

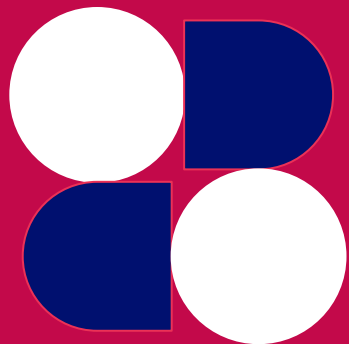


Programa PUCP DATA LITERACY

Formación tecnológica

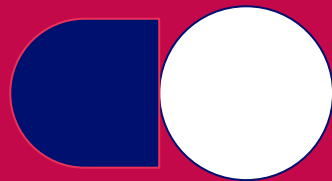
Dirección Académica de Planeamiento y Evaluación
Dirección de Gestión del Talento Humano





SESIÓN 2:

Ética, riesgos y beneficios de la IA en la toma de decisiones





Contenido:



1

Aplicaciones de la IA en el Contexto Universitario

2

Introducción a la ética en IA

3

Dilemas éticos

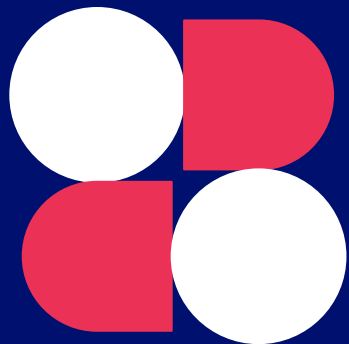
4

Beneficios y riesgos de la IA en contexto educativo

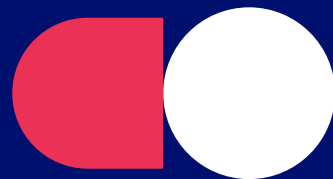


Al finalizar la sesión, lograrás:

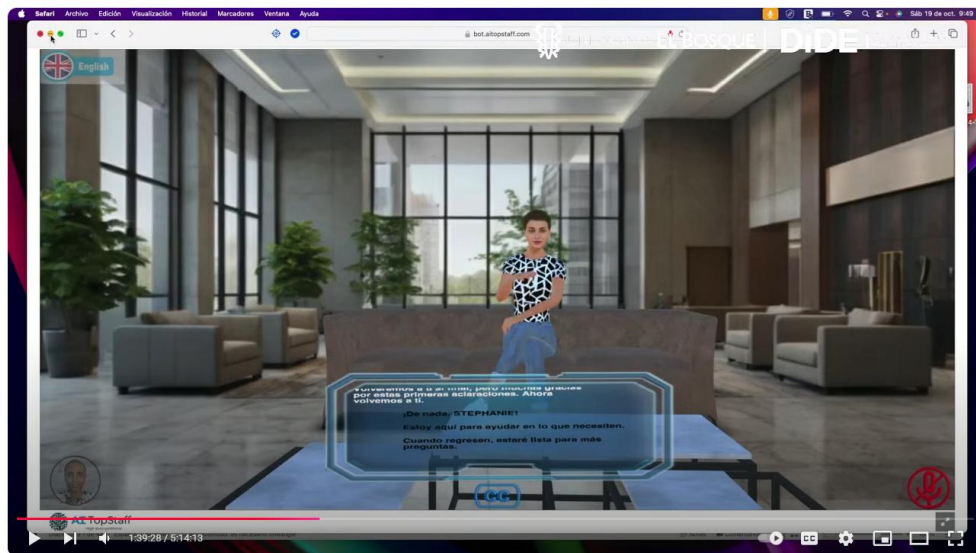
- Describir diversas aplicaciones de la IA en el contexto universitario
- Explicar los principios éticos relacionados con el uso de la IA
- Evaluar los posibles riesgos y beneficios de integrar la IA en la toma de decisiones
- Identificar estrategias para implementar la IA de manera responsable



Aplicaciones de IA en el contexto universitario



Mejorar el apoyo a estudiantes y profesores



Octubre 19 | Congreso en tendencias de educación digital

Universidad El Bosque
9.67K subscribers

Subscribe

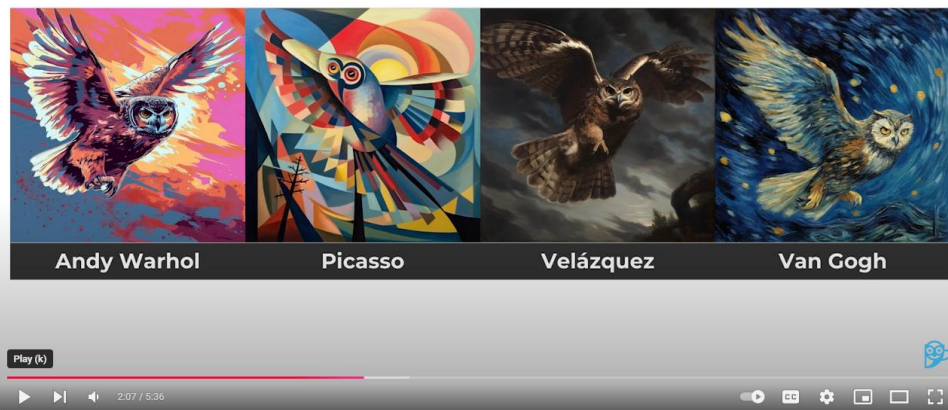
19 Share Download Clip Save ...

