

Inteligencia Artificial

Unidad 1: Representación Avanzada del Conocimiento

TEMA 1: Fundamentos de Inteligencia Artificial

Módulo 3: Ética en la Inteligencia Artificial

Unidad 1

Representación Avanzada del Conocimiento

TEMA 1: Fundamentos de Inteligencia Artificial

Sesión 3

MÓDULO 3: Ética en la Inteligencia Artificial



Contenido

1. ¿Por qué I.A. y ética?
2. Principios básicos de ética e I.A.
3. Inteligencia artificial: ejemplos de dilemas éticos
4. UNESCO: 5 cosas que debes saber sobre la I.A.
5. Uso malicioso de la I.A: Desinformación + I.A.



Preguntas

1. ¿Por qué I.A. y ética?

“La ética es la disciplina filosófica que estudia el bien y el mal y sus relaciones con la moral y el comportamiento humano”. Definiciones de Oxford Languages

Aristóteles fue el fundador de la ética

¿Por qué aplicar Ética
en la I.A.?



1. ¿Por qué I.A. y ética?

- ❖ Porque somos la primera generación de humanos que tienen a las máquinas tomando las decisiones que normalmente hacemos. Ejemplo: toma de decisiones legales (quien va o no a la cárcel), control de migración en aeropuertos, ingreso a instalaciones con reconocimiento facial, etc.).
- ❖ Porque las personas ya no solo requerimos gran tecnología, sino que requerimos gente que diseñe tecnología de forma “responsable”.
- ❖ Porque necesitamos habilitar a las máquinas para que piensen éticamente.



1. ¿Por qué I.A. y ética?

La ética en la Inteligencia Artificial



Video: <https://youtu.be/VyR-tfIWBQM> (2 min)

1. ¿Por qué I.A. y ética?

Podemos ser formidables desde el punto de vista ético pero no desarrollamos nada, y al revés, tampoco podemos concentrarnos en el desarrollo ignorando por completo la ética.

Ranga Yogeshwar
Físico y periodista científico



2. Principios básicos de ética e I.A.

FATE, son las siglas en ingles de la cualidades o principios que deben ir a la par con las investigaciones en I.A.

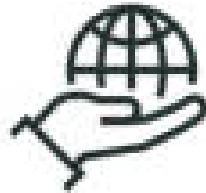
FATE



FAIRNESS (Justicia)

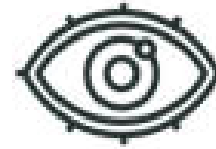
La tecnología debe ser justa y no discriminar.

Debe haber garantías de que se trate con igualdad.



ACCOUNTABILITY (Responsabilidad)

Tener bien claro a quien se le atribuye la responsabilidad del uso de los sistemas de I.A.



TRANSPARENCY (Transparencia)

Los sistemas de I.A deben ser legibles y que se expliquen en términos no técnicos.

Que expliquen:

- Cómo funcionan
- Qué datos usan
- Cómo los usan, etc.



ETHICS (Ética)

Hay que asegurarse que el uso y aplicación de los sistemas de I.A y su desarrollo, están de acuerdo con el **conjunto de valores éticos** y principios que como sociedad respetamos y aceptamos.

2. Principios básicos de ética e I.A.

FAIRNESS (Justicia)

❖ Necesitamos crear I.A.

A pesar que no reduce la injusticia en nuestra sociedad pero que al menos la mantenga igual.

❖ Evitar los prejuicios humanos

Se deben realizar esfuerzos para **evitar** crear refuerzos o perpetuar los **prejuicios humanos** durante todo el ciclo de vida de un sistema de I.A.

La justicia a través de la I.A. es usada hoy en:

- Justicia penal
- Investigación Criminal
- Empleo y contratación
- Finanzas y crédito



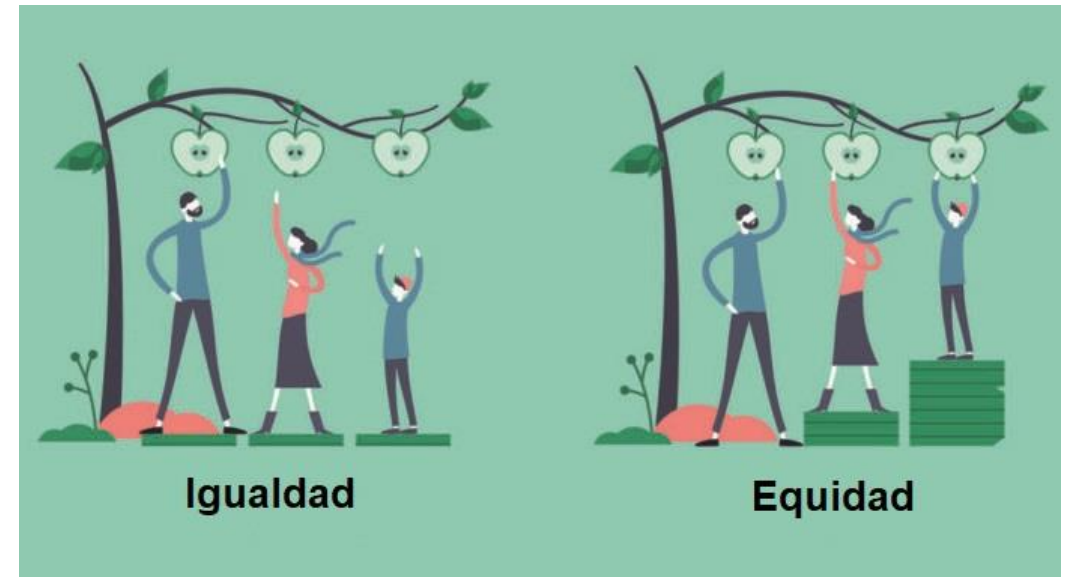
2. Principios básicos de ética e I.A.

