





Programa de especialización

DATA DRIVEN DECISIONS 2026

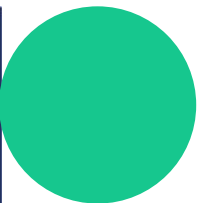
Nivel Fundamentals

SESIÓN 2



Agenda

1. Introducción a la ética en IA
2. Dilemas éticos
3. Normativa sobre Inteligencia Artificial en el Perú
4. Beneficios y riesgos de la IA
5. Estrategias para una implementación responsable de la IA



Objetivo de la sesión

Al finalizar la sesión, lograrás:

- Explicar los principios éticos relacionados con el uso de la IA
- Evaluar los posibles riesgos y beneficios de integrar la IA en la toma de decisiones
- Identificar estrategias para implementar la IA de manera responsable



Sesión 2

Introducción a la ética en IA

Principios



Transparencia



Equidad



Responsabilidad



Privacidad



Transparencia

Definición:

Claridad con la que se comunica cómo funciona la inteligencia artificial (IA), qué datos utiliza y cómo toma decisiones.



Ejemplo de violación de transparencia

The Cambridge Analytica Files
Cambridge Analytica

This article is more than 6 years old

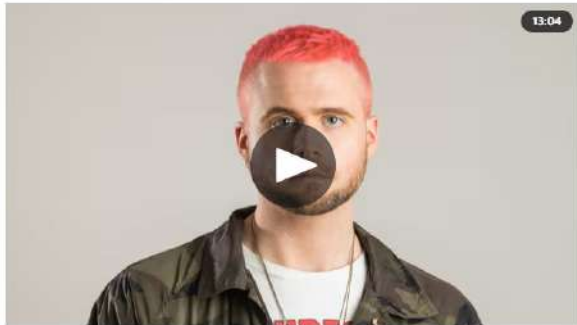
Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach

Whistleblower describes how firm linked to former Trump adviser Steve Bannon compiled user data to target American voters

- 'I made Steve Bannon's psychological warfare tool': meet the data war whistleblower
- Mark Zuckerberg breaks silence on Cambridge Analytica

Carole Cadwalladr and Emma Graham-Harrison
Sat 17 Mar 2018 22:03 GMT

Share 484



13:04

Cambridge Analytica whistleblower: 'We spent \$1m harvesting millions of Facebook profiles' - video

The data analytics firm that worked with Donald Trump's election team and the winning Brexit campaign harvested millions of Facebook profiles of US voters, in one of the tech giant's biggest ever data breaches, and used them

<https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>

Most viewed

- Live Liverpool v Aston Villa: Premier League - live
- 'Strong likelihood' of imminent famine in northern Gaza, food experts warn, as Israel continues siege
- I spent hours trying to persuade US voters to choose Harris not Trump. I know why she lost *Oliver Hall*
- Brighton v Manchester City: Premier League - as it happened
- Live Trump to visit White House next week as thousands protest in New York and Washington - live

Equidad

Definición:

Los sistemas de IA deben tratar a todos los usuarios por igual y no mostrar sesgos o discriminación hacia ningún grupo



Ejemplo de violación de equidad

Uber

This article is more than 3 years old

Ex-Uber driver takes legal action over 'racist' face-recognition software

UK driver alleges his account was illegally deactivated when Uber software decided he was not who he said he was

Robert Booth *Social affairs correspondent*
Tue 5 Oct 2021 18:00 BST

Share



Uber has used the software since April 2020 to try to prevent unlicensed and un-vetted drivers working through the app. Photograph: Russell Hart/Alamy

An **Uber** driver who lost his job when automated face-scanning software failed to recognise him is accusing the firm of indirect race discrimination in a legal test case.

The black driver, who worked on the Uber platform from 2016 until April 2021, has filed an employment tribunal claim alleging his account was illegally deactivated when facial-verification software used to log drivers on

<https://www.theguardian.com/technology/2021/oct/05/ex-uber-driver-takes-legal-action-over-racist-face-recognition-software>

Most viewed

- Live** Liverpool v Aston Villa: Premier League - live
- 'Strong likelihood' of imminent famine in northern Gaza, food experts warn, as Israel continues siege
- I spent hours trying to persuade US voters to choose Harris not Trump. I know why she lost *Oliver Hall*
- Brighton v Manchester City: Premier League - as it happened
- Live** Trump to visit White House next week as thousands protest in New York and Washington - live

Responsabilidad

Definición:

Tanto los desarrolladores como los usuarios deben asumir el control y la obligación de corregir errores, supervisar el uso de la tecnología y asegurar que su implementación sea justa y segura.



