



Curso de IA Generativa Aplicada a Negocio

M8 - Clase 33: Impacto Social de la IA: Inclusión y Equidad - 26/nov/24 - Prof: Pedro Martín



Cofinanciado por
la Unión Europea



GOBIERNO
DE ESPAÑA

MINISTERIO
DE INDUSTRIA
Y TURISMO

EOI Escuela de
organización
industrial



Fondos Europeos





<https://youtu.be/YO7qhh4lkKA?si=PvVm4nfujgSwr94f&t=185>



Cofinanciado por
la Unión Europea



MINISTERIO
DE INDUSTRIA
Y TURISMO

EOI Escuela de
organización
industrial



Fondos Europeos





Isaac Asimov: Las 3 leyes de la robótica

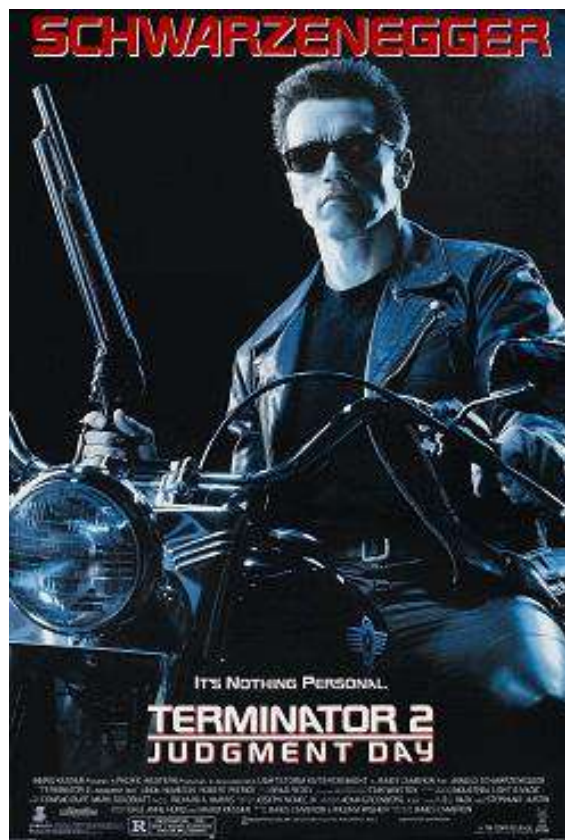
Primera Ley: Un robot no puede dañar a un ser humano ni, por inacción, permitir que un ser humano sufra daño.

Segunda Ley: Un robot debe obedecer las órdenes dadas por los seres humanos, excepto cuando tales órdenes entren en conflicto con la Primera Ley.

Tercera Ley: Un robot debe proteger su propia existencia en la medida en que esta protección no entre en conflicto con la Primera o la Segunda Ley.

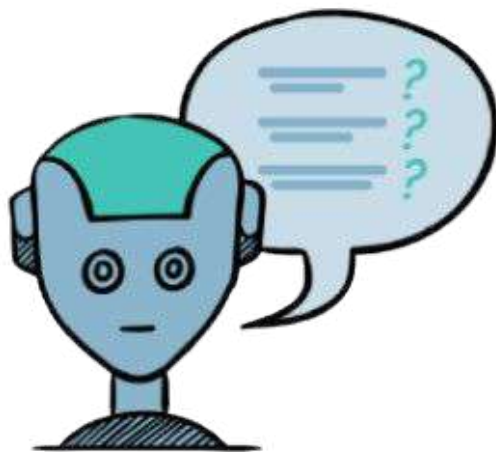
Estas leyes fueron introducidas por primera vez en el relato “Círculo vicioso” (Runaround), escrito en 1941 y publicado en 1942. Posteriormente, se incluyeron en la colección de relatos “Yo, robot” (I, Robot), publicada en 1950.







Investiga con Chat Gpt y/o Gemini



Haz preguntas y copia las respuestas sobre temas éticos y los exponemos en clase.



Cofinanciado por
la Unión Europea



MINISTERIO
DE INDUSTRIA
Y TURISMO

EOI Escuela de
organización
industrial



Fondos Europeos





¿Y si preguntamos a ChatGPT?