FAIRNESS (Justicia)

La **Equidad** o **Justicia** en la I.A. se relaciona no solo con el componente técnico, sino también con el contexto social en el que se implementa el sistema.

La Equidad en la I.A. es un desafío socio-tecnológico:

- Necesitamos una mayor diversidad de personas en desarrollo y desplegando sistemas de I.A.
- Las suposiciones y decisiones tomadas por los equipos en cada etapa del ciclo de vida de desarrollo e implementación de la I.A. puede introducir **sesgos**.
- Hay que evitar el "**Trash in, Trash out**"
 - ✓ No podemos delegar esto en una o dos personas.
 - ✓ Todos debemos pensar activamente en esto.

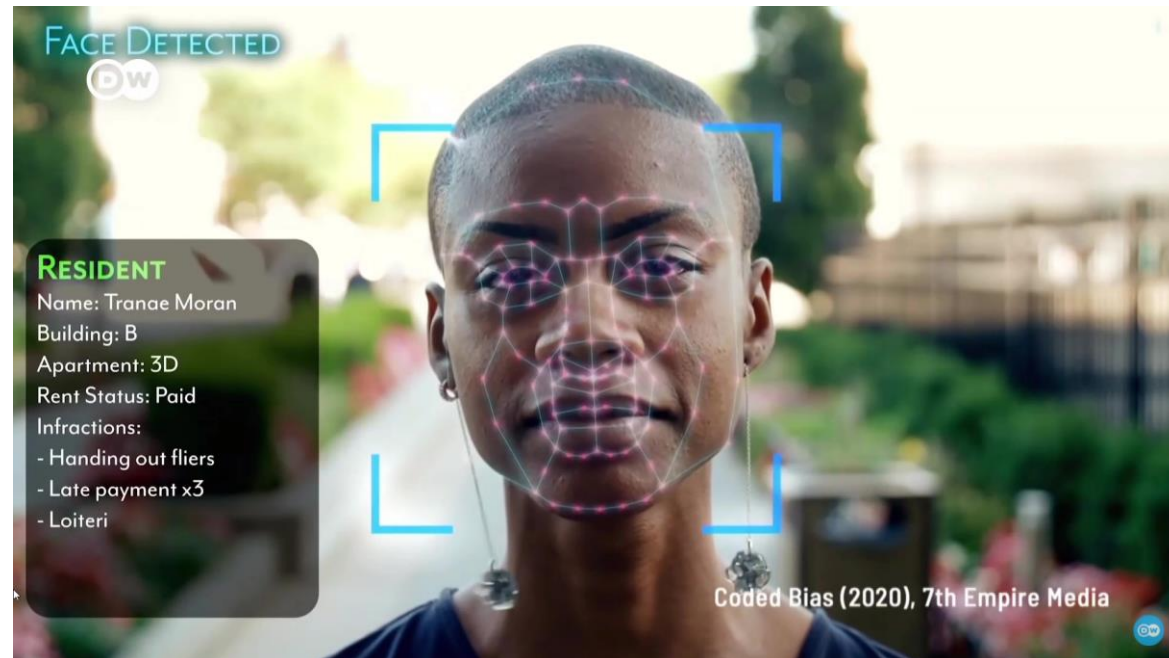


2. Principios básicos de ética e I.A.

FAIRNESS (Justicia)

Discriminación en IA – ¿Quiénes son las víctimas?

¿Cómo discrimina
la I.A.?



Video: <https://www.youtube.com/watch?v=mWGzicjqAc0> (13 min)

2. Principios básicos de ética e I.A.

ACCOUNTABILITY (Responsabilidad)

Somos responsables de cómo nuestra I.A. impacta en el mundo.

- Necesitamos **ayudar a los usuarios finales también a ser responsables de sus decisiones** con nuestra I.A.
- Los seres humanos siempre debemos **seguir siendo responsables de las decisiones** ayudadas o mediadas por los sistemas de I.A.
- Siempre debe existir **supervisión humana** a lo largo del ciclo de vida de los sistemas que involucran A.I.
- Las partes interesadas debemos implementar:
 - ✓ Mecanismos de evaluación.
 - ✓ Transparencia y auditabilidad.
 - ✓ Rendición de cuentas.