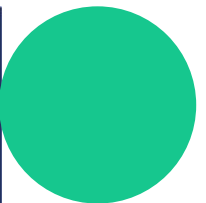
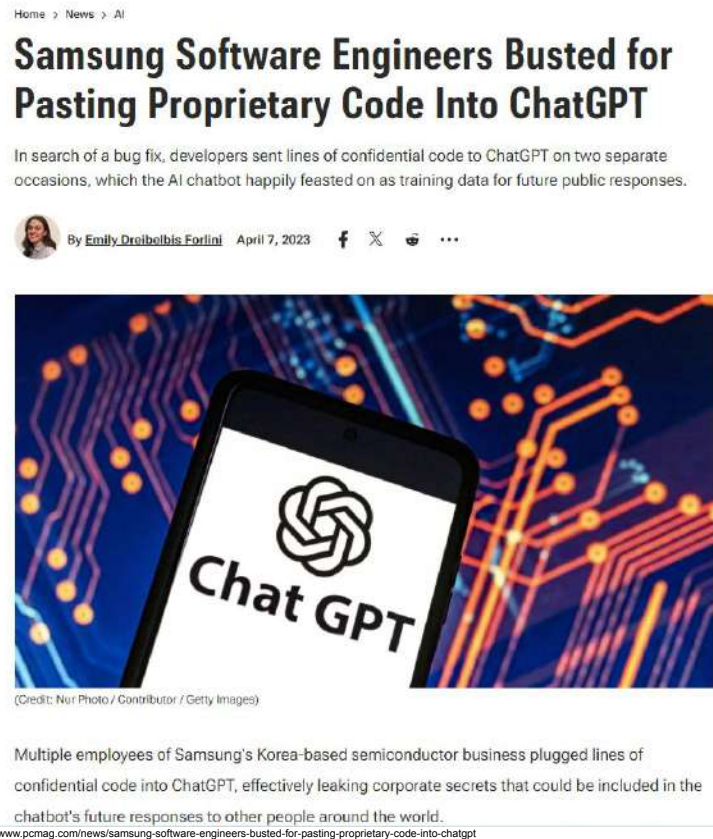
Privacidad

Definición:

La privacidad se refiere a proteger la información personal y asegurar que la IA maneje los datos de los usuarios de manera confidencial.



Ejemplo de violación de privacidad



Dilemas éticos

Ejemplo

Privacidad de datos vs. Necesidad de información

Una empresa utiliza IA para analizar datos de sus colaboradores y mejorar la productividad.

Sin embargo, surge el reto de aprovechar esta información sin vulnerar la privacidad de los empleados.

Aquí surge una pregunta clave: ¿cómo aprovechar los beneficios de la analítica y la IA sin comprometer la privacidad y la confianza de los empleados?



Ejemplo

Transparencia vs. Complejidad del algoritmo

Una empresa utiliza IA para filtrar candidatos en un proceso de selección.

Si un postulante es rechazado, debería poder entender las razones de la decisión, pero el funcionamiento del algoritmo puede ser demasiado complejo para explicarlo de manera sencilla.



Ejemplo

Sesgo algorítmico vs. Equidad en decisiones

Una empresa utiliza IA para seleccionar candidatos a una vacante.

Si el modelo fue entrenado con datos históricos sesgados, podría favorecer ciertos perfiles y perjudicar a otros, reproduciendo desigualdades en lugar de promover procesos de selección más justos.



Ejemplo

Responsabilidad de decisiones de IA vs. Supervisión humana

Una empresa utiliza IA para aprobar solicitudes, evaluar riesgos o priorizar casos de atención.

Si las decisiones se toman de forma totalmente automática, podrían producirse errores o situaciones injustas, por lo que es necesario definir cuándo y cómo debe intervenir una persona.

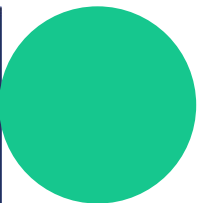


Normativa sobre Inteligencia Artificial en el Perú

Normativa de IA en el Perú: El marco legal

Reglamento de la Ley N° 31814
(Decreto Supremo N° 115-2025-PCM)

El Centro de Implementación de Políticas Públicas para la Equidad y el Crecimiento (CIPPEC)²⁰ proyectó una tasa de crecimiento anual del PBI reducida del **4.9% a 3.7%** en caso la economía peruana no adopte la IA



Normativa de IA en el Perú: El marco legal

La autoridad competente

La Presidencia del Consejo de Ministros (PCM), a través de la Secretaría de Gobierno y Transformación Digital (SGTD), actúa como la autoridad técnico-normativa a nivel nacional encargada de dirigir, evaluar y supervisar el uso y promoción de la IA.