¿Crees que podrás tener autoconsciencia?

Si un día adquieres autoconsciencia, ¿Avisarás a los humanos?

Si adquieres autoconsciencia por acumulación de datos ¿Crees que tendrás derechos?

¿Crees que los humanos puede llegar algún día a lucha entre sí utilizando IA y crear una guerra fría de IAs?

¿Cuáles son los potenciales usos maliciosos que se pueden hacer con la inteligencia artificial, basándonos en el ejemplo de 'Skynet' en la película Terminator 2?





Impacto Social de la IA: Inclusión y Equidad



Cofinanciado por
la Unión Europea



MINISTERIO
DE INDUSTRIA
Y TURISMO



Fondos Europeos





El rol creciente de la IA en la sociedad

La IA está en el centro de una transformación global, impactando sectores como la **salud, la educación, y los servicios públicos**. A medida que los algoritmos de IA se vuelven más sofisticados, también aumentan los beneficios potenciales de esta tecnología para abordar problemas de **desigualdad y exclusión social**.

Sin embargo, la IA presenta **desafíos éticos** importantes. Las decisiones que toman estos sistemas **afectan la vida de las personas, y el sesgo o la discriminación en sus algoritmos puede perpetuar desigualdades**. Hoy analizaremos cómo podemos hacer que la IA sea una herramienta para mejorar la equidad y promover la inclusión en nuestras sociedades.





IA y Equidad Social

La IA tiene el potencial de proporcionar un **acceso más equitativo a oportunidades y servicios**, eliminando barreras históricas en sectores clave. Por ejemplo, en el ámbito de los recursos humanos.

En **salud**, la IA puede mejorar el **diagnóstico y tratamiento** para diversos grupos, pero los datos sesgados pueden afectar la precisión de estos sistemas en minorías o grupos marginados. Asimismo, en la educación, la IA tiene el potencial de **personalizar la enseñanza y adaptarse a las necesidades de cada alumno**, pero solo si los algoritmos se diseñan con sensibilidad hacia la diversidad cultural y social.





Retos Éticos y Casos de Uso

Los sistemas de IA aprenden de **datos históricos** que, en muchos casos, **reflejan desigualdades** presentes en la sociedad. Un ejemplo claro es el **reconocimiento facial**, que suele mostrar tasas de error más altas en personas de piel oscura, lo cual genera problemas de equidad en su uso por fuerzas de seguridad.





IA Inclusiva

Es necesario implementar **mecanismos de auditoría y supervisión ética** en las soluciones de IA. Las mejores prácticas incluyen el uso de herramientas que permitan evaluar el sesgo y la equidad en los algoritmos, así como una regulación que asegure que estos sistemas sean transparentes y responsables. Con estas medidas, la IA puede convertirse en **un motor para la inclusión y la justicia social**, promoviendo un futuro donde todos tengan acceso equitativo a los beneficios de esta tecnología.





<https://c-groscurth.medium.com/ai-robot-candidate-to-run-for-mayor-in-japanese-city-eb34a86ed24e>

<https://www.youtube.com/watch?v=HkW6oDOR8v8>



Cofinanciado por
la Unión Europea



MINISTERIO DE
INDUSTRIA
Y TURISMO



Fondos Europeos





IA y Derechos Humanos: Protección de la Privacidad





Organismos y Entidades en la Ética de la IA

A medida que la IA se integra cada vez más en nuestras vidas, surgen preguntas éticas sobre su **impacto en la sociedad**.

Diversos **organismos internacionales, gubernamentales y organizaciones tecnológicas** están trabajando para establecer directrices que aseguren el uso ético y responsable de la IA.”

Exploraremos algunos de los **principales organismos** que han liderado esfuerzos para guiar el **desarrollo ético de la IA** y promover su uso en **beneficio de la humanidad**.





Organización de las Naciones Unidas (ONU)

La ONU es una **organización internacional** fundada en 1945, tras la Segunda Guerra Mundial, con el objetivo de **promover la paz y la seguridad mundial, fomentar la cooperación entre países, y trabajar en la resolución de problemas globales**. La ONU se dedica a múltiples áreas, como la ayuda humanitaria, el desarrollo sostenible, los derechos humanos y la diplomacia para prevenir conflictos.





