos, bases de datos, plataformas financieras o infraestructuras críticas.

La auditoría debe reconstruir la cadena completa: objetivo, razonamiento operativo, herramientas, credenciales, acciones y resultados. Debe aplicarse mínimo privilegio, separar ambientes, limitar operaciones, exigir aprobación humana en acciones críticas e incorporar detención inmediata. Los registros deben explicar quién autorizó al agente, qué información utilizó y por qué actuó. Las pruebas adversariales deben considerar desviación del objetivo, instrucciones maliciosas externas, uso indebido de herramientas, coordinación inesperada y persistencia no autorizada. También se requieren límites de tiempo, costo, acciones y sistemas accesibles.

La tesis de Amodei conduce a una conclusión empresarial clara: la velocidad de implementación debe estar condicionada por la capacidad de control. A mayor autonomía, impacto y acceso a información o sistemas, mayor debe ser la profundidad de evaluación independiente.

Una política de IA es insuficiente si no se traduce en inventarios, responsables, evaluaciones de impacto, pruebas, monitoreo y evidencia. Auditar no significa frenar la innovación, sino impedir que avance a ciegas. La organización madura no es la que implementa más IA, sino aquella que puede demostrar qué utiliza, para qué, bajo qué límites y quién responde cuando algo falla. En la nueva frontera de la inteligencia artificial, gobernar el ritmo será tan importante como dominar la tecnología. La auditoría constituye el mecanismo que permite convertir prudencia en evidencia y confianza en una propiedad verificable del sistema.

Equipo Prensa
Portal Innova



f

🐦

G+

P

🗨️

📧 Post

Artículo anterior

Summit para la Salud del Futuro | IA busca anticiparse a los riesgos de salud: expertos abordarán cómo Chile puede avanzar hacia una medicina más predictiva y preventiva

Artículo siguiente

Magical fortalece el ecosistema de innovación de Coquimbo con nueva generación de mentores

Portal Innova | Grupo Prensa Digital | I.V

ARTÍCULO RELACIONADOS

MÁS DEL AUTOR

- Atención pesca artesanal de Ñuble: Amplían plazo para postular a fondo de reactivación económica

Economía y Negocios
-
- Cuenta de Ahorro Habitacional abre debate sobre las nuevas barreras para acceder a la vivienda

Economía y Negocios
-
- Estímulo al crédito hipotecario: ¿más viviendas o menor estabilidad financiera?

Economía y Negocios

Aeropuerto Santiago de Chile registró un flujo de más de 2 millones en agosto

Transporte

Abren convocatoria al Premio Nacional de Innovación Alimentaria 2026

Industria Alimentaria

Show de Motores llega este domingo a Monticello con superautos, Fórmula 3, rally, drift y música en vivo

Tecnología



SUSCRÍBETE

Nombre (*)

Teléfono (*)

Correo (*)

☐ Empresario ☐ Profesional ☐ Otro

(*): Requerido

Enviar



Carros para la cosecha | Carros Tolva | Equipos agroindustriales | Equipos agrícolas | Packing | Preparación de suelo | Trituradora de podas | Carros cónicos de descarga | Carro transportador de bins | Calibradora de frutos secos | Remecedor de piola | Limpia fosas | Plataforma de cosecha y poda |

ÚLTIMAS PUBLICACIONES

Atención pesca artesanal de Ñuble: Amplían plazo para postular a fondo de reactivación económica

Economía y Negocios septiembre 21, 2026

Cuenta de Ahorro Habitacional abre debate sobre las nuevas barreras para acceder a la...

Economía y Negocios septiembre 21, 2026

Estímulo al crédito hipotecario: ¿más viviendas o menor estabilidad financiera?