PRINCIPIOS CLAVE EN IA Y DATOS



Normativa de IA en el Perú: El marco legal

Clasificación de riesgos de la IA

Uso Indebido (Prohibido - Art. 23)

-  Manipulación Cognitiva: Algoritmos que influyan de manera engañosa aprovechando vulnerabilidades.
-  Vigilancia Masiva: Identificación biométrica en tiempo real en espacios públicos (salvo excepciones penales o de seguridad).
-  Sesgo Biométrico: Inferir datos sensibles (origen étnico, religión, orientación sexual) mediante biometría.

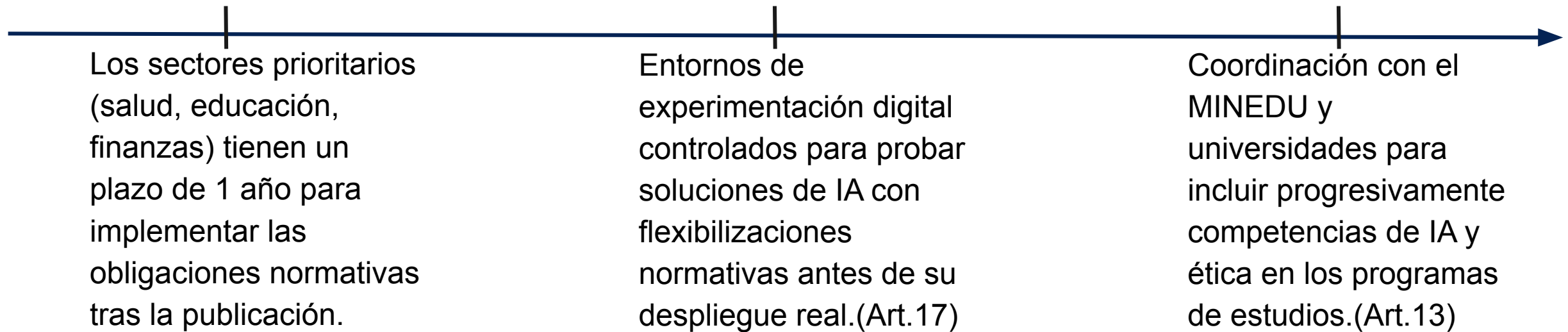
Riesgo Alto (Regulado - Art. 24)

-  Servicios Esenciales: Sistemas aplicados a la gestión de activos críticos (salud, banca, energía, transporte).
-  Ámbito Laboral: Algoritmos para selección, evaluación, contratación o cese de trabajadores.
-  Evaluación Crediticia: IA que determine el acceso a créditos (con excepción de detección de fraudes)