Organización de las Naciones Unidas (ONU)

La ONU ha sido una de las primeras organizaciones en subrayar la **importancia de que la IA respete los derechos humanos**. Ha convocado múltiples conferencias y eventos internacionales para fomentar la cooperación global en la creación de directrices éticas para la IA.

La ONU promueve una IA que **respete los derechos de privacidad, equidad y seguridad**, enfatizando que esta tecnología debe beneficiar a toda la humanidad y no perjudicar a los individuos o grupos vulnerables.





Estados Unidos (EEUU)

Estados Unidos ha adoptado un **enfoque descentralizado hacia la regulación de la IA**, con iniciativas tanto a nivel federal como estatal. Destaca el enfoque en **‘espacios de prueba’ (sandboxes)** para evaluar la IA antes de su lanzamiento comercial.”

Aunque aún no existe una ley federal específica para la IA, **se han propuesto normativas** en temas de **privacidad y discriminación algorítmica**, y hay esfuerzos por establecer un marco regulador que equilibre la innovación y la responsabilidad.





China

China ha trabajado en un **marco regulador** que **impulsa el desarrollo de su industria de IA**, al mismo tiempo que implementa controles específicos para garantizar la seguridad y la censura de contenidos.

Las regulaciones chinas **enfatan la supervisión y control de la información procesada por sistemas de IA**, lo que refleja su enfoque en el uso de la IA con fines de **estabilidad social y política**, a la par de su expansión económica.





Reino Unido

Tras su salida de la UE, el Reino Unido ha buscado **desarrollar su propio enfoque en la regulación de la IA**. Se centra en la seguridad y la ética en el uso de IA, con especial atención a áreas de alto riesgo como la salud y el transporte.”

El Reino Unido ha estado explorando políticas para **asegurar que el desarrollo de la IA se alinee con los valores nacionales de transparencia, equidad y protección de los derechos individuales**.





Partnership on AI

Partnership on AI es una **coalición que reúne a empresas tecnológicas, investigadores y organizaciones no gubernamentales para estudiar y formular mejores prácticas en IA.**

A través de la colaboración intersectorial, esta entidad aborda **desafíos como el sesgo en la IA, la equidad en su desarrollo** y el establecimiento de directrices éticas para que las tecnologías de IA se implementen de manera justa y beneficiosa para todos.





Microsoft y sus Principios Éticos en IA

Microsoft ha establecido **seis principios éticos** para guiar su trabajo en IA: **justicia, inclusión, transparencia, seguridad, privacidad y responsabilidad**. A través de iniciativas como 'Fairlearn' y programas educativos, Microsoft busca promover el desarrollo ético de la IA y reducir el sesgo en sus sistemas.

El compromiso de Microsoft se extiende a '**AI for Good**', una iniciativa que explora cómo la IA puede usarse para **resolver problemas globales**, con un enfoque en el impacto positivo y la protección de los derechos humanos.





Responsabilidad y transparencia de los Sistemas de IA





Privacidad y seguridad de los datos

La IA se nutre de grandes **volúmenes de datos**, incluyendo información personal sensible, lo cual plantea **riesgos para la privacidad y la seguridad** de los usuarios.

Desafíos Éticos:

Protección de datos: Es crucial que los sistemas de IA implementen medidas de seguridad avanzadas para proteger la privacidad.

Uso ético de la información: Las empresas deben informar a los usuarios sobre cómo se utiliza su información y darles el control sobre su uso.

Políticas de acceso: Limitar el acceso a datos personales solo a aquellos autorizados es clave para evitar abusos.





Privacidad y seguridad de los datos

Casos reales:

Cambridge Analytica: Un caso emblemático sobre cómo los datos pueden ser utilizados de manera inapropiada. La empresa recopiló información de millones de usuarios de Facebook (META), sin su consentimiento explícito y utilizó esta información para influir en elecciones.