2. Principios básicos de ética e I.A.

TRANSPARENCY (Transparencia)

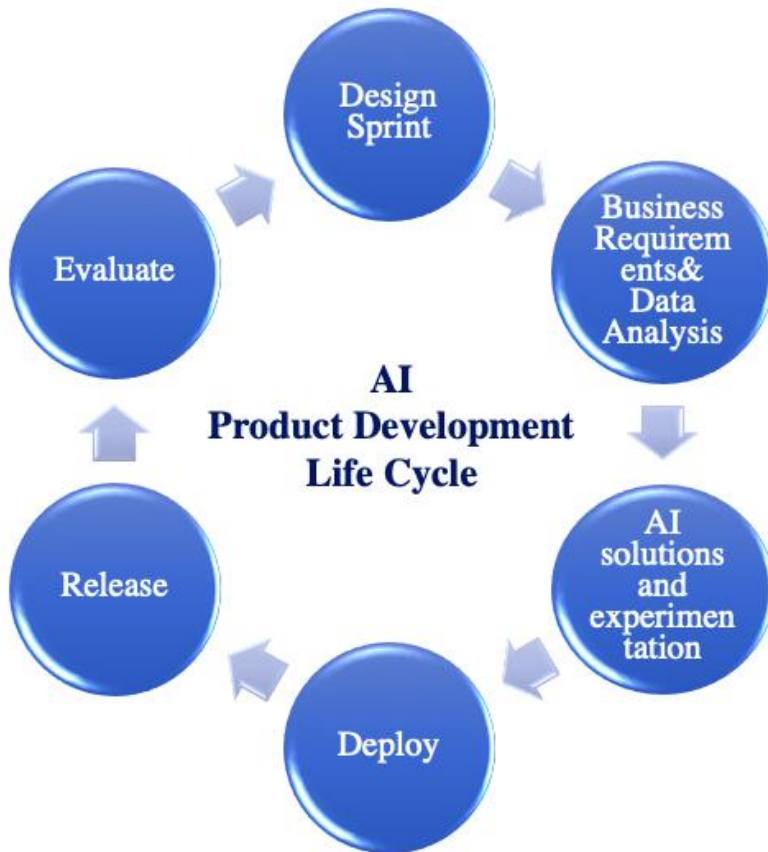


La transparencia en la I.A. es importante porque permite:

- Mitigar la injusticia
- Ayudar a los desarrolladores a depurar sus sistemas de I.A.
- Incrementar la confianza de los usuarios finales que ahora pueden entender mejor la plataforma.

2. Principios básicos de ética e I.A.

TRANSPARENCY (Transparencia)



Si aplicamos la transparencia al ciclo de vida del desarrollo de la I.A.:

Datos:

Cada **conjunto de datos** puede tener una **hoja de datos(*)** que explique su:

- Motivación
- Creación
- Mantenimiento
- Uso previsto

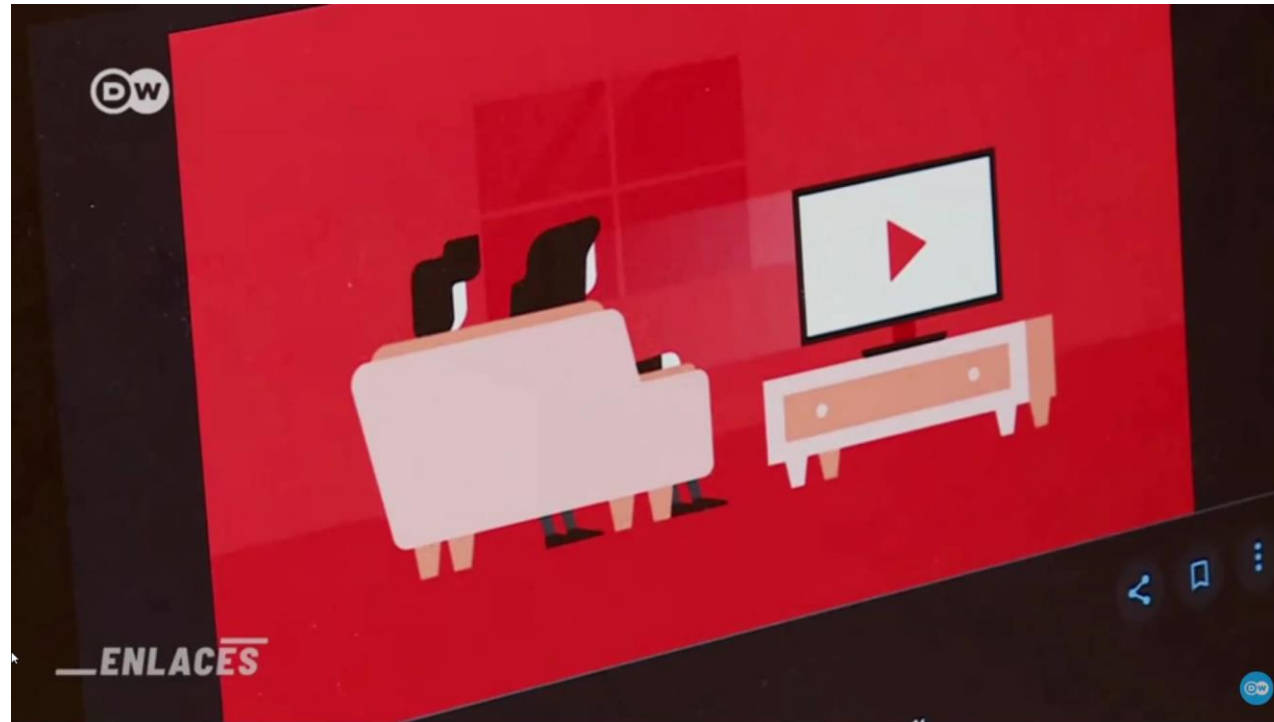
Las hojas de datos pueden ayudar:

1. Los profesionales de los datos comprendamos y descubramos:
 - Posibles sesgos en los datos.
 - Supuestos involuntarios que estemos haciendo.
2. Consumidores de datos:
 - Determinar si un conjunto de datos es adecuado para sus necesidades.