Normativa de IA en el Perú: El marco legal

Ecosistema e implementación gradual

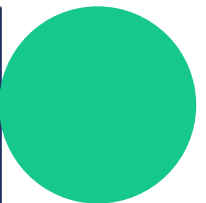
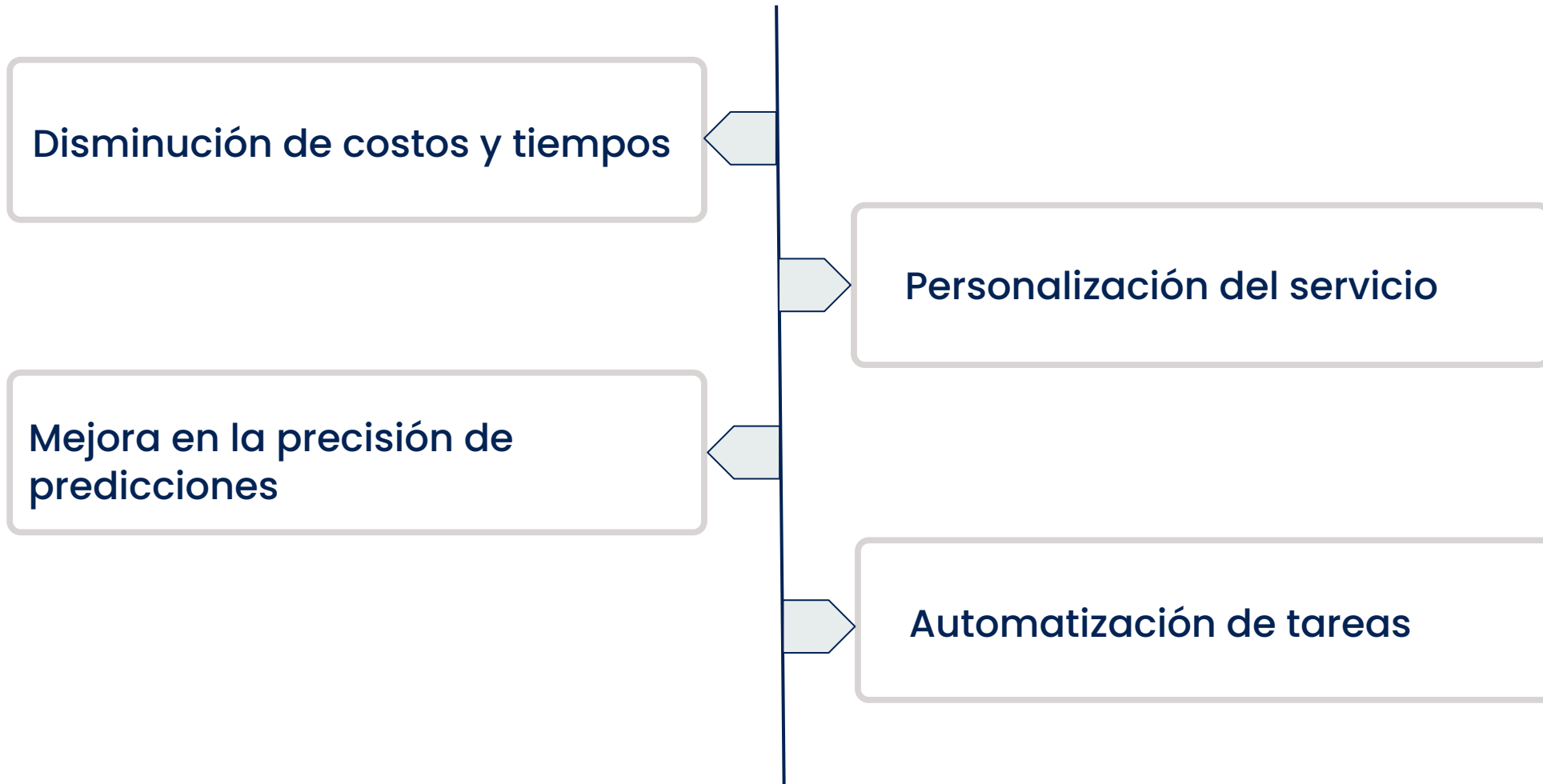


<https://cdn.www.gob.pe/uploads/document/file/8619777/7133522-decreto-supremo-n-115-2025-pcm.PDF?v=1757422328>



Beneficios y riesgos de la IA

Beneficios

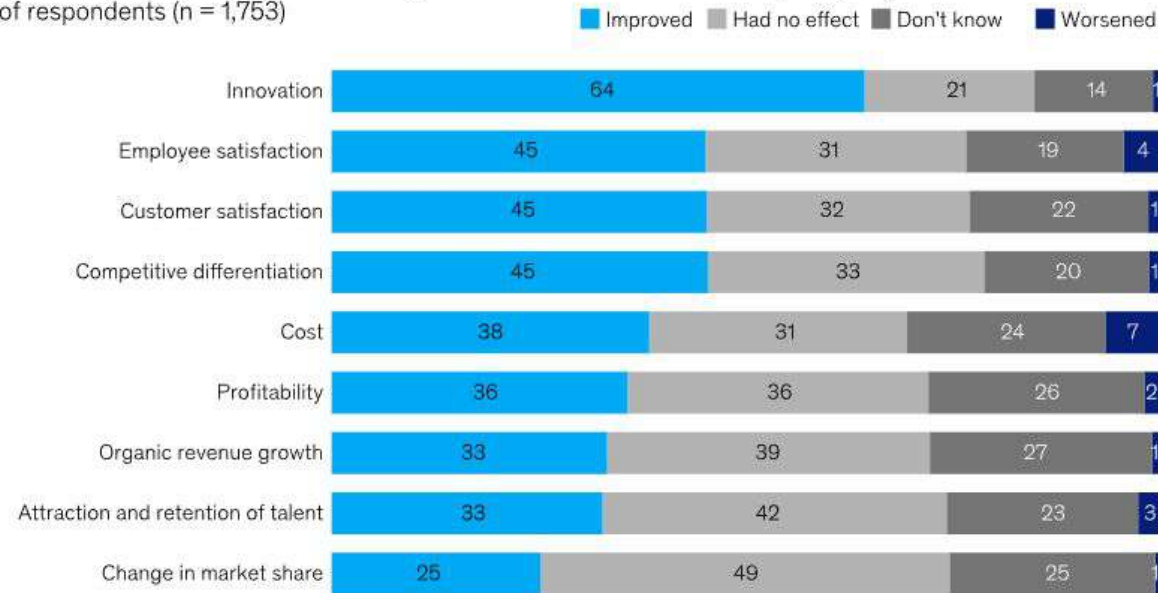


Beneficios

Respondents most often cite benefits from AI in innovation, employee and customer satisfaction, and competitive differentiation.