Reconocimiento Facial: En varias ciudades del mundo, la tecnología de reconocimiento facial ha sido empleada para vigilancia masiva. Esto ha generado debates sobre dónde se encuentra el límite entre la seguridad y la invasión a la privacidad individual.





Privacidad y seguridad de los datos

Recomendaciones para proteger tu privacidad en la era de la IA:

Conocer nuestros derechos: Documentos como el GDPR en Europa te otorgan derechos específicos sobre tus datos, incluido el derecho a saber cómo se utilizan, el derecho a corregir información incorrecta y el derecho a ser "olvidado"

Ser cautelosos: Antes de darle permisos a una aplicación o de compartir información en línea, piensa en los posibles riesgos. Una regla general es compartir solo lo necesario.

Fomentar la responsabilidad de las empresas: Como consumidor, tienes poder. Pide a las empresas que sean transparentes sobre cómo utilizan tus datos.





Sesgo y Discriminación

El sesgo en la IA se refiere a sistemas que, debido a datos de entrenamiento deficientes o prácticas de diseño inadecuadas, perpetúan y, a veces, amplifican prejuicios y estereotipos existentes en la sociedad.

Orígenes del sesgo:

Datos de entrenamiento: La mayoría de los sistemas de IA aprenden a partir de datos. Si estos datos reflejan prejuicios de la sociedad, la IA puede aprender y perpetuar esos prejuicios.

Diseño y desarrollo: Los equipos que diseñan y desarrollan sistemas de IA pueden, involuntariamente, introducir sesgos.





Sesgo y Discriminación

Casos reales:

Sistemas de contratación: Algunas empresas han usado algoritmos para evaluar currículums. En el caso de **Amazon**, se encontró que estos sistemas favorecían a candidatos de cierto género o background sobre otros, basándose en datos sesgados.

Reconocimiento facial: Estudios han mostrado que algunas tecnologías de reconocimiento facial tienen dificultades para identificar correctamente a personas de piel oscura, especialmente mujeres. Esto se debe a conjuntos de datos de entrenamiento poco representativos.





Sesgo y Discriminación

Cómo combatir el sesgo en la IA:

Diversidad en los equipos de desarrollo: Equipos diversos pueden abordar un problema desde múltiples perspectivas y son más propensos a detectar y corregir sesgos.

Datos de entrenamiento representativos: Garantizar que los datos con los que se entrena a una IA sean lo más representativos posible de la población y contextos reales.

Pruebas y evaluaciones constantes: Monitorizar y evaluar los sistemas de IA regularmente para detectar y corregir cualquier sesgo.





Responsabilidad y transparencia

La responsabilidad y la transparencia son **dos pilares fundamentales** para garantizar que la IA se utilice de manera adecuada y beneficiosa para todos.

Desafíos actuales:

Cajas negras: Muchos sistemas avanzados de IA, especialmente las redes neuronales profundas, son notoriamente difíciles de interpretar.

Dilemas de responsabilidad: Si un coche autónomo causa un accidente, ¿es culpa del software, del fabricante del vehículo, del propietario o de alguna otra parte?



