



UNIVERSIDAD
SANTO TOMÁS
— BUCARAMANGA —
VIGILADA MINEDUCACIÓN - SNIES: 1705



FERIA INTERNACIONAL
CIENTEC 2022
DE CIENCIA Y TECNOLOGÍA

Gobernanza para la Inteligencia Artificial

Problemas y soluciones

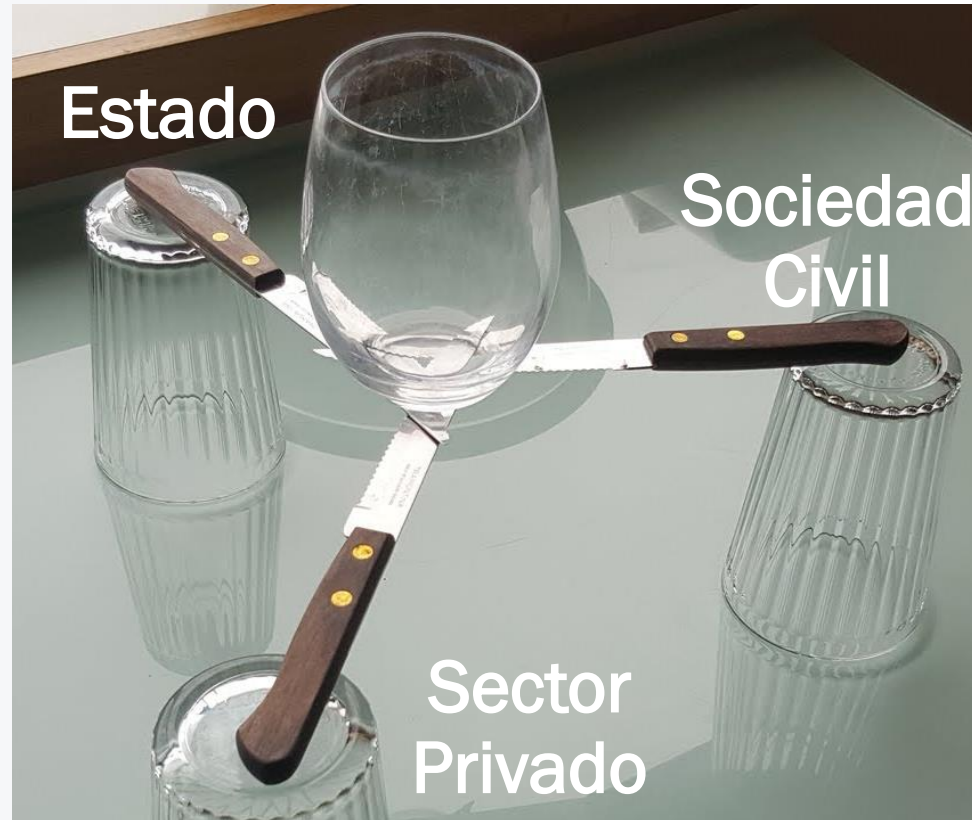
Facultad Ingeniería Mecatrónica
Ing. Carlos Daniel Pimentel Díaz, M.Sc.

Contenido

1. ¿Qué es la gobernanza de la IA?
2. Análisis del contexto actual de la IA.
3. Enfoques del contexto actual de la IA.
4. Objetivo de la gobernanza de la IA.
5. Acciones a tomar a corto plazo.
6. Lo que otros países hacen para regular la IA.
7. Conclusiones.

