



Universidad  
**Tecnológico**®



# Ética aplicada a la inteligencia artificial

Ética aplicada a la  
inteligencia artificial



La mezcla de grandes avances científicos con una tecnología moderna aplicada en un contexto de una economía capitalista global ha generado severos problemas ambientales.

En este tema conocerás:

- El principio de responsabilidad de Hans Jonas.
- El principio de precaución.
- Los principios de la bioética.





Es de vital importancia establecer **principios básicos** para el profesional que trabaja con algoritmos de aprendizaje automático, por ejemplo:

a. Estar preparados y ser lo suficientemente honestos para saber cuándo las herramientas y metodologías de aprendizaje automático o inteligencia artificial no pueden por sí solas dar solución a los problemas que enfrentan.

b. Reconocer cuando sean necesarios otros mecanismos, como acciones sociales o políticas, para dar solución a los problemas como el del sesgo, por citar uno.

c. Comprender que no trabajan solos ni existen independientemente de una comunidad social mayor regida por sus propias reglas políticas, tradiciones, usos y costumbres de donde varias prácticas sociales derivan su significado.

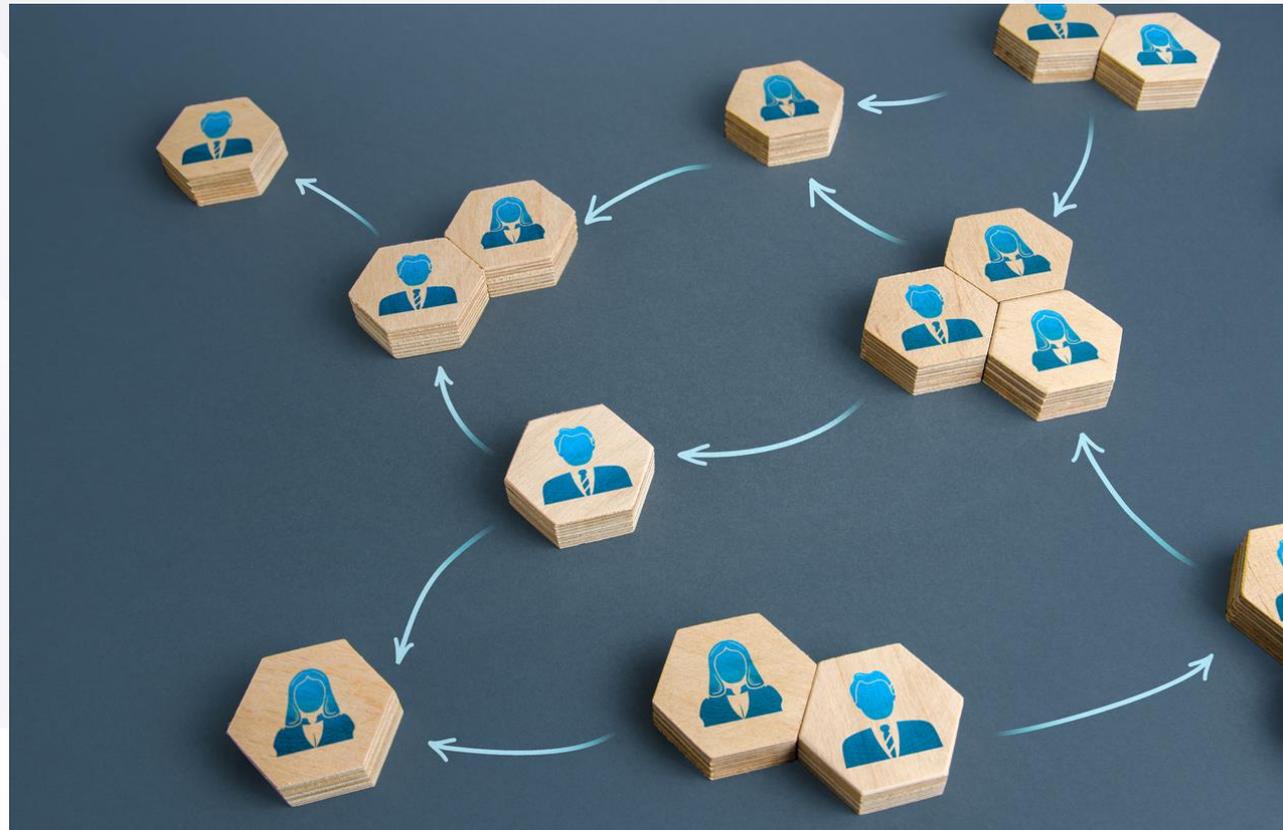
d. Entender que no son instancias de datos distribuidos de forma idéntica e independiente, sino que representan nodos conscientes situados en una red compleja e interconectada de creencias, prácticas legales, sociales, culturales, económicas y morales, donde ellos mismos se pueden ver afectados por los propios algoritmos.