2. Principios básicos de ética e I.A.

ETHICS (Ética)

¿Dinero antes que ética? El Algoritmo de YouTube



Video: <https://www.youtube.com/watch?v=iY7nxBVRtTA> (6 min)

3. Inteligencia artificial: ejemplos de dilemas éticos

IA sesgada



© UNESCO

Fuente: <https://es.unesco.org/artificial-intelligence/ethics/cases>

3. Inteligencia artificial: ejemplos de dilemas éticos

IA SESGADA: Los sistemas de inteligencia artificial dan resultados sesgados

Si buscas: “**los mejores líderes de todos los tiempos**” en tu motor de búsqueda favorito, probablemente verás una lista de las personalidades masculinas más destacadas del mundo. ¿Cuántas mujeres puedes contar?

Yo conté 8 sólo mujeres 😞



Algoritmos de los Buscadores

La tecnología de los motores de búsqueda no es neutral, ya que procesa **macrodatos** y prioriza los resultados con **la mayor cantidad de clics** dependiendo tanto de las preferencias del usuario como de la ubicación.

3. Inteligencia artificial: ejemplos de dilemas éticos

Coche autónomo



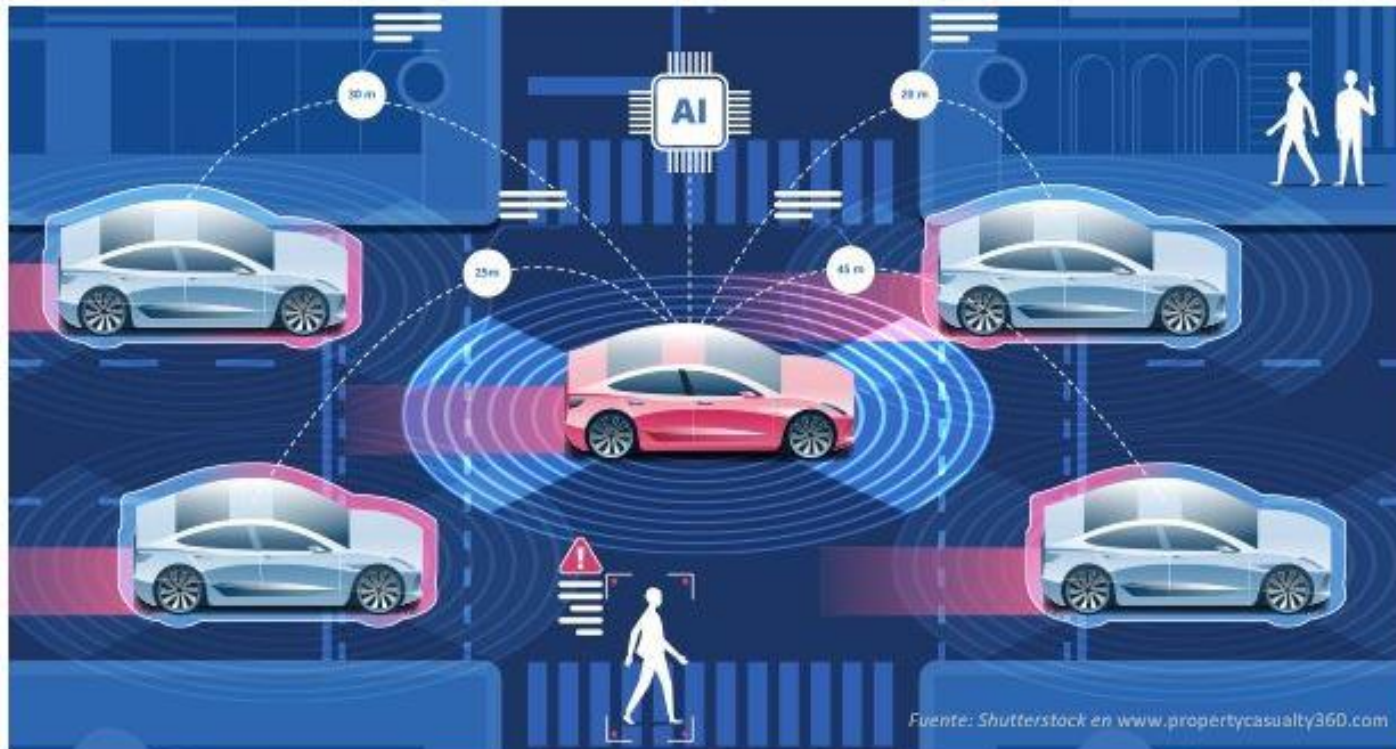
© Shutterstock.com/Senha

Fuente: <https://es.unesco.org/artificial-intelligence/ethics/cases>

3. Inteligencia artificial: ejemplos de dilemas éticos

Vehículo Autónomo: No es un conductor humano quien tomará la decisión, sino el algoritmo del coche.

Cuando un conductor decide frenar de golpe para evitar chocar con un peatón, está tomando la decisión moral de trasladar el riesgo del peatón a las personas que van en el coche.



Algoritmos de los vehículos autónomos

¿A quién elegirías, a la abuela o al niño?
¿Crees que sólo hay una respuesta correcta?

Este es un típico dilema ético, que muestra la importancia de la ética en el desarrollo de las tecnologías.

3. Inteligencia artificial: ejemplos de dilemas éticos

La IA crea arte



© Shutterstock.com/agsandrew

Fuente: <https://es.unesco.org/artificial-intelligence/ethics/cases>

3. Inteligencia artificial: ejemplos de dilemas éticos

IA Creadora de Arte: Un autor humano es reemplazado por máquinas y algoritmos.