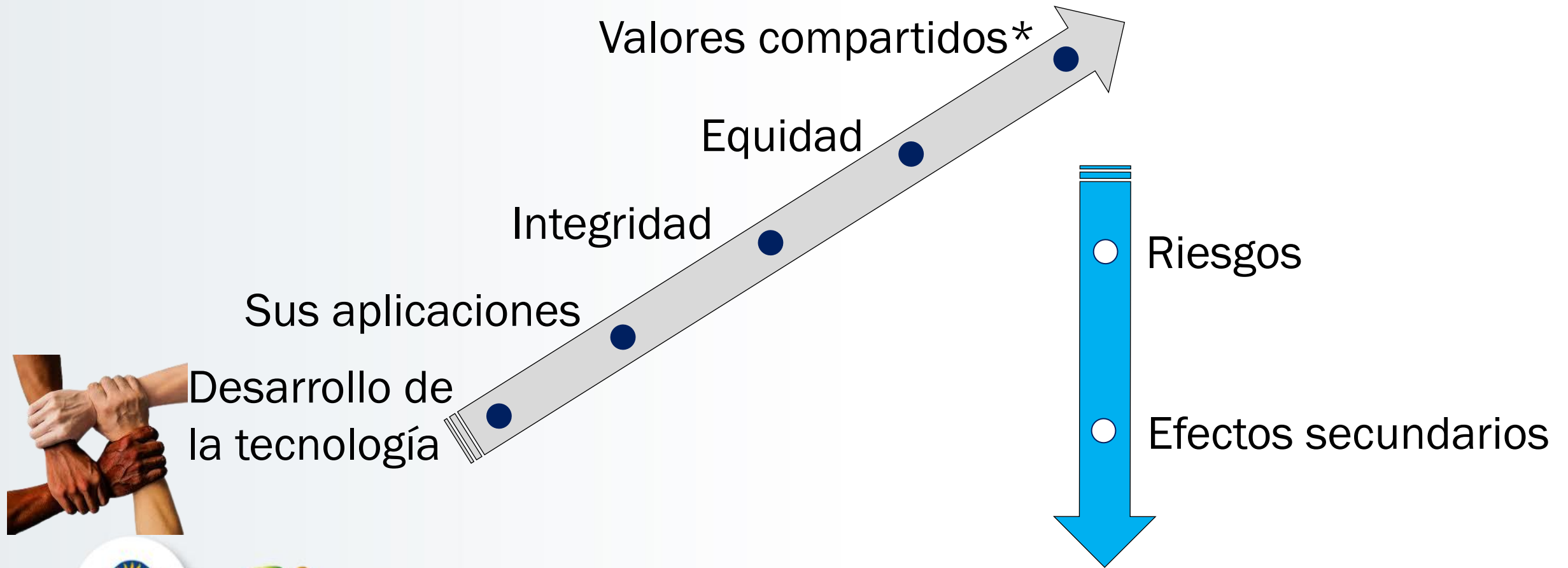
¿Qué es gobernanza y cómo construirla en el contexto de la IA?

- Gobernanza de la IA:



Organizar y distribuir el poder y sus consecuencias.

Gobernanza de la Inteligencia Artificial



El contexto actual de la IA - Análisis

- Regulaciones actuales:

Encontrar “soluciones a problemas existentes”



Minimizar los daños causados por la IA.



El contexto actual de la IA - Análisis

- Regulaciones futuras:

“Diagnosticar problemas” para
examinar...



origen y causas de problemas
sociales y económicos.



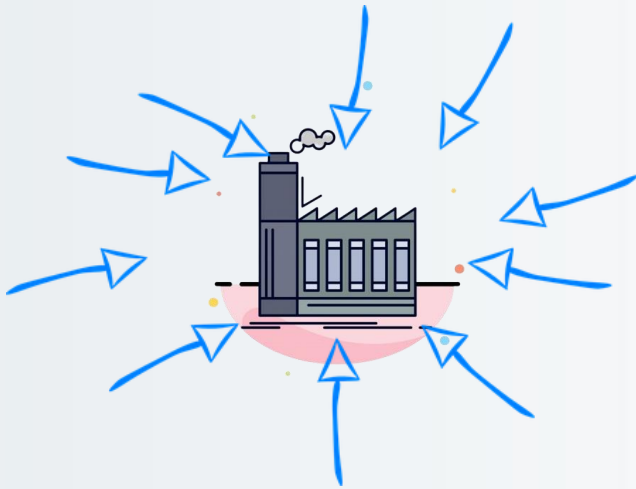
El contexto actual de la IA - Análisis

- El contexto actual de IA se encamina a tres enfoques diferentes:
 1. Que la industria se auto regule.
 2. Diseño de algoritmos menos sesgados.
 3. Implementar marcos regulatorios basados en DD.HH.



1° Enfoque: que la industria se auto regule

- La auto regulación es la opción predominante.



Auto asignación de ciertos principios éticos.



Marco para desarrollo y despliegue de IA.



Más para crecimiento personal que comunitario.

Todo esto debido a que no existe un marco regulatorio bien definido.

1° Enfoque: que la industria se auto regule

Principios auto asignados:

- ✓ Transparencia,
- ✓ Imparcialidad,
- ✓ Equidad,
- ✓ Responsabilidad y
- ✓ Seguridad.

El inconveniente:

Marcos éticos que no son ni vinculantes, ni obligatorios.

Es tan sencillo comprometerse a cumplirlos como violarlos.

1° Enfoque: que la industria se auto regule

Auto regulaciones complacientes “suficientes” para minimizar los riesgos de la IA.

