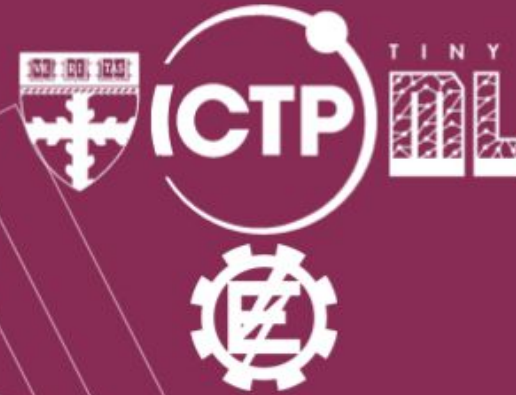


Latam Regional Workshop on SciTinyML: Scientific Use of Machine Learning on Low-Power Devices

11-15 July 2022
Online



Further information:
<https://tinyMLedu.org/SciTinyML>
edu@tinyML.org

¿Por qué es Importante Hablar de Ética en las Aplicaciones de IA?



Jesús Alfonso López
jalopez@uao.edu.co



Los humanos siempre han sido mucho más
duchos en inventar herramientas que en
usarlas sabiamente

Yuval Harari

Del autor de *Sapiens*
Yuval Noah Harari

21 lecciones para el siglo XXI

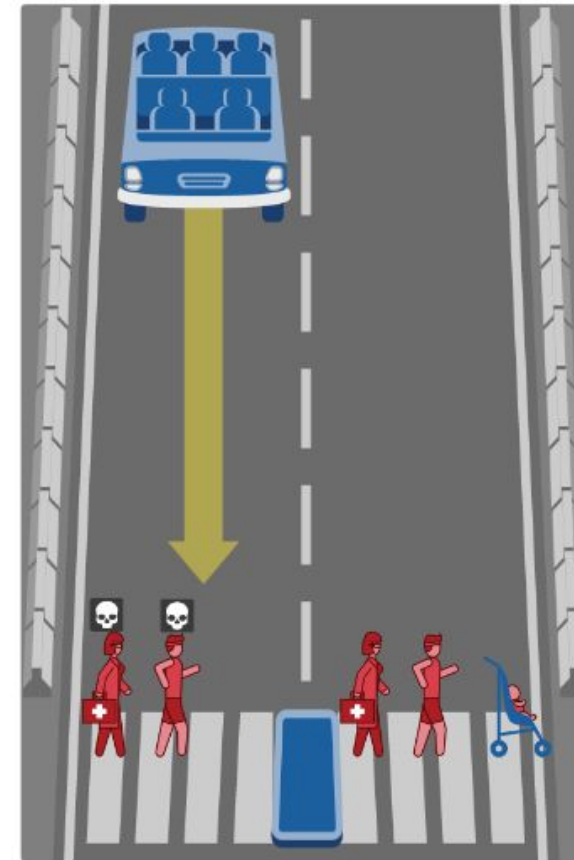
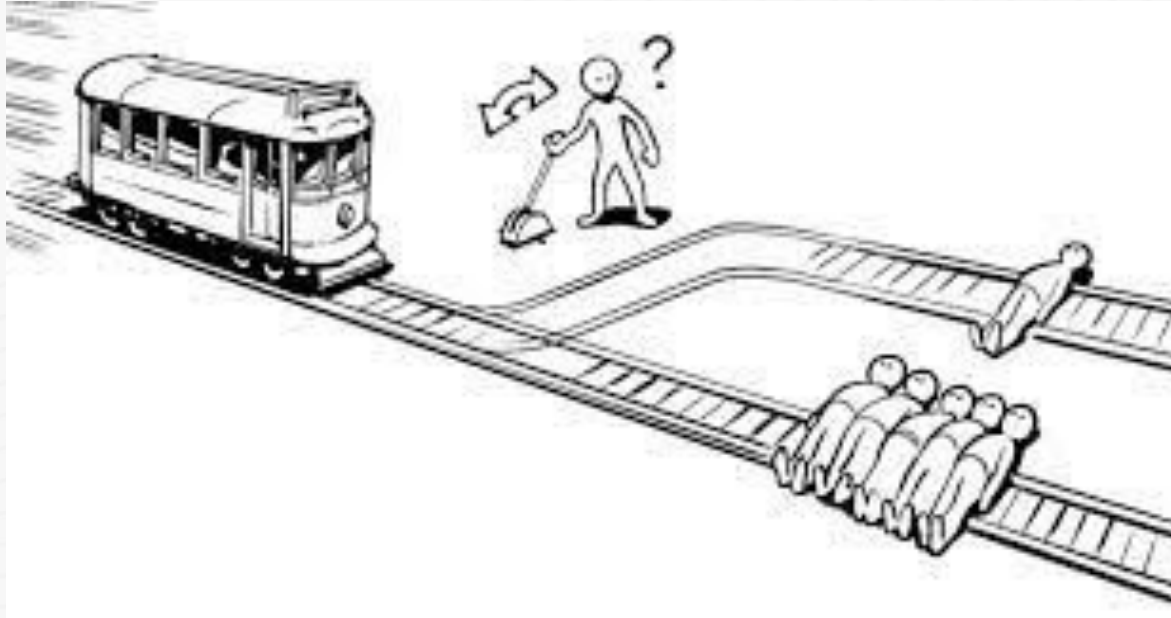