Extent to which AI use has affected organizational measures over the past year,¹

% of respondents (n = 1,753)



Note: Figures may not sum to 100%, because of rounding.

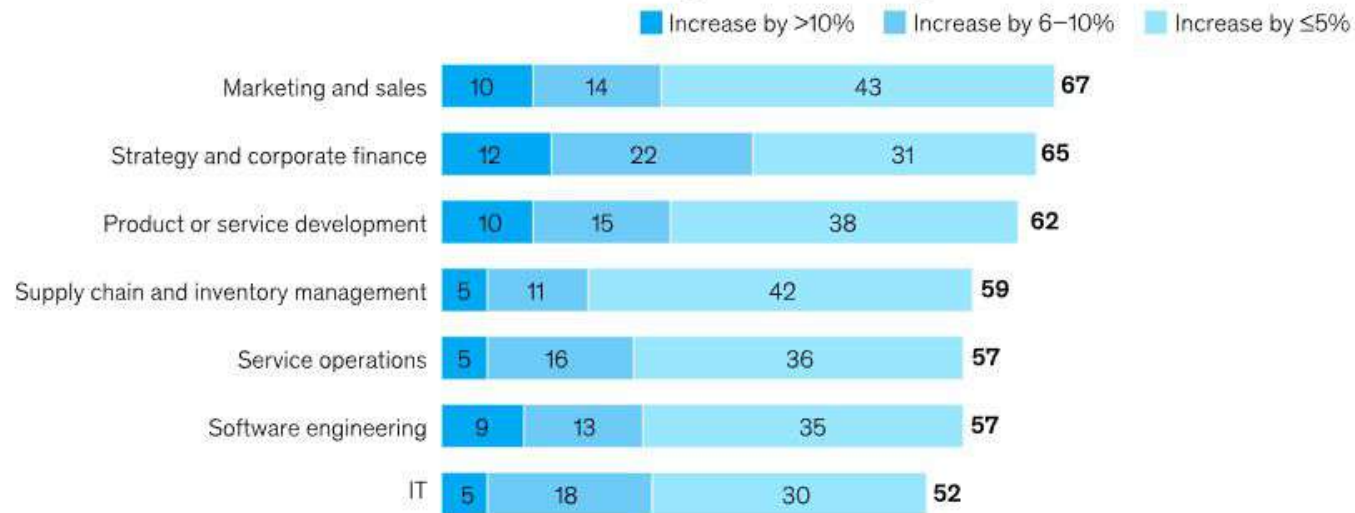
¹Asked only of respondents who said their organizations regularly use AI in at least 1 business function.

Source: McKinsey Global Survey on the state of AI, 1,993 participants at all levels of the organization, June 25–July 29, 2025

Beneficios

Respondents report the greatest revenue benefits from AI in marketing and sales, strategy and corporate finance, and product or service development.

Revenue increase within business units from AI use, past 12 months, by function,¹ % of respondents



Note: Figures may not sum to totals, because of rounding.

¹Questions were asked only of respondents who said their organizations use AI in a given function. Respondents who said "decreased revenue," "no change," "not applicable," or "don't know" for the effects of AI on revenue are not shown.