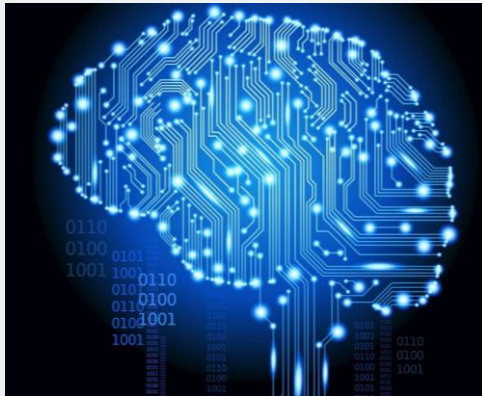
Evitar otras formas de regulación que les resten margen de autonomía y de control.

El problema:

Diferentes tipos de industria, instaladas en diferentes economías interpretan, de forma subjetiva, diferentes significados de principios como la equidad y la transparencia.

2° Enfoque: algorítmicos más justos, responsables y transparentes

- Propuesto por la comunidad de la ciencia de la programación.
- Requiere de debates previos imprescindibles y no siempre sencillos.
- Ejemplo: Algoritmo COMPAS.



2° Enfoque: algorítmicos más justos, responsables y transparentes

- Características de COMPAS:
 - Soporte en la toma de decisiones.
 - Algoritmo complejo.
 - Utiliza muchas variables para llegar a un resultado.
 - Otros algoritmos más sencillos realizan predicciones aproximadamente tan bien como este.
 - Puede dar resultados sesgados: falta en principio de equidad en la justicia.

Two Petty Theft Arrests



VERNON PRATER

LOW RISK

3



BRISHA BORDEN

HIGH RISK

8

Two Drug Possession Arrests



DYLAN FUGITT

LOW RISK

3



BERNARD PARKER

HIGH RISK

10

Tomada de: http://www.nvp.co.kr/news/photo/201710/121489_55331_347.png

2° Enfoque: diseño de algorítmicos más justos, responsables y transparentes

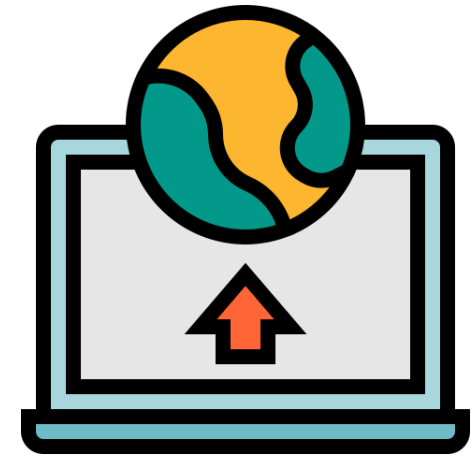
- La equidad es un atributo de los sistemas sociales, no de los sistemas técnicos y solo se puede entender dentro de un contexto social.



La construcción de mejores algoritmos...



Debate sobre si es realmente necesario...

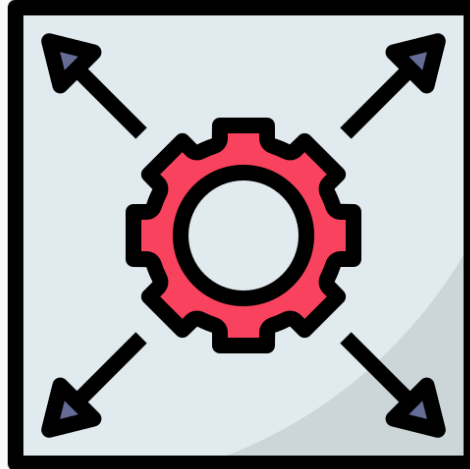


Desplegar IA para desarrollar ciertas tareas.

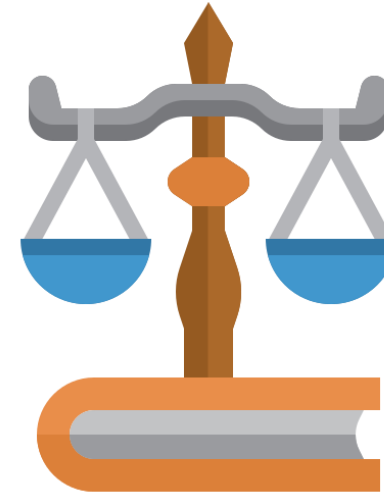
3° Enfoque: implementar marcos regulatorios basados en el respeto a los DD.HH.



Se aplican marcos universales y vinculantes.