No se debe olvidar que quienes trabajan en esta área son los agentes morales responsables de la toma de decisiones y el resto del mundo son usuarios morales. ¿Qué derechos deben tener los usuarios morales?, ¿qué obligaciones deben tener los agentes morales hacia sus usuarios?

Para resolver ambos cuestionamientos, se debe tener claro el significado de responsabilidad y conocer la diferencia entre **responsabilidad causal** y **responsabilidad moral**.



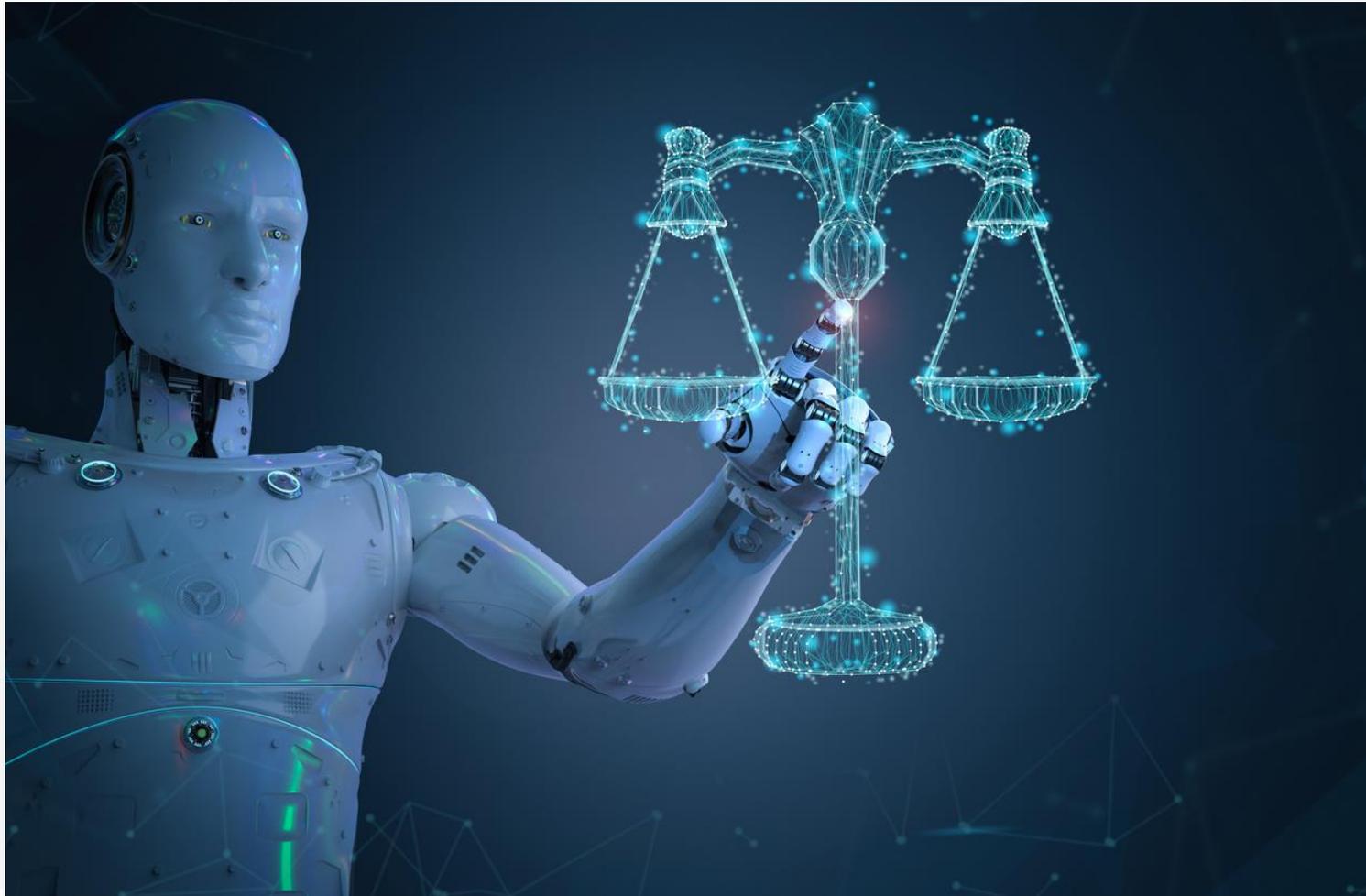


En 1979, Hans Jonas escribió *El principio de responsabilidad*, donde afirma que todos los sistemas éticos desarrollados se deben actualizar, ya que toman de forma errónea el alcance de las acciones humanas y, por lo tanto, la responsabilidad humana puede definirse de forma estricta. Jonas argumenta que los enfoques tradicionales de la ética están basados en consideraciones utilitarias o deontológicas que se desarrollaron bajo condiciones sociales completamente distintas. Durante la década de los 80, Jonas desarrolló aún más las implicaciones éticas de su “principio de responsabilidad” que se convirtió después en el “principio de precaución”. La idea detrás de este principio es, básicamente, “más vale prevenir, que lamentar”.





El diseño de sistemas de decisión basados en aprendizaje automático no es un reto meramente técnico, se requiere además poner atención a los principios bioéticos: **beneficencia, no maleficencia, autonomía y justicia.**





¿Es fácil distinguir entre responsabilidad moral y responsabilidad causal?

¿Puede existir una sin la otra?

Investiga un ejemplo en que pueda existir lo anterior

¿Cómo se aplicaría el principio de precaución para este ejemplo?





El desarrollo y uso de algoritmos de IA tiene el potencial de generar un impacto tanto positivo como negativo en la sociedad: minimizar o ampliar las desigualdades existentes, resolver problemas viejos o provocar nuevos. Para trazar una ruta que sea socialmente preferida no solo se deben utilizar buenas regulaciones y estándares, sino también crear un marco de principios éticos dentro del cual se definan acciones concretas que permitan pasar de los buenos principios a las buenas prácticas.