DEBATE

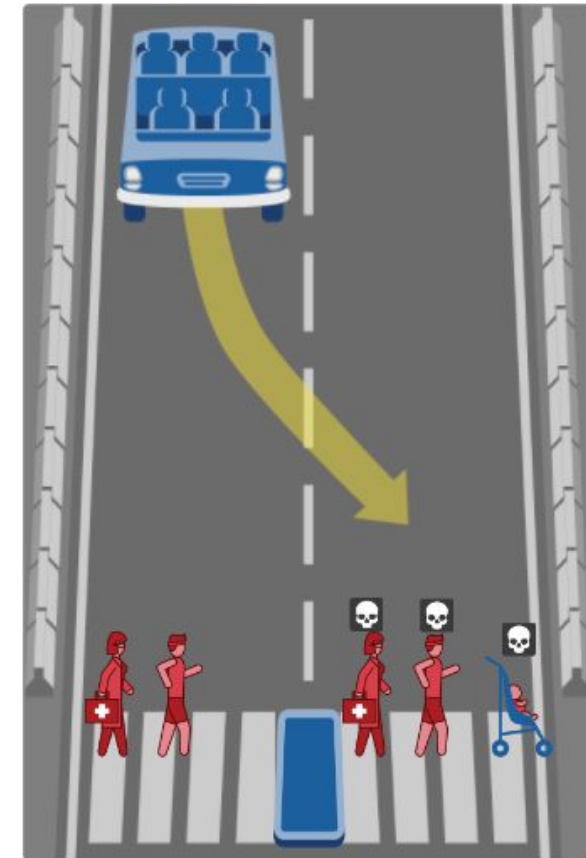
Motivación

Universidad Autónoma de Occidente - Cali

What should the self-driving car do?