- ¿Qué sucede cuando la I.A. tiene la capacidad de crear obras de arte por sí misma?
- ¿Hasta qué punto se pueden atribuir los derechos de autor?
- ¿Puede y debe reconocerse a un algoritmo como autor y gozar de los mismos derechos que un artista?



El algoritmo del pintor

La inminente subasta de la primera obra creada por inteligencia artificial plantea varias interrogantes acerca de la creación artística (2018)

Fuente: shorturl.at/sxB68

3. Inteligencia artificial: ejemplos de dilemas éticos

AI en el Tribunal de Justicia



© Shutterstock.com/icedmocha

Fuente: <https://es.unesco.org/artificial-intelligence/ethics/cases>

3. Inteligencia artificial: ejemplos de dilemas éticos

I.A. en los tribunales de justicia: Un juez, un abogado es reemplazado por máquinas y algoritmos.

- ¿La I.A. podría evaluar los casos y aplicar la justicia de una manera mejor, más rápida y más eficiente que un juez?
- La I.A. tomaría decisiones basadas en decisiones informadas y carentes de todo sesgo y subjetividad.

PERO existen muchos desafíos éticos:

- ☐ **La falta de transparencia de las herramientas de la I.A.:** Las decisiones de la I.A. no siempre son inteligibles para los humanos.
- ☐ **La I.A. no es neutral:** Las decisiones basadas en la I.A. son susceptibles de inexactitudes, resultados discriminatorios, sesgos, etc.
- ☐ Nuevas preocupaciones por la **equidad** y el riesgo para los **derechos humanos** y otros valores fundamentales.

Responde a esta pregunta:

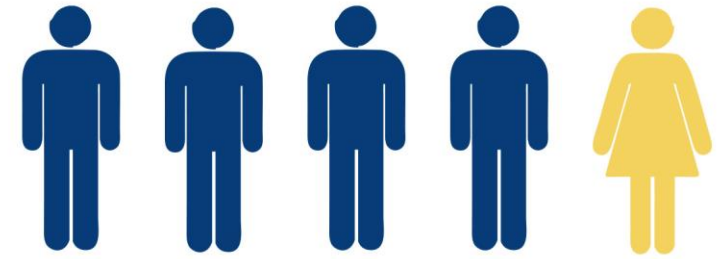
¿Te gustaría ser juzgado por un robot en un tribunal? ¿aún y cuando no estemos seguros de cómo llegaría a sus conclusiones?



4. UNESCO: 5 cosas que debes saber sobre la I.A.

1. La IA tiende a ampliar las diferencias de género existentes

- ☐ Las mujeres están subrepresentadas en la industria, por tanto, los prejuicios y estereotipos de género se están reproduciendo en las tecnologías de la I.A.
- ☐ No es una coincidencia **que las asistentes personales virtuales** como Siri, Alexa o Cortana sean "**femeninas**" por defecto.
- ☐ El servilismo y a veces la sumisión que expresan son un ejemplo de cómo la I.A. puede (seguir) reforzando y difundiendo los prejuicios de género en nuestras sociedades.



Solo el 22%
de los profesionales
que se dedican a la IA
son MUJERES

4. UNESCO: 5 cosas que debes saber sobre la I.A.

2. Es probable que el crecimiento impulsado por la I.A. sea muy desigual

- ☐ Para 2030, se espera que los beneficios económicos sean mayores en **China** y **América del Norte**, y que representen el 70% del impacto económico mundial de la I.A.
- ☐ La concentración de la I.A. en manos de pocos países de altos ingresos probablemente dejará a los países en desarrollo muy atrás.
- ☐ Los países en desarrollo no se beneficiarán, o lo harán muy poco, de las tecnologías de la I.A. y carecerán de propiedad sobre dichas tecnologías.

Se espera que la IA
genere casi  **4 billones de dólares**
de valor añadido para
2022



4. UNESCO: 5 cosas que debes saber sobre la I.A.

3. La I.A. puede ser una herramienta poderosa para abordar el cambio climático y las cuestiones ambientales

- ☐ Los impactos del cambio climático están empeorando a medida que el planeta se sigue calentando.
- ☐ La extracción de datos consume casi el diez por ciento de la energía a nivel mundial. Es esencial abordar el alto consumo de energía de la I.A. y el consiguiente impacto en la emisión de carbono.
- ☐ Los modelos impulsados por la I.A. podrían, por ejemplo, ayudar a mejorar la gestión de los ecosistemas y la restauración del hábitat, esencial para menguar la disminución de las poblaciones de peces y fauna silvestre.

**LA EXTRACCIÓN
DE DATOS**



**consume casi el
10% de la energía
a nivel mundial**

Source: Anders S.G. Andrae et al.



4. UNESCO: 5 cosas que debes saber sobre la I.A.

4. La I.A. ha demostrado su valor para enfrentar la pandemia de COVID-19

- ☐ La I.A. ayuda a los investigadores a procesar grandes cantidades de datos en la carrera por encontrar una vacuna o un tratamiento
- ☐ La I.A. ha participado para contener la propagación del virus a través de tecnologías de prueba, rastreo y seguimiento.
- ☐ **Desventaja:** el uso de la IA durante esta pandemia ha reabierto las preocupaciones sobre la privacidad, la protección de los datos y el uso de los datos más allá de las necesidades de rastreo del virus.