- Asistentes virtuales de aprendizaje para estudiantes y profesores
- Crear contenidos educativos, como resúmenes, materiales de lectura, y actividades personalizadas.

<https://www.youtube.com/watch?v=qkVXGw8uyI4&t=6574s> min 1h 40min



Generación de contenidos visuales y educativos



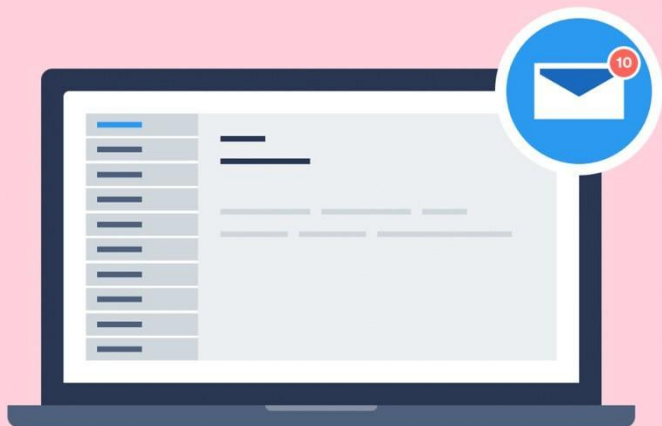
Midiournev for Teachers: using AI in the classroom

- Usar *DALL-E* o *MidJourney* para crear material visual (infografías, diagramas, imágenes para marketing o comunicación interna).
- Generación de videos explicativos sobre procedimientos o normativas universitarias.

<https://www.youtube.com/watch?v=GjPDFczYU8I>



Optimizar operaciones administrativas



- Generación automática de informes o resúmenes de reuniones.
- Automatización de correos electrónicos y respuestas a consultas administrativas.



Optimizar búsqueda de información



Ver ejemplo

- Respuestas 24/7
- Diálogo cercano al estudiante o profesor con información precisa
- Liberación del tiempo de personal administrativo
- “*ChatGPT*” para asistir en la resolución de dudas frecuentes de los estudiantes, como consultas sobre horarios, requisitos académicos, etc.



Muchas más aplicaciones



Asistente de apoyo al desarrollo y
evolución del curso

Plan de estudios personalizados
para los exámenes

Recomendación de ideas de TGs

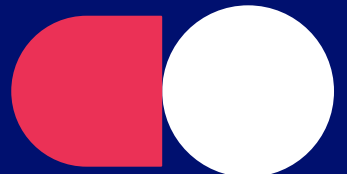
Recomendación de la mejor vía de
apoyo a estudiantes de bajo
rendimiento

Recomendación de ideas de TGs

Monitor personalizado del
profesor



Introducción a la ética en IA





Principios



Transparencia

Equidad

Responsabilidad

Privacidad



Definición

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

Entorno universitario

Los estudiantes, administrativos y otros usuarios deberían tener una comprensión básica de cómo y por qué la IA llega a ciertas conclusiones o recomendaciones.



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

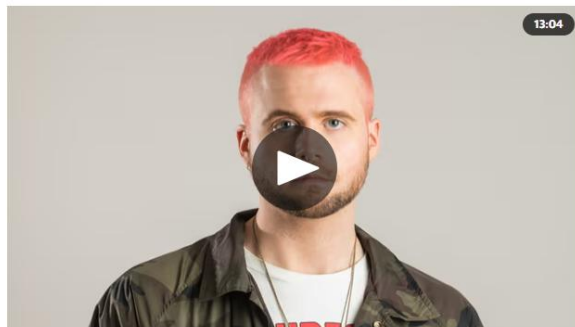
- 'I made Steve Bannon's psychological warfare tool': meet the data war whistleblower
- Mark Zuckerberg breaks silence on Cambridge Analytica

Carole Cadwalladr and Emma Graham-Harrison

Sat 17 Mar 2018 22:03 GMT

Share

484



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

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



Definición

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

Entorno universitario

Si una universidad usa IA para analizar el rendimiento académico de los estudiantes, el sistema debe asegurar que no existan sesgos basados en género, etnia o antecedentes socioeconómicos en sus análisis o recomendaciones.