Economía y Negocios septiembre 21, 2026

Aeropuerto Santiago de Chile registró un flujo de más de 2 millones en agosto

Transporte septiembre 21, 2026

Abren convocatoria al Premio Nacional de Innovación Alimentaria 2026

Industria Alimentaria septiembre 21, 2026

 BTC	\$86,312.00 (5.63%)
 ETH	\$2,770.65 (2.94%)
 EOS	\$0.093054 (4.87%)
 XRP	\$1.54 (7.55%)

LTC

\$62.29 (5.62%)

USD

Powered by CoinGecko

SÍGUENOS EN REDES SOCIALES

Seguir a @PrensaInnova

PUBLICACIONES MÁS VISTAS

- 

Columnas y Reportajes, Energía, Medio ambiente, Noticias

Reportaje | ¡Rompiendo el molde! Joven ingeniero de la USS conquista la innovación energética en el Biobío y busca transformar el futuro del pellet

Por Portal Innova — Sábado de Reportajes Fecha: Sábado, 19 de septiembre de...
- 

Economía y Negocios, Noticias

Fondo Sigma crece más de un 50% en 2026

En los primeros siete meses del año, el fondo de deuda privada orientado a...
- 

Uncategorized

¿Llegará el dólar a \$1.000? La Fed decide mañana y el peso es la divisa más débil del mundo

El peso chileno está bajo fuego. Este martes 15 de septiembre de 2026, el d...
- 

Economía y Negocios, Noticias

Cómo Usar el Rutificador Chile para Encontrar Datos Públicos

En Chile, acceder a datos públicos de forma rápida y confiable es más fácil...
- 

Columnas y Reportajes, Economía y Negocios, Medio ambiente, Noticias

Reportaje | Fiestas Patrias: ¿De qué están orgullosos los chilenos?

El Estudio de Orgullo Chileno 2026, elaborado por Marca Chile junto a Ipsos...
- 

Tecnología, Noticias

El futuro tecnológico del SII

Por Jean Boudeguer, CEO de Clay Hace unos días participé en un encuentro re...
- 

Tecnología, Noticias

La IA ya responde en más del 40% de las búsquedas de Google. Cómo una agencia AEO te ayuda en este escenario

Cuando los usuarios investigan productos o servicios, los motores de búsque...
- 

Tecnología, Noticias

Golpe a las apuestas ilegales: Subtel ordena bloquear 42 sitios online en 48 horas — Betano, Stake y Bet365 en la lista

El gobierno dio un golpe contundente a las casas de apuestas ilegales en Ch...
- 

Economía y Negocios, Noticias

El Banco Central decide hoy: dólar en zona alta, tensión geopolítica y economía débil presionan la reunión más esperada del mes

El dólar arranca esta semana con presión alcista y el mercado en máxima ten...
- 

Noticias

Chile no tiene una recesión del cobre: tiene una recesión de proyectos

Por Manuel Viera Flores Presidente de la Cámara Minera de Chile El ministro...

Algunas imágenes de archivo son proporcionadas por: dreamstime.com

Algunas imágenes de archivo son proporcionadas por: Depositphotos.com



| [Prensa Publicidad](#) | [Prensa Colaborativa](#) | [Suscríbete](#) |

GRUPO PRENSA DIGITAL

Contacto Prensa: prensa@portalinnova.cl | +569 4817 5372

Contacto Comercial: +569 4521 9061

Prensa y Difusión para la Innovación y Sustentabilidad en las principales áreas de la economía y desarrollo.

Somos un medio de prensa colaborativo, temático y digital, perteneciente a Grupo Prensa Digital www.grupoprensadigital.cl.

Damos visibilidad a temas sobre minería, medio ambiente, energías renovables, recursos hídricos, industria alimentaria, transporte, construcción, telecomunicaciones, tecnología e informática, economía, servicios y más. mediante la publicación de contenidos de calidad, con una audiencia de profesionales de todas las edades y tomadores de decisión del ámbito público y privado.

Los 5 portales de Noticias de **Grupo Prensa Digital**, incluido Portal InnoVA, publican en forma gratuita para entidades sin fines lucros, que busquen un medio de prensa donde visibilizar sus contenidos.

Dejamos invitados a periodistas, fundaciones, municipios, colegios, universidades, ONG, agrupaciones, ministerios, servicios públicos , etc... a ser parte de nuestra red de prensa colaborativa para la Innovación en Chile. También invitamos a las empresas y marcas a sumarse a nuestro selecto grupo de Auspiciadores y ser parte de la comunidad de **Portal InnoVA Noticias**.

Contáctanos: prensa@portalinnova.cl



[Portal Agro Chile](#) [Portal Red Salud](#) [Portal Educa](#) [Prensa&Eventos](#) [Paga online aquí](#)

© Grupo Prensa Digital | [Exclusión de responsabilidad](#)