4. UNESCO: 5 cosas que debes saber sobre la I.A.

¿Puede la Inteligencia Artificial detectar una pandemia?



Video: <https://youtu.be/7ozzhS8vaYA> (10 min)

4. UNESCO: 5 cosas que debes saber sobre la I.A.

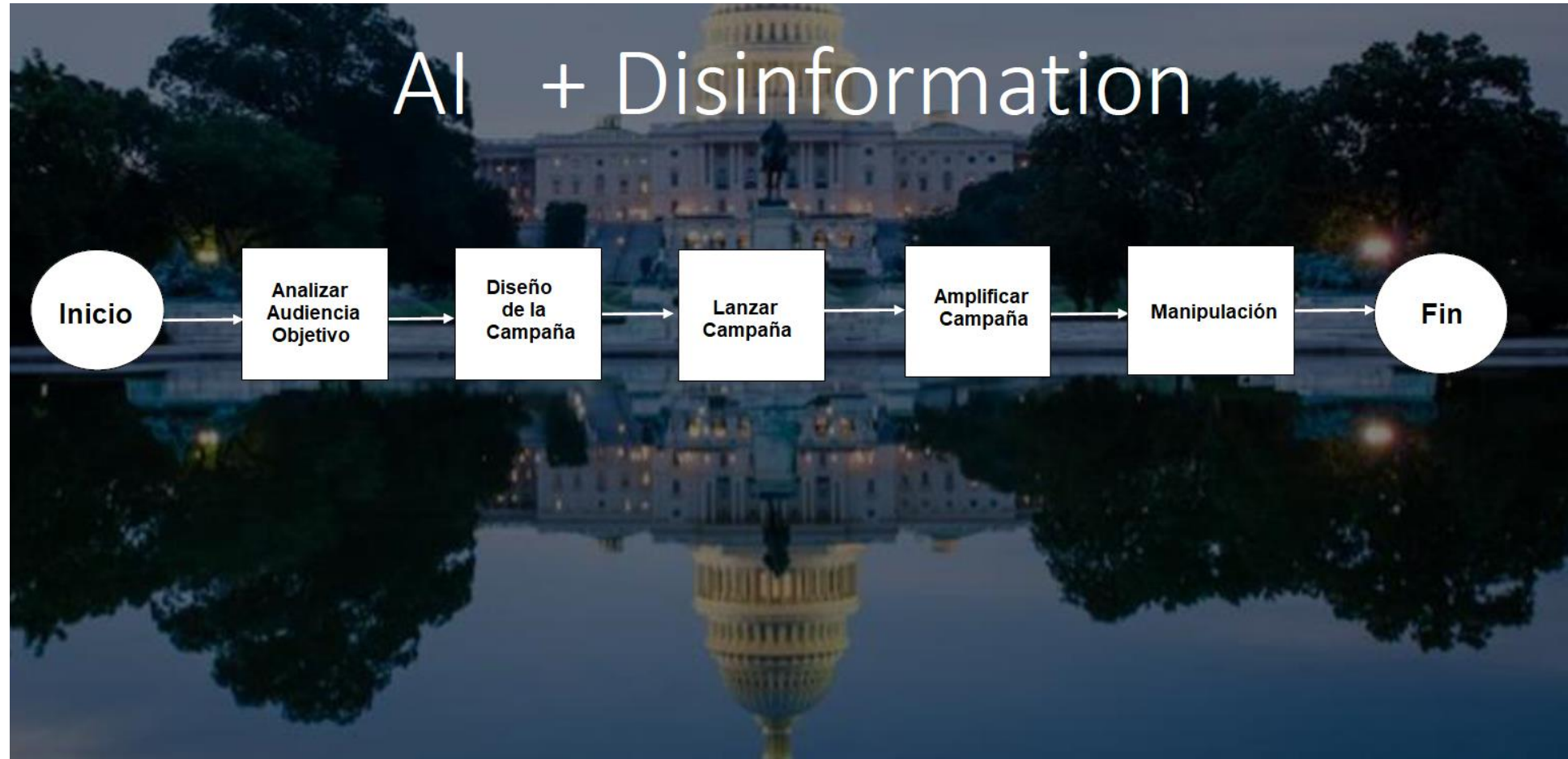
5. La I.A. no puede ser una zona sin ley

- ☐ La I.A. ya está en nuestras vidas, dirigiendo nuestras elecciones, a menudo en formas que pueden ser perjudiciales.
- ☐ Existen algunos vacíos legislativos alrededor de la industria que deben ser tratados rápidamente.
- ☐ Urge convenir exactamente qué valores deben ser consagrados, y qué normas deben ser aplicadas.
- ☐ La **I.A. es mundial**, por lo que necesitamos un instrumento mundial para regularla.