# Ética aplicada a la inteligencia artificial

Dilemas morales

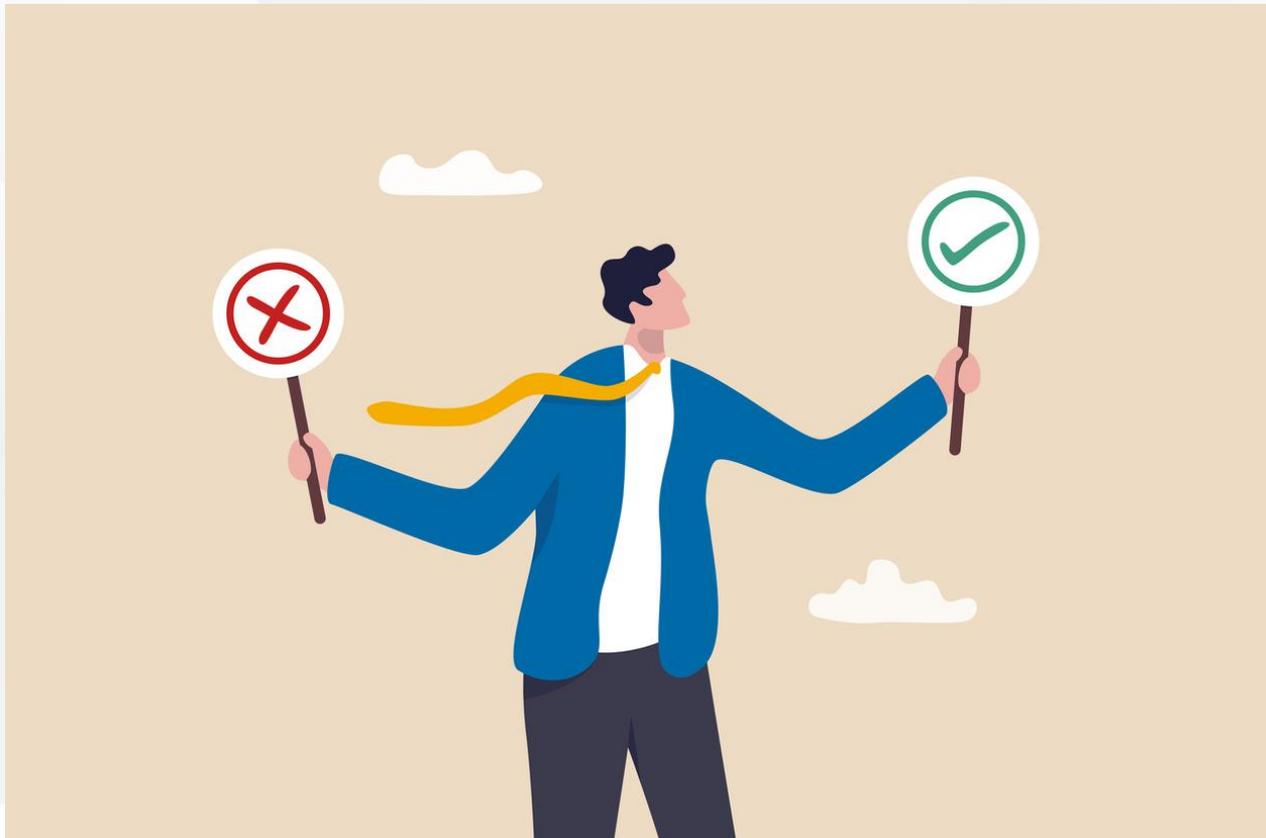


Si bien es cierto que el uso de algoritmos de aprendizaje automático no es nuevo, las cuestiones éticas y los dilemas que rodean a la inteligencia artificial se encuentran hoy en su momento histórico más urgente. Por lo que existen varias cuestiones éticas y dilemas morales que afrontar.

En este tema conocerás:

Dilemas morales con la inteligencia artificial.

La diferencia entre la inteligencia artificial general y la inteligencia artificial débil.





La inteligencia artificial débil (ANI) es aquella donde la **tecnología supera al humano en una tarea específica** y se centra en un subconjunto de habilidades cognitivas.

Actualmente existen varios ejemplos de este tipo de inteligencia artificial:

Asistentes personales: Siri, Alexa y Cortana.

Automóviles autónomos.

Herramientas de etiquetado facial en fotografías digitales.

El algoritmo de Google que determina, en función de cierta calificación, el sitio que debe aparecer al principio de los resultados de una búsqueda.

Sistemas que recomiendan el tipo de producto a comprar con base en el historial de navegación.

Filtros de correo no deseado que permiten tener un buzón libre de *spam*.





Los **dilemas morales** a los que se enfrenta la IA débil son múltiples, por ejemplo:

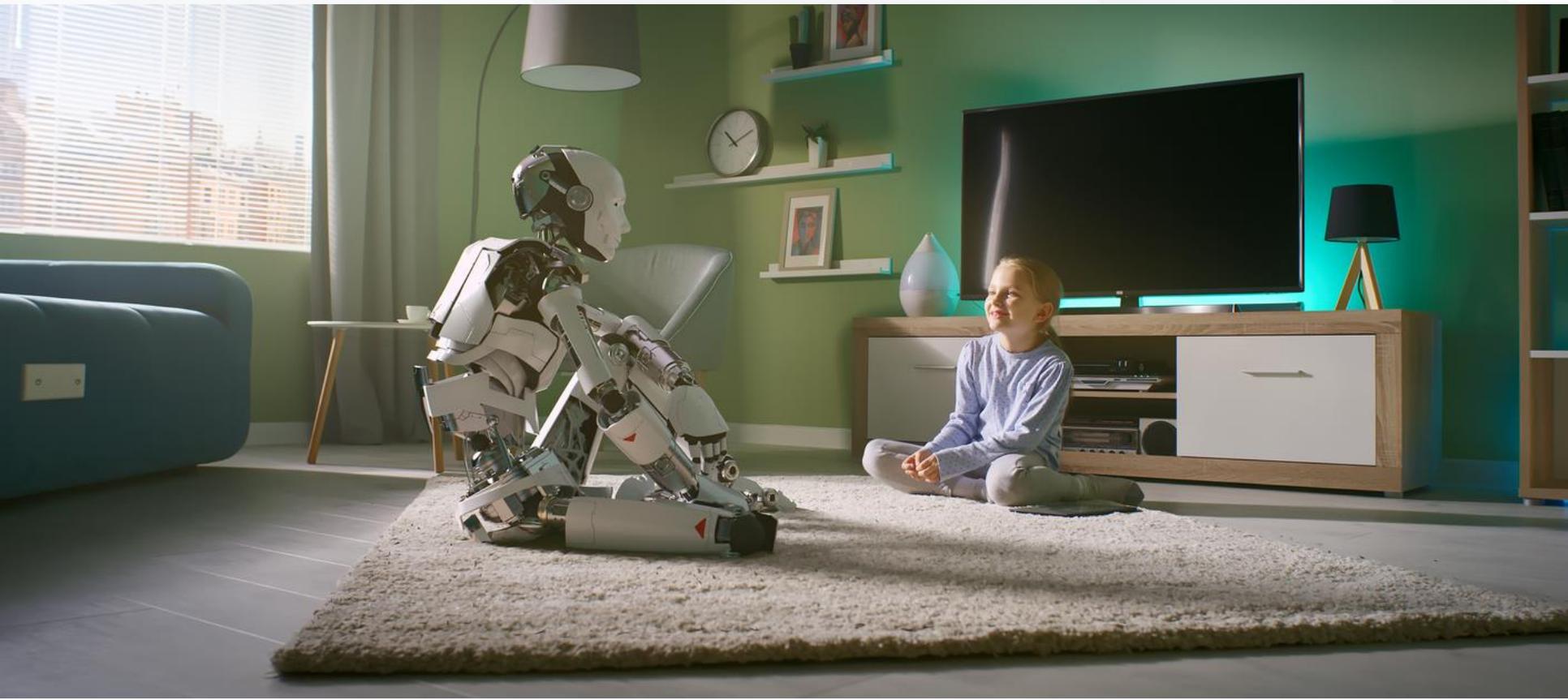
1	<b>Desigualdad económica y laboral:</b> un mayor uso de la tecnología puede desplazar puestos laborales, así como generar la necesidad de mayores conocimientos técnicos y capacitación.
2	<b>Influencia en el comportamiento humano:</b> existen aplicaciones que usan algoritmos para provocar adicción en el uso de estas.
3	<b>Sesgo de datos</b> de entrenamiento de algoritmos para contratación de personal o para determinar culpabilidad en ciertos crímenes.
4	<b>Noticias falsas:</b> actualmente proliferan las falsificaciones profundas, por ejemplo, una técnica que usa IA para sobreponer imágenes, audio y videos de otras personas sobre una en particular, creando una falsa impresión de los audios y medios originales.
5	<b>Privacidad y seguridad en los datos:</b> el riesgo de utilizar IA en los datos personales es enorme, incluso con las regulaciones existentes.





El concepto de inteligencia artificial general (AGI) se utiliza para referirse a aquellas **máquinas que muestran inteligencia humana**, son sistemas que pueden procesar, de forma exitosa, cualquier tipo de tarea que un humano pueda realizar.

Es el tipo de sistemas como los que se muestran en las películas como el *Hombre Bicentenario* o *Her*, donde se ve la interacción humana con máquinas y sistemas operativos que son conscientes, sensibles e impulsados por las emociones y autoconciencia.





Como resultado de conferencias en AGI, se ha llegado a la conclusión de que una plataforma AGI requiere de razonamientos **deductivo**, **inductivo** y **abductivo**.