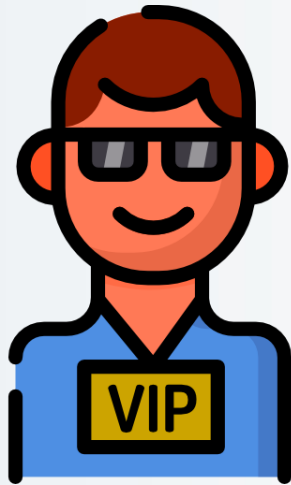
Responsabilidades de Gobiernos y empresas delimitadas.



Instancias jurídicas locales, regionales e internacionales: subsanar los incumplimientos

3° Enfoque: implementar marcos regulatorios basados en el respeto a los DD.HH.

- Dificultades de este enfoque:



Situar al individuo en el centro de su accionar.



Existencia de derechos y deberes.



Participación de ciertos actores “perversos”.

1° Dificultad: Situar al individuo en el centro de su accionar.

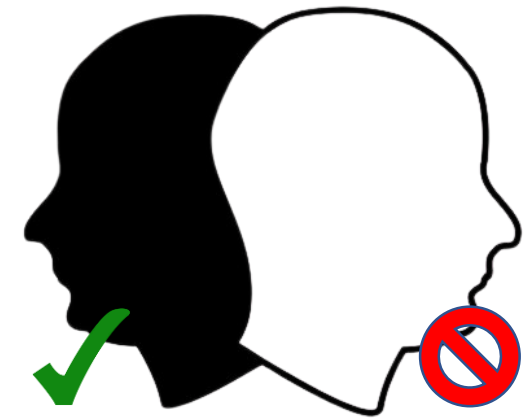
- Se pueden generar algoritmos que tienden a favorecer perjuicios individuales: créditos bancarios y seguros de vida.



Entrenados con datos de comportamiento y consumo



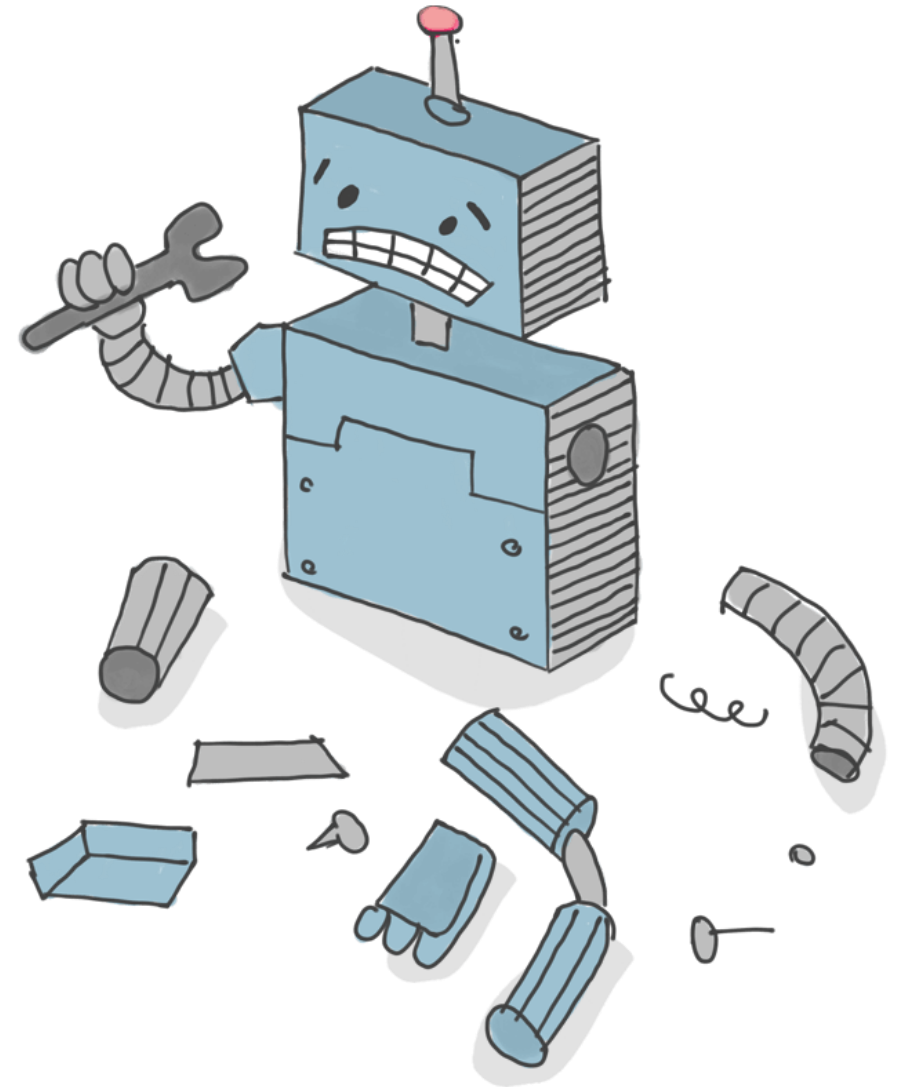
Tras múltiples iteraciones...