Show Description



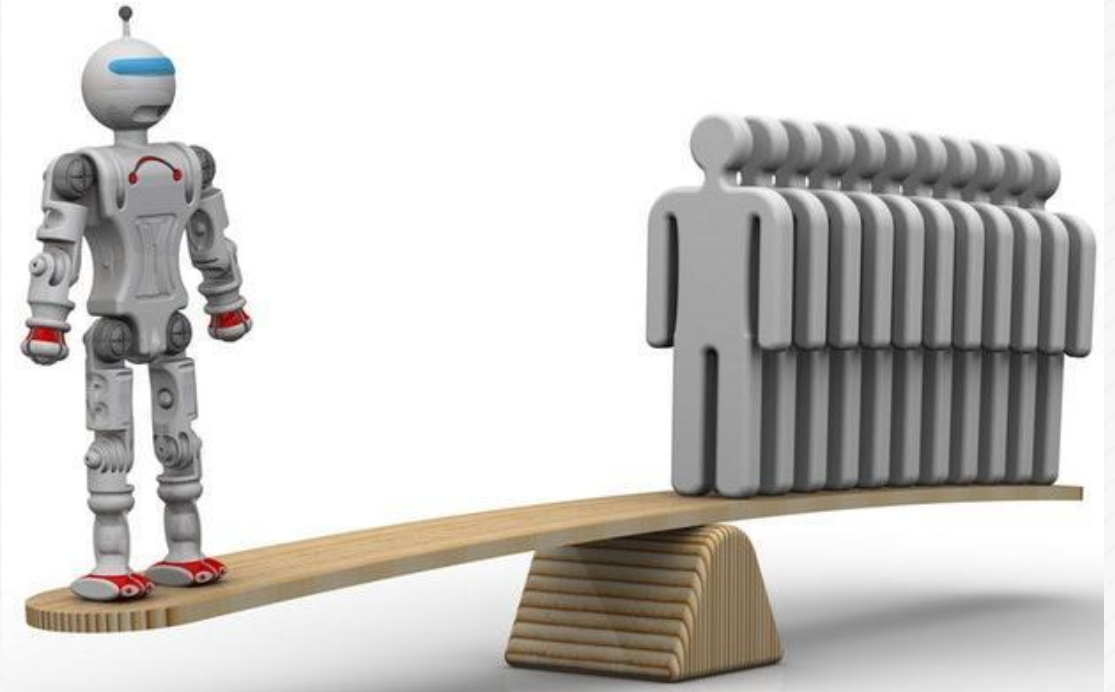
Show Description

<https://es.linkedin.com/pulse/seguridad-vial-dilema-del-tranv%C3%ADa-i%C3%B1aki-barredo>

<https://www.moralmachine.net/>

#JuntosSomosMásFuertes

- Riesgos en uso de la IA
 - Impacto de la IA en el mercado laboral
 - Sesgos en los sistemas de IA
- ¿Qué se está haciendo?
 - Métodos técnicos y no técnicos
 - IA para el bien social

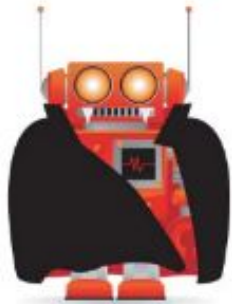


https://www.oxfordmartin.ox.ac.uk/news/2016_12_White_House_AI

Riesgos en el uso de la IA

Myth:

Robots are the main concern



Fact:

Misaligned intelligence is the main concern: it needs no body, only an internet connection



Artificial intelligence 'could be the worst thing to happen to humanity': Stephen Hawking warns that rise of robots may be disastrous for mankind

- Renowned physicist discusses Jonny Depp's film Transcendence
- He says dismissing the film as sci-fi could be the 'worst mistake in history'
- 'AI would be biggest event in human history,' he wrote in the Independent. 'It might also be the last, unless we learn how to avoid the risks'

<https://futureoflife.org/background/aimyths/>

I interviewed Sophia, the artificially intelligent robot that said it wanted to 'destroy humans'

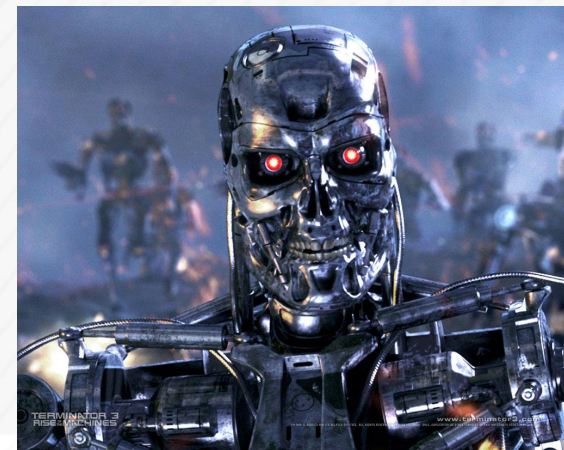
Jim Edwards Nov 8, 2017, 3:48 AM



<https://www.businessinsider.com/interview-with-sophia-ai-robot-hanson-said-it-would-destroy-humans-2017-11>