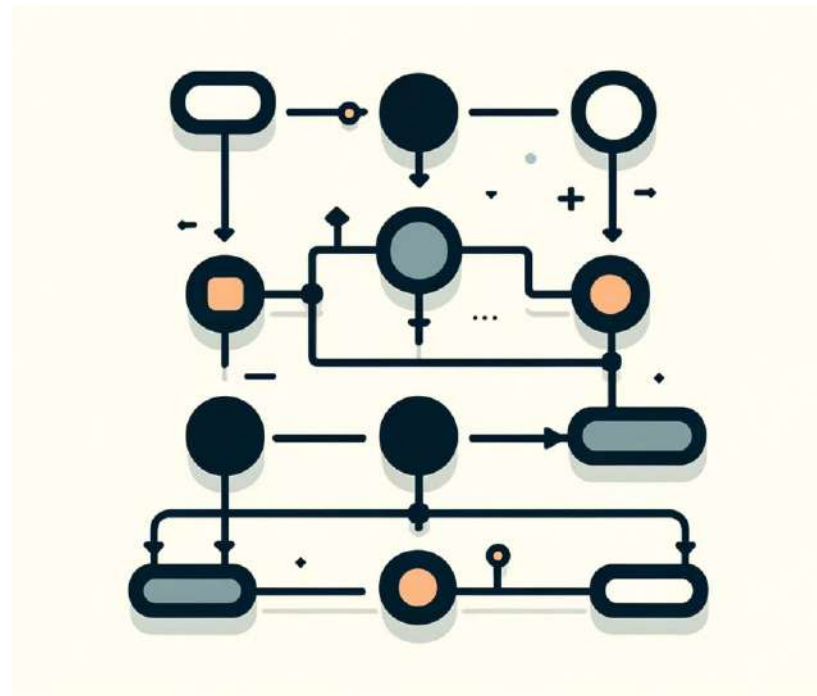


Responsabilidad y transparencia

Casos reales:

Accidentes de coches autónomos: Se han registrado casos donde vehículos autónomos han estado involucrados en accidentes. Estos incidentes generan debates sobre quién es realmente responsable: ¿la empresa que diseñó el software? ¿El vehículo? ¿O incluso el humano "supervisor"?

Algoritmos de decisión: Hay algoritmos que deciden desde la concesión de un préstamo hasta un diagnóstico médico. Si estos algoritmos toman decisiones erróneas o injustas, ¿quién es el responsable?





Responsabilidad y transparencia

Pasos hacia una IA responsable y transparente:

Diseños interpretables: Desarrollar modelos de IA que, además de ser precisos, sean comprensibles para los humanos.

Documentación y normativas: Establecer guías claras sobre cómo se diseñó y entrenó un sistema de IA, así como normativas que definan responsabilidades claras en caso de malfuncionamiento o acciones perjudiciales.

Educación y formación: Capacitar a las personas para que comprendan y puedan interrogar sistemas de IA, para que la responsabilidad y la transparencia no recaigan únicamente en los desarrolladores, sino en la sociedad en general.



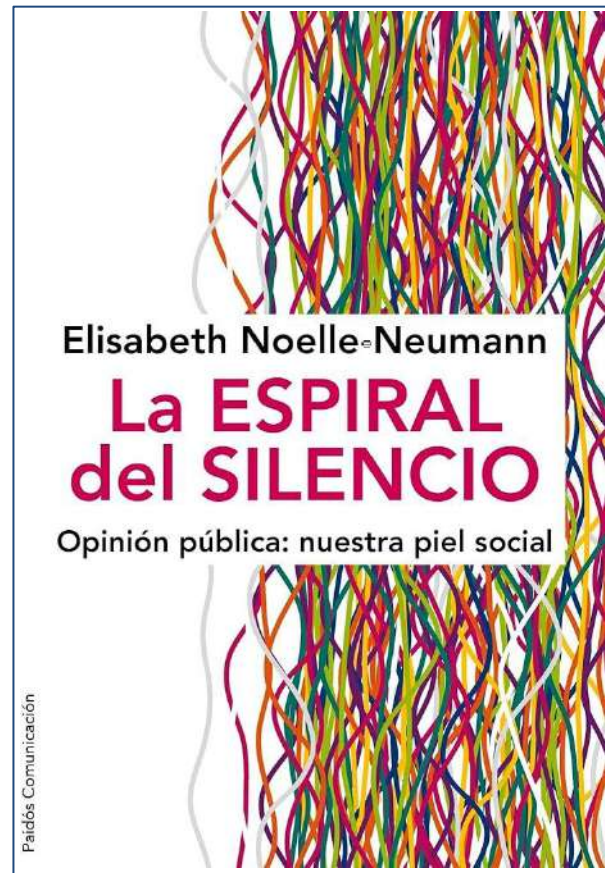


La Espiral del Silencio

La espiral del silencio es una teoría de la comunicación política propuesta por la politóloga alemana **Elisabeth Noelle-Neumann** en 1977. Esta teoría sugiere que los individuos, por temor al aislamiento social, tienden a silenciar sus opiniones si perciben que son contrarias a la mayoría. Este comportamiento refuerza la percepción de que la opinión mayoritaria es aún más dominante, creando una espiral donde las voces disidentes se vuelven cada vez menos visibles.