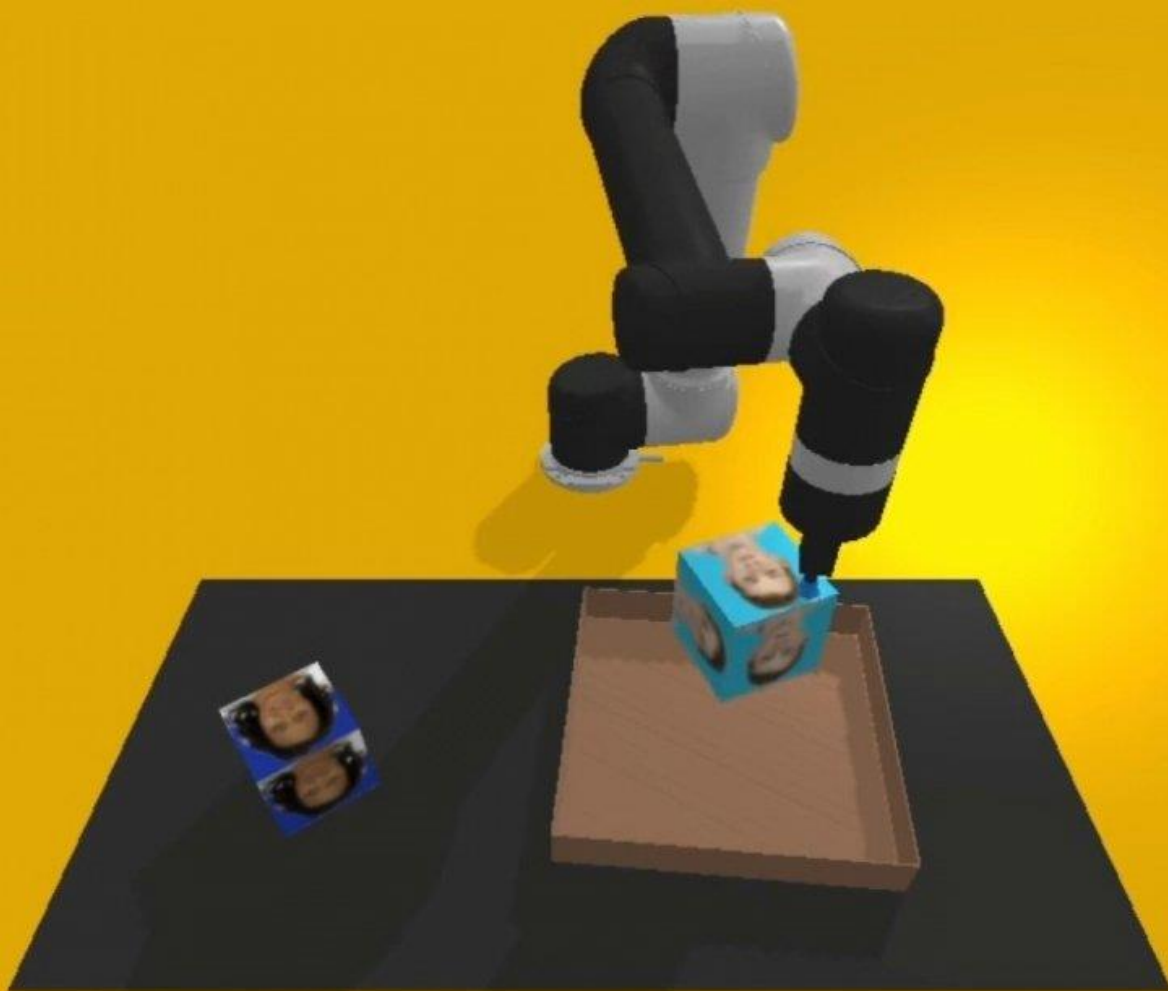


Nuevas categorías de individuos confiables.