Fuente: <https://es.unesco.org/artificial-intelligence/ethics>

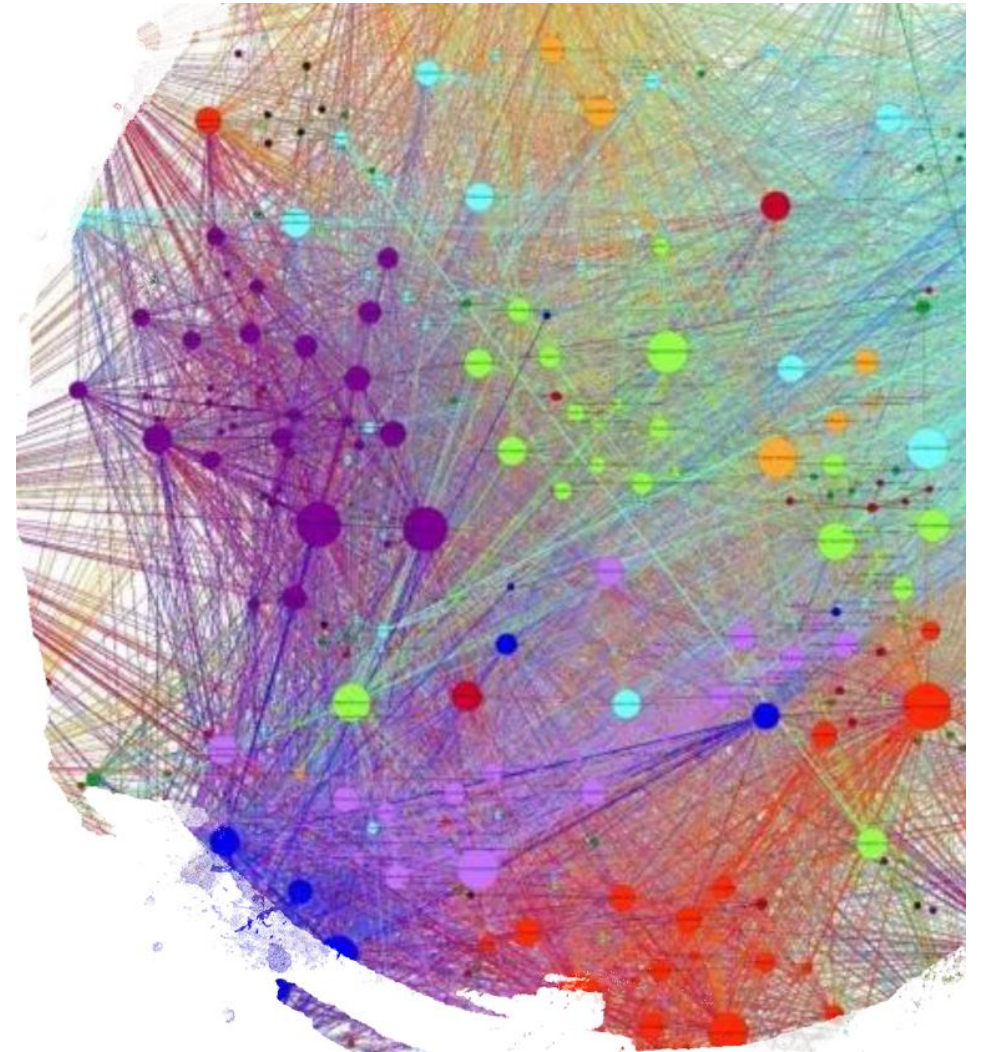
5. Uso malicioso de la I.A: Desinformación + I.A.



5. Uso malicioso de la I.A: Desinformación + I.A.

Analizar audiencia objetivo

- Identificar personas influyentes clave, que pueden ser abordadas con ofertas (maliciosas) o dirigidas con desinformación.



5. Uso malicioso de la I.A: Desinformación + I.A.

Diseño de la Campaña

Diseño, hiper personalizado de campañas de desinformación

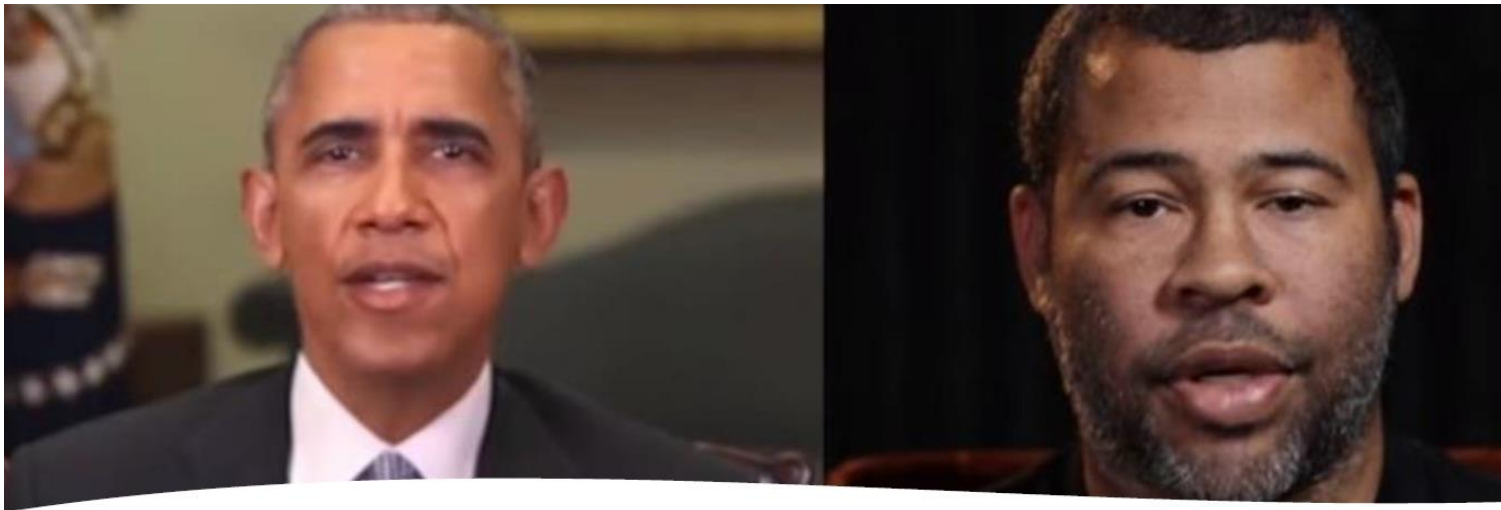
- Crea mensajes personalizados para afectar el comportamiento de su voto.