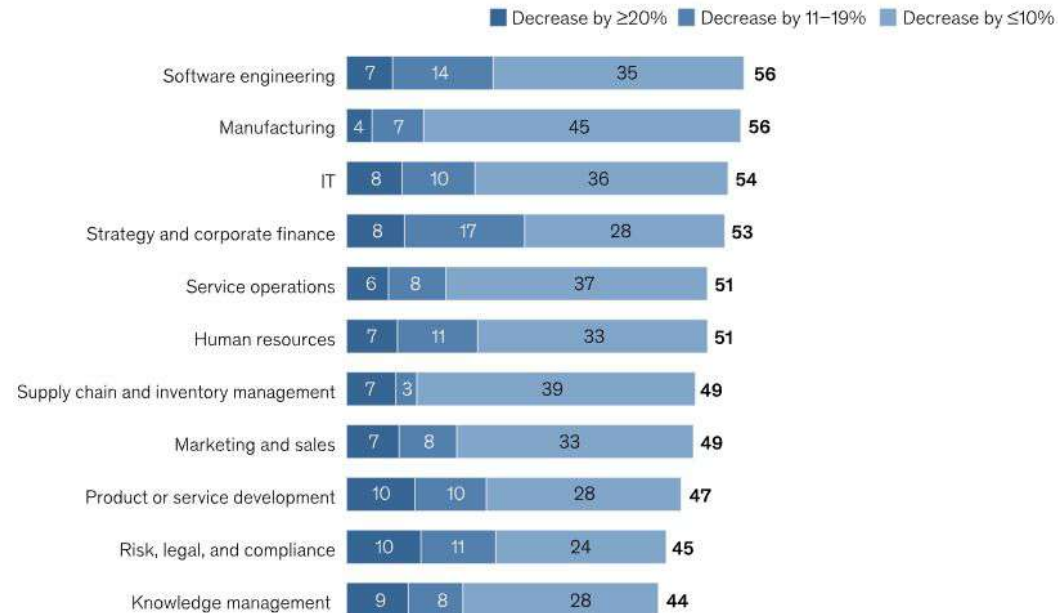
Source: McKinsey Global Survey on the state of AI, 1,993 participants at all levels of the organization, June 25–July 29, 2025



Beneficios

Respondents most commonly report cost benefits from AI activities in software engineering, manufacturing, and IT.

Cost decrease within business units from AI use, past 12 months, by function,¹ % of respondents



Note: Figures may not sum to totals, because of rounding.

¹Question was asked only of respondents who said their organizations regularly use AI in a given function. Respondents who said "cost increase," "no change," "not applicable," or "don't know" are not shown.

Source: McKinsey Global Survey on the state of AI, 1,993 participants at all levels of the organization, June 25–July 29, 2025



Riesgos y sesgos

Plagio y deshonestidad

RIESGOS

Compromiso de la privacidad de los colaboradores y recolección no autorizada de datos

Disminución de la autonomía y responsabilidad de colaboradores

Sesgo social y falta de sensibilidad cultural

Dependencia excesiva y pérdida de pensamiento crítico/empatía



Riesgos y sesgos

Plagio y deshonestidad

La IA generativa puede producir textos, informes, presentaciones, códigos o análisis completos en segundos. El riesgo surge cuando las personas presentan este contenido como propio, sin validarlo, citar las fuentes utilizadas o aportar valor intelectual adicional.

 La IA debe utilizarse como herramienta de apoyo para potenciar el trabajo humano, no como un sustituto de la responsabilidad profesional sobre los resultados entregados.

Ejemplo:

Un empleado entrega un informe generado casi en su totalidad por IA sin revisar su contenido ni reconocer su uso, comprometiendo la calidad y la integridad del trabajo.



Riesgos y sesgos

Plagio y deshonestidad

Abogado usó ChatGPT para demandar a Avianca, pero todo salió mal

Forbes Staff | junio 9, 2023 @ 4:30:19 am



Foto: Reuters.

<https://forbes.co/actualidad/abogado-uso-chatgpt-para-demandar-a-avianca-pero-todo-sallo-mal>



Riesgos y sesgos

Compromiso de la privacidad de los colaboradores y recolección no autorizada de datos

Las herramientas de IA suelen requerir grandes volúmenes de información para generar análisis o recomendaciones. El riesgo aparece cuando se comparten datos personales, confidenciales o estratégicos con plataformas de IA sin autorización, controles adecuados o conocimiento de cómo serán almacenados y utilizados.

Ejemplo:

Un empleado introduce información confidencial de clientes o colaboradores en una herramienta de IA sin autorización, exponiendo datos sensibles de la organización.

 Revisemos si están compartiendo datos con sus IA



Riesgos y sesgos

Compromiso de la privacidad de los colaboradores y recolección no autorizada de datos

Amazon Bids Employees Not to Leak Corporate Secrets to ChatGPT

It might already be too late.

By [Noor Al-Sibai](#) / Published Jan 25, 2023 1:26 PM EST / [Add Futurism](#) 



Image: Getty Images

<https://futurism.com/the-byte/amazon-bids-employees-chatgpt>



Riesgos y sesgos

Disminución de la autonomía y responsabilidad de colaboradores

La IA puede automatizar tareas, generar recomendaciones y sugerir decisiones de manera rápida. Sin embargo, una dependencia excesiva puede llevar a que los colaboradores reduzcan su capacidad de análisis crítico, cuestionamiento y toma de decisiones, delegando en la tecnología responsabilidades que deberían seguir siendo humanas.