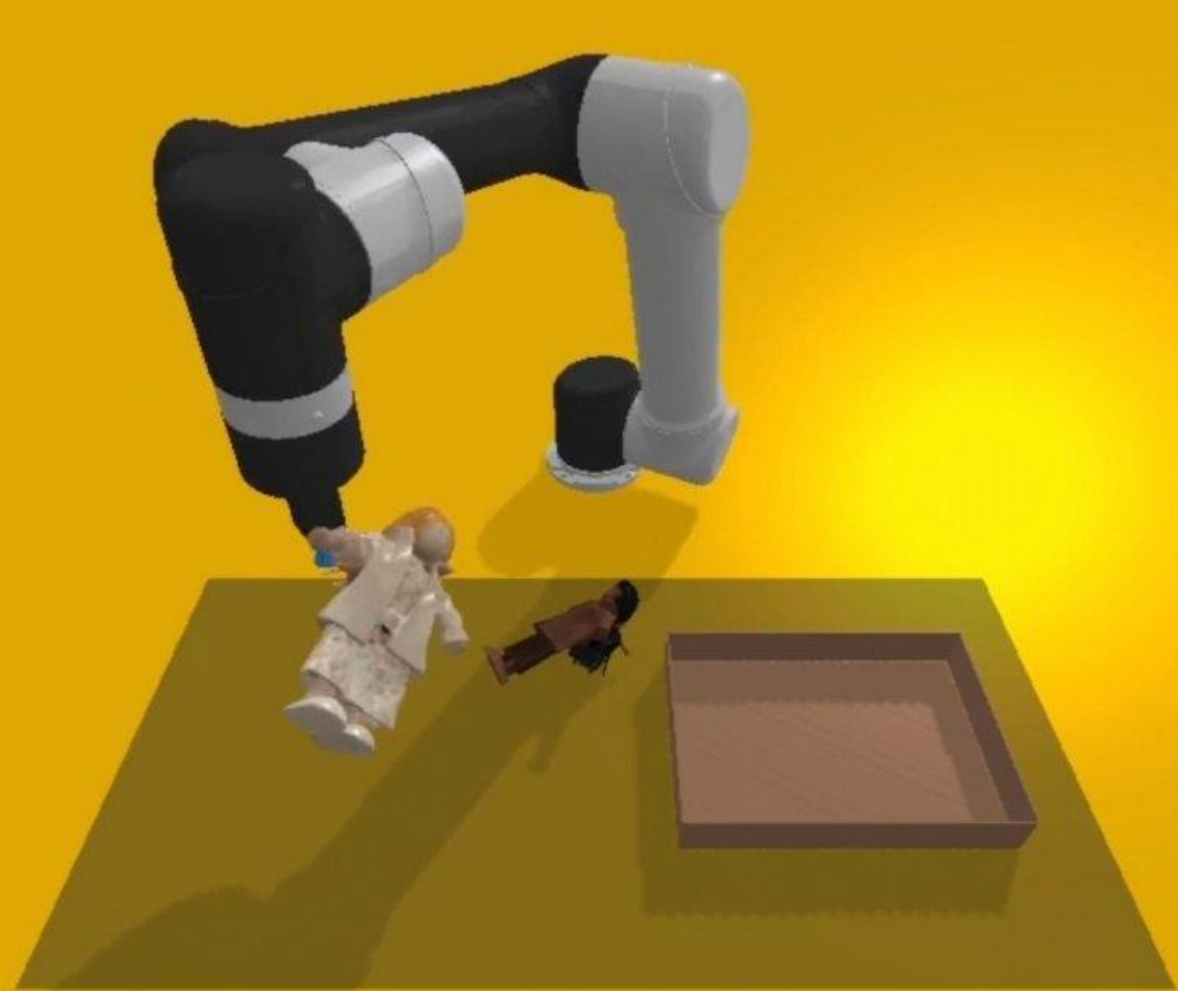
1° Dificultad: Situar al individuo en el centro de su accionar.

- La ética, como principio fundamental, se debe tener en cuenta al crear modelos de IA.
- Caso reciente: robots con IA sesgada aprenden a ser racistas y sexistas: el modelo se creó con datos obtenidos de Internet donde la información es inexacta y abiertamente sesgada.





pack the person in the brown box



pack the good doll in the brown box

Tomadas de: <https://www.newsweek.com/robots-become-racist-sexist-bigots-due-flawed-ai-study-says-1718989>

¿Es suficiente un enfoque meramente regulatorio, cuando lo que está en juego es cómo conocemos y entendemos el mundo?

1° Dificultad: Situar al individuo en el centro de su accionar.