Ejemplo de violación de equidad



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

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

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



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.

Entorno universitario

Si un sistema de IA en la universidad comete un error o afecta negativamente a un grupo de estudiantes, los administradores deben estar preparados para responder, resolver el problema y evitar que ocurra en el futuro.



Definición

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

Entorno universitario

Un asistente virtual de IA que responde a preguntas de los estudiantes no debería recopilar ni almacenar información privada de los usuarios sin su consentimiento explícito, y debería limitar su acceso a solo los datos necesarios para realizar sus funciones



Ejemplo de violación de privacidad



Home > News > AI

Samsung Software Engineers Busted for Pasting Proprietary Code Into ChatGPT

In search of a bug fix, developers sent lines of confidential code to ChatGPT on two separate occasions, which the AI chatbot happily feasted on as training data for future public responses.



By [Emily Dreifelbis Fortini](#)

April 7, 2023



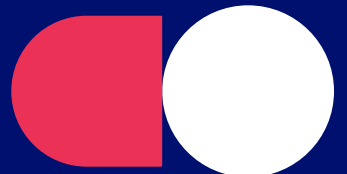
(Credit: Nur Photo / Contributor / Getty Images)

Multiple employees of Samsung's Korea-based semiconductor business plugged lines of confidential code into ChatGPT, effectively leaking corporate secrets that could be included in the chatbot's future responses to other people around the world.

<https://www.pcmag.com/news/samsung-software-engineers-busted-for-pasting-proprietary-code-into-chatgpt>



Dilemas éticos





Privacidad de datos vs. Necesidad de información

Una universidad que utiliza IA para identificar patrones de éxito académico podría obtener datos sensibles, como antecedentes personales o de salud de los estudiantes. Aquí surge la pregunta: ¿Cómo encontrar un equilibrio entre mejorar los servicios y proteger la privacidad individual?



Transparencia vs. Complejidad del algoritmo

Si una IA decide recomendaciones de admisión o becas, los afectados deberían entender por qué se tomó una decisión. Sin embargo, la explicación completa del proceso puede ser demasiado técnica.



Sesgo algorítmico vs. Equidad en decisiones

Si una universidad usa IA para identificar estudiantes en riesgo de abandonar sus estudios, el sistema podría sesgarse en función de datos históricos, como antecedentes socioeconómicos o académicos, replicando patrones discriminatorios en lugar de ayudar a resolverlos.



Responsabilidad de decisiones de IA vs. Supervisión humana

En decisiones de evaluación de solicitudes o manejo de casos de estudiantes, depender exclusivamente de la IA puede llevar a errores o decisiones injustas. El dilema es cómo mantener el rol humano en la toma de decisiones mientras se aprovecha la IA.



¿Qué podríamos hacer?



Un grupo multidisciplinario:

- ¿Cuáles son los valores o principios en conflicto?
- ¿Quiénes son los afectados en este dilema?
- ¿Qué riesgos y beneficios identifican al usar IA en esta situación?
- ¿Qué soluciones o compromisos podrían mitigar estos riesgos?

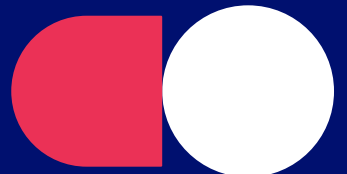


¿Cuáles son los dilemas éticos que podrían surgir al aplicar IA en tu área de trabajo?

3 participantes



Beneficios y riesgos de la IA





Beneficios



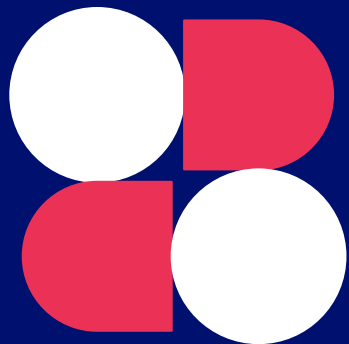
**Disminución de costos y
tiempos**

Personalización del servicio

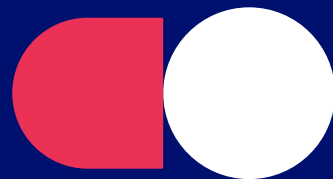
**Mejora en la precisión de
predicciones**

Automatización de tareas





Estrategias para una implementación responsable de la IA





- Explicar el propósito de la IA y los datos que utiliza, especialmente cuando afecta a estudiantes o personal.
- 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 se utiliza IA para evaluar solicitudes de ayuda financiera, asegurarse de que los datos no penalicen a ciertos grupos étnicos, sociales o económicos, 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



- UNESCO: Directrices para la Inteligencia Artificial en la Educación
<https://www.unesco.org/es/digital-education/artificial-intelligence>
- 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/>



Gracias