5. Uso malicioso de la I.A: Desinformación + I.A.

Diseño de la Campaña

Informes de noticias falsos con video y audio realistas



- Se hacen videos muy realistas de líderes estatales que parecen hacer comentarios incendiarios que en realidad nunca hicieron.

5. Uso malicioso de la I.A: Desinformación + I.A.

Lanzar Campaña

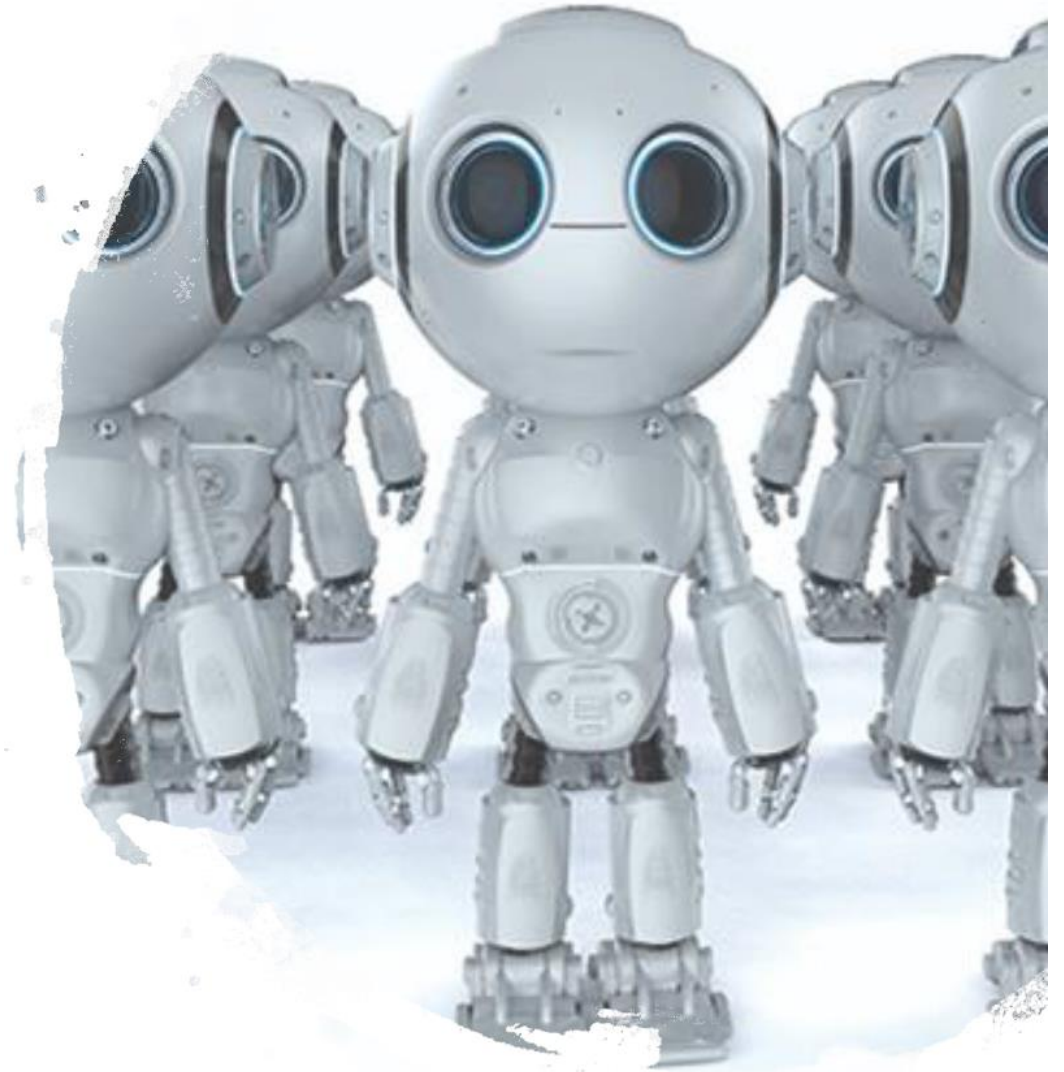
- Se entrega contenido usando diferentes cuentas para crear la ilusión de que hay múltiples fuentes de una misma historia.



5. Uso malicioso de la I.A: Desinformación + I.A.

Amplificar

- Se amplifica el contenido retransmitiendo la historia usando cuentas falsas, de forma masiva, gran volumen de me gusta, comparte, etc.
- Uso de bots en las redes sociales es lo usual.



5. Uso malicioso de la I.A: Desinformación + I.A.

Manipulación

- Manipular la reacción del objetivo infiltrándose en conversaciones sobre el contenido.
- Incitar conflicto y/o fortalecer la ilusión de consenso por comentarios en secciones de publicaciones en línea (trolling).



PREGUNTAS

Dudas y opiniones