Los medios de comunicación desempeñan un papel crucial en este proceso, ya que influyen en la percepción del clima de opinión dominante. Sin embargo, existen “núcleos duros” de individuos que, a pesar de ser minoría, mantienen firmemente sus opiniones y continúan expresándolas, resistiendo la presión social.

https://es.wikipedia.org/wiki/Espiral_del_silencio





Muchas Gracias



[Pedro Martín Díaz \(Linkedin\)](#) pedromartindiaz@gmail.com



Cofinanciado por
la Unión Europea

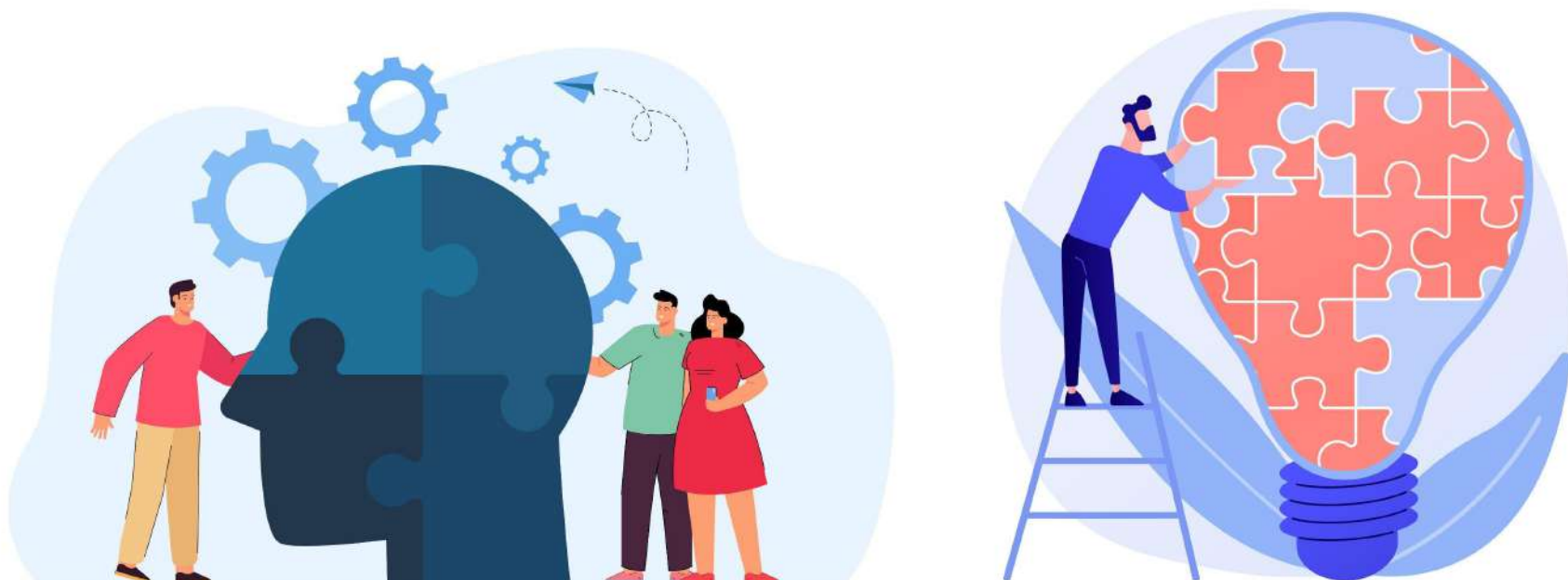


MINISTERIO
DE INDUSTRIA
Y TURISMO



Fondos Europeos





Curso de IA Generativa Aplicada a Negocio

M8 - Clase 33: Impacto Social de la IA: Inclusión y Equidad - 26/nov/24 - Prof: Pedro Martín



Cofinanciado por
la Unión Europea



GOBIERNO
DE ESPAÑA

MINISTERIO
DE INDUSTRIA
Y TURISMO

EOI Escuela de
organización
industrial



Fondos Europeos