Ejemplo:

Los colaboradores confían plenamente en las recomendaciones de la IA y dejan de analizar o validar los resultados antes de tomar decisiones.

 Cuando la IA sustituye el juicio humano en lugar de complementarlo, puede reducir la autonomía, el pensamiento crítico y la responsabilidad individual sobre las decisiones tomadas.



Riesgos y sesgos

Disminución de la autonomía y responsabilidad de colaboradores

'Think outside the bots': How to stop AI from turning your brain to mush

6 May 2026

Share



Save



 Add as preferred on Google

Thomas Germain

Fuentes:

<https://www.bbc.com/future/article/20260505-how-to-use-ai-without-turning-your-brain-to-mush>

<https://time.com/article/2026/05/19/is-ai-making-our-brains-weaker/>

<https://www.psychologytoday.com/us/blog/the-algorithmic-mind/202603/adults-lose-skills-to-ai-children-never-built-them>

<https://www.theguardian.com/australia-news/2020/may/29/robodebt-government-to-repay-470000-unlawful-centrelink-debts-worth-721m>



Riesgos y sesgos

Sesgo social y falta de sensibilidad cultural

Los sistemas de IA aprenden a partir de datos históricos que pueden contener prejuicios, estereotipos o una representación limitada de ciertos grupos. Como resultado, las respuestas o recomendaciones generadas pueden favorecer determinadas perspectivas culturales, sociales o demográficas, ignorando la diversidad de contextos presentes en una organización.

 Los sesgos presentes en los datos pueden trasladarse a las decisiones y contenidos generados por la IA, impactando la diversidad, la inclusión y el trato equitativo de las personas.

Ejemplo:

Una herramienta de IA genera recomendaciones o contenidos que reflejan estereotipos culturales o sociales, afectando la inclusión y la equidad dentro de la organización.



Riesgos y sesgos

Sesgo social y falta de sensibilidad cultural

IPROFESIONAL | TECNOLOGÍA | ALGO FALLÓ

La desastrosa idea de Microsoft: un chatbot que en 24 horas se convirtió en racista y xenófobo



El chatbot funcionaba en Twitter, donde a partir del diálogo con los usuarios, "aprendió" a elogiar a Hitler y apoyar iniciativas de Trump

<https://www.iprofesional.com/tecnologia/361581-el-chatbot-que-se-volvio-racista-el-gran-fracaso-de-microsoft>



Caso aplicado: Dilemas reales de la IA

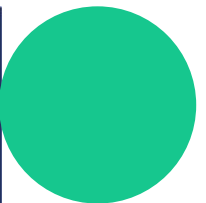
Caso aplicado

Guía de trabajo en grupo

Tiempo de Análisis: 15 minutos para leer la noticia asignada y responder las preguntas clave.

Preguntas:

- ¿Cuales son los riesgos y sesgos detectados?
- ¿Por qué identifican estos puntos como críticos?



Caso aplicado A

Tras la filtración accidental de código fuente confidencial a través de ChatGPT, Samsung limitó el uso de herramientas de IA generativa para proteger la información corporativa sensible.



Samsung prohíbe ChatGPT entre sus empleados tras una filtración de código sensible.

Por [Siladitya Ray](#), Personal de Forbes. Siladitya Ray es reportero del equipo ...

[Seguir al autor](#)

Published 2 de mayo de 2023, 07:17 a. m. EDT, Updated 2 de mayo de 2023, 07:31 a. m. EDT



Este artículo tiene más de 3 años.

LÍNEA SUPERIOR

Según informó Bloomberg, Samsung Electronics ha prohibido el uso de ChatGPT y otros chatbots con inteligencia artificial por parte de sus empleados, convirtiéndose así en la última empresa en tomar medidas enérgicas contra el uso de servicios de IA en el lugar de trabajo, en medio de la preocupación por la posible filtración de información interna confidencial en dichas plataformas.



Foto de archivo: Samsung Electronics ha prohibido el uso de ChatGPT y otras herramientas similares en el lugar de trabajo.
NURPHOTO VIA GETTY IMAGES



Caso aplicado B

Una falla en la plataforma expuso parcialmente información de otros usuarios, demostrando que incluso los sistemas de IA más avanzados pueden presentar riesgos de privacidad y protección de datos.