<b>Deductivo</b>	El razonamiento inicia con el planteamiento de una hipótesis general y examina las posibilidades de llegar a una conclusión lógica específica. La inferencia deductiva genera la predicción de las consecuencias a partir de una teoría, esto es, se predicen las observaciones que deberían ser, en caso de que la teoría fuera correcta.
<b>Inductivo</b>	Realiza generalizaciones a partir de observaciones específicas. La inferencia va de lo particular a lo general. Se hacen observaciones de las que se obtiene un patrón para realizar generalizaciones e interferir en una explicación o teoría.
<b>Abductivo</b>	El razonamiento abductivo comienza con un conjunto incompleto de observaciones de las que se deriva la explicación más probable para ese grupo de observación. Formula y prueba hipótesis usando la mejor información posible.





Escribe un ensayo sobre los dilemas éticos, sociales y morales de la inteligencia artificial





La inteligencia artificial promete ser capaz de realizar múltiples tareas, recordar y memorizar la información de forma perfecta, funcionar continuamente sin afectaciones ni interrupciones, realizar una gran cantidad de cálculos a velocidades récord, examinar datos y documentos extensos, así como tomar decisiones imparciales.

En la medida que la IA reemplaza puestos de trabajo, se da un gran debate sobre ética y sobre si los gobiernos deberían intervenir para monitorear y regular su uso y crecimiento, ya que es muy alto el riesgo de que la IA podría transformar las relaciones humanas, aumentar la discriminación, invadir la privacidad personal e incluso plantear amenazas a la seguridad, al permitir el uso de armas autónomas