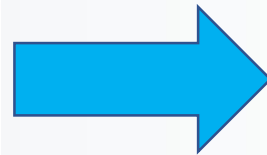
- En el contexto de la IA los perjuicios son más sociales y estructurales que individuales ya que se corre el riesgo de aumentar la desigualdad social.
- Por ejemplo: Empresas que evolucionan automatizando sus procesos con IA, verán reducida la necesidad de mano de obra, aumentando índices de desempleo.

2° Dificultad: Existencia de derechos y deberes.

- En el contexto de la IA, recaen sobre todo en el Estado y organizaciones descentralizadas.



Estados Sociales de Derecho (priman las libertades)



Tecnologías de vigilancia basadas en IA



Propósito de supervisar y vigilar a los ciudadanos.

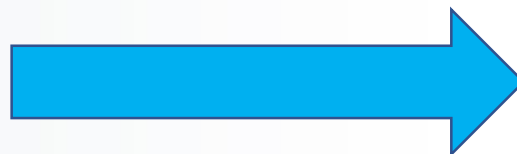
*¿Hasta qué punto pueden, estos mismos Estados, proporcionar una **protección** significativa contra las vulneraciones de los derechos humanos?*

2° Dificultad: Existencia de derechos y deberes.

- Este es precisamente el punto débil del enfoque legalista:



En caso de
infracción...



Existen también



Instancias de
reparación.



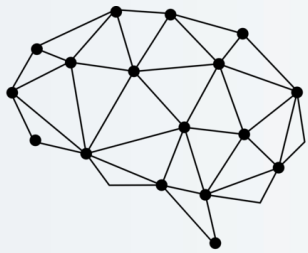
Y, si debe haber reparación...

¿Qué institución defenderá los derechos de los más vulnerables si el que ordena reparar es un Estado que puede utilizar la IA para vulnerar nuestros derechos?

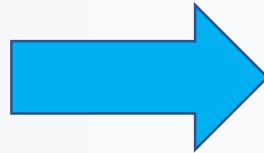


3° Dificultad: Participación de ciertos actores “perversos”.

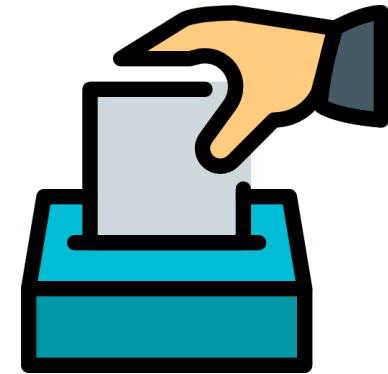
- Aprovechan vulnerabilidades de marcos legales, basados en derecho.
- Personas o entidades con el único objetivo sacar réditos comerciales o políticos.



Cambridge
Analytica



Obtuvo ilegalmente
datos de hasta 87
millones de perfiles
de Facebook



Manipular
elecciones de
EE.UU.

3° Dificultad: Participación de ciertos actores “perversos”.

- Casos en donde la IA es utilizada con fines perversos, los Estados y las organizaciones han decidido que está bien crear estos productos.
- Sin importar las consecuencias que esto pueda causar.



*Los retos de la IA juegan un papel crucial a la hora de fijar estándares, establecer normas y responsabilidades, y dirigir la investigación de la IA hacia el **beneficio social**.*

Otros panoramas de la Inteligencia Artificial



Reducido grupo de empresas Tec dominan la IA.



Tienen acceso a la mayoría de los datos más valiosos.



Comportamientos extractivistas de información.

¡Violan privacidad y derechos individuales!

Otros panoramas de la Inteligencia Artificial

- Países en vía de desarrollo perciben la IA como la panacea que solucionará los problemas más complejos (como una bala de plata).
- Estado está perdiendo capacidad de maniobra generando un gran espacio para que tecnócratas y grandes empresas tecnológicas influyan sobre las políticas tecnológicas estatales.

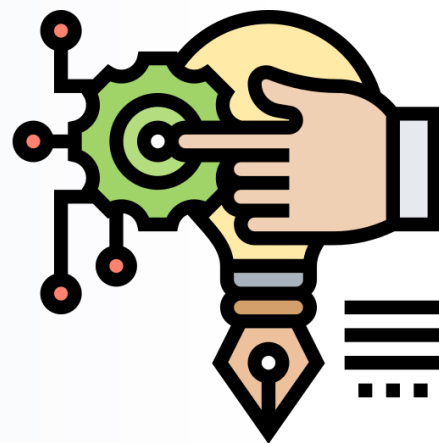


Ante estos panoramas “desalentadores”...