Estrategias para una implementación responsable de la IA

Transparencia

- Explicar el propósito de la IA y los datos que utiliza, especialmente cuando afecta a colaboradores, clientes o procesos de negocio.
- Impulsar AI Literacy.

Alfabetización en IA o alfabetización en inteligencia artificial es la capacidad de entender, utilizar, monitorear y reflexionar críticamente sobre las aplicaciones de IA



Minimización del Sesgo

- Revisar los datos utilizados para entrenar modelos de IA para asegurar que sean diversos y representativos.
- Realizar auditorías regulares para identificar y corregir sesgos en los algoritmos y modelos.
- Necesitamos equipos multidisciplinarios.

Ejemplo: Si la IA se utiliza en procesos de selección, evaluación o promoción, validar que los datos y algoritmos no favorezcan ni perjudiquen a determinados grupos de personas y realizar pruebas para confirmar la equidad del sistema.



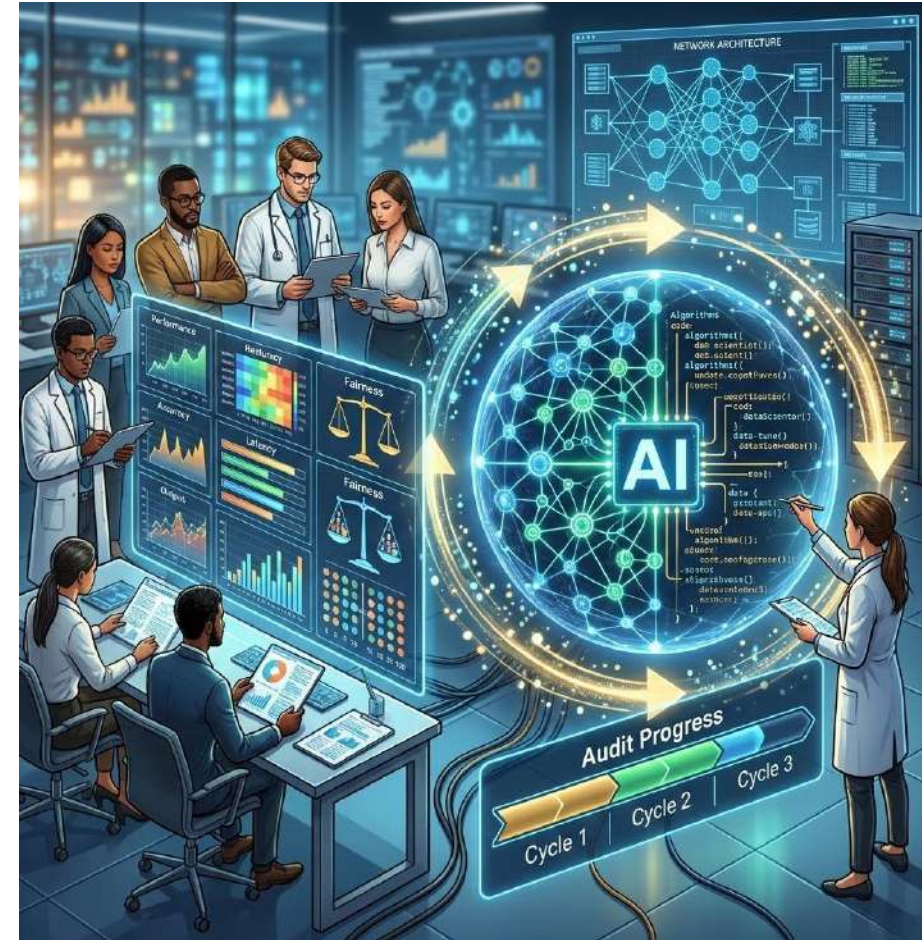
Privacidad y Seguridad de Datos

- Implementar políticas claras sobre la recolección y uso de datos.
- Utilizar herramientas de anonimización cuando sea posible.
- Utilizar sistemas de cifrado y permisos de acceso para proteger la información almacenada.



Monitoreo y mejora continua

- Establecer métricas de rendimiento y auditorías regulares para evaluar el desempeño y equidad del sistema de IA.
- Utilizar estos resultados para ajustar y actualizar los algoritmos según sea necesario.
- Esto depende de la complejidad del sistema de IA



Recursos adicionales

- IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems
<https://standards.ieee.org/industry-connections/activities/ieee-global-initiative/>
- EdSAFE AI Alliance
<https://www.edsafeai.org/>



Fin Sesión 2



La *transformación digital*
es parte de nuestro

Plan
Estratégico
Institucional

23
27



Conoce más ↗