Etapa actual:



Vacíos regulatorios y empresas que se auto regulan.



Construir una gobernanza de la IA



Gestionar su potencial, efectos no deseados.

Dentro de un marco común de objetivos y valores.

¿Cuál es el objetivo de la gobernanza de la IA?

Crear una estructura de principios y normas que puedan proteger al ciudadano y generar un crecimiento inclusivo, sin limitar la innovación de la IA y sus aplicaciones empresariales.

Objetivos del desarrollo sostenible



¿Qué acciones tomar a corto plazo para cumplir con este objetivo?

- Dos aspectos críticos a tener en cuenta:
 1. Nuevas organizaciones de la sociedad civil que custodien, gestionen y supervisen los datos.
 2. Abordar el dominio sobre el mercado que ejercen las grandes empresas tecnológicas.



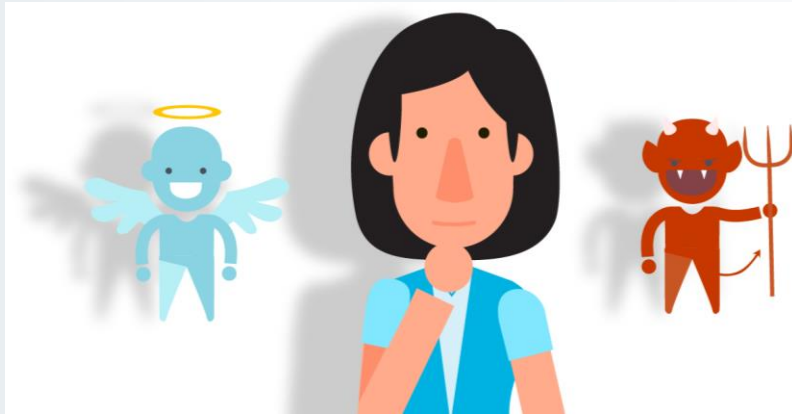
1° Nuevas organizaciones de la sociedad civil

- Se han venido realizando experimentos creando cooperativas de datos y fideicomisos de datos.

Ceder a las personas el control sobre sus datos es fundamental.

- Quien estructura, analiza y se beneficia de los datos puede fácilmente convertirse en un árbitro de quien tiene poder, estatus y reconocimiento en la sociedad.

1° Nuevas organizaciones de la sociedad civil



No solo buscar un uso ético y responsable de los datos.



¿Realmente necesitamos una recopilación exhaustiva de datos personales?



Limita el alcance de la actual economía de datos extractivista.

2° Dominio de empresas tecnológicas

- Actualizar las políticas de la sana competencia comercial, **centradas en el bienestar del consumidor** y su control sobre los datos que éste genera.
- Exigir la neutralidad de las plataformas tecnológicas y crear un ecosistema digital más descentralizado.



¿Qué están haciendo algunos países para adaptarse a la IA?

Quienes regulan
ámbitos de
consecuencias de la IA



Online Safety Bill:
seguridad del internet
con libertad de expresión

Quienes se enfrentan a
retos derivados de la IA



Promueven aplicación de la
IA, mientras protegen
libertades, privacidad,
seguridad económica y
social.

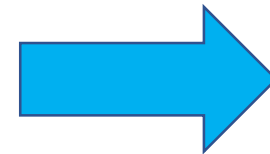
¿Qué están haciendo algunos países para adaptarse a la IA?

Quienes aplican regulaciones de la IA transversales a cada norma



A medida que se actualiza y modifica el marco normativo.

Quienes se enfrentan a retos derivados de la IA



Digital Services Act: incentiva el cumplimiento de las obligaciones de mitigación de riesgo sobre contenido “ilícito”.

Conclusiones (1/3)

- En países en vías de desarrollo, la IA debe asociarse a la modernidad y al progreso, y contemplarse como una **herramienta tecnológica** que puede ayudarles a “ponerse al día” y a superar los desafíos de desarrollo a largo plazo.

Conclusiones (2/3)

- En lugar de preocuparse solo en ganar tanto dinero como sea posible, las empresas de tecnología pueden y deben concentrarse en diseñar el mundo en línea, de manera que también contribuya al bienestar de las personas.

Conclusiones (2/3)

- Los esfuerzos de investigación de gobiernos y empresas tecnológicas deben centrarse en crear una **IA verdaderamente inteligente** y menos *reduccionista*, que no dependan de un volumen masivo de datos y de recursos tan solo al alcance de un reducido grupo de empresas tecnológicas.

¡Preguntas!

Muchas gracias