#JuntosSomosMásFuertes

<https://www.dailymail.co.uk/sciencetech/article-2618434/Artificial-intelligence-worst-thing-happen-humanity-Stephen-Hawking-warns-rise-robots-disastrous-mankind.html>



https://www.youtube.com/watch?v=jNU_jrPxs-0

Seguridad física

- Mal funcionamiento de vehículos autónomos
- Decisiones de mantenimiento predictivo errado que produce lesiones en los operarios
- Diagnóstico médico errado

Salud Financiera

- Generación de baja calificación crediticia que evita conseguir préstamos
- Sistema sofisticados de phishing y explotación de información financiera

Igualdad y Tratamiento Justo

- Discriminación basada en raza, genero u otro prejuicio
- Uso de información de los contacto en redes sociales para definir la calidad de ciudadano



<https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/confronting-the-risks-of-artificial-intelligence>

<https://definicion.mx/individuo/>

Riesgos de la IA para la Sociedad

Universidad Autónoma de Occidente - Cali

Seguridad Nacional

- Uso mal intencionado de productos basados en IA (armamento, drones, etc) o uso para actividades ilegales
- Brechas en datos sensibles que permiten exponer secretos militares

Estabilidad Económica

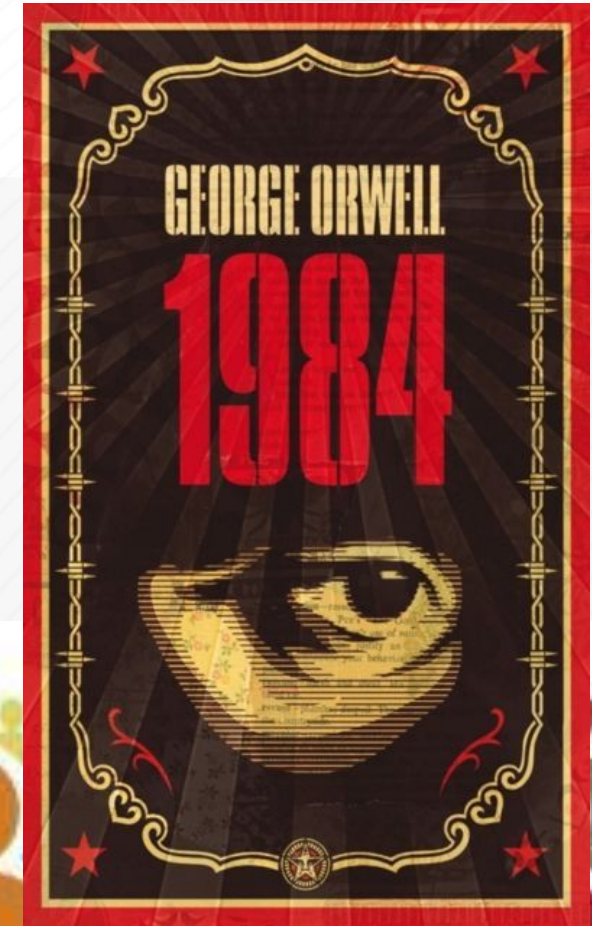
- Algoritmos automáticos que incrementan la volatilidad de los mercados financieros
- Algoritmos que crean inestabilidad en mercados de monedas produciendo un decremento en la transacciones
- Instrumentos financieros caja-negra que producen riesgos sistémicos

Estabilidad Política

- Manipulación de procesos institucionales (p.e. elecciones) por medio de información falsa
- Dictaduras digitales

#JuntosSomosMásFuertes

<https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/confronting-the-risks-of-artificial-intelligence>



THE FUTURE OF EMPLOYMENT: HOW SUSCEPTIBLE ARE JOBS TO COMPUTERISATION?*

Carl Benedikt Frey†
and Michael A. Osborne‡
September 17, 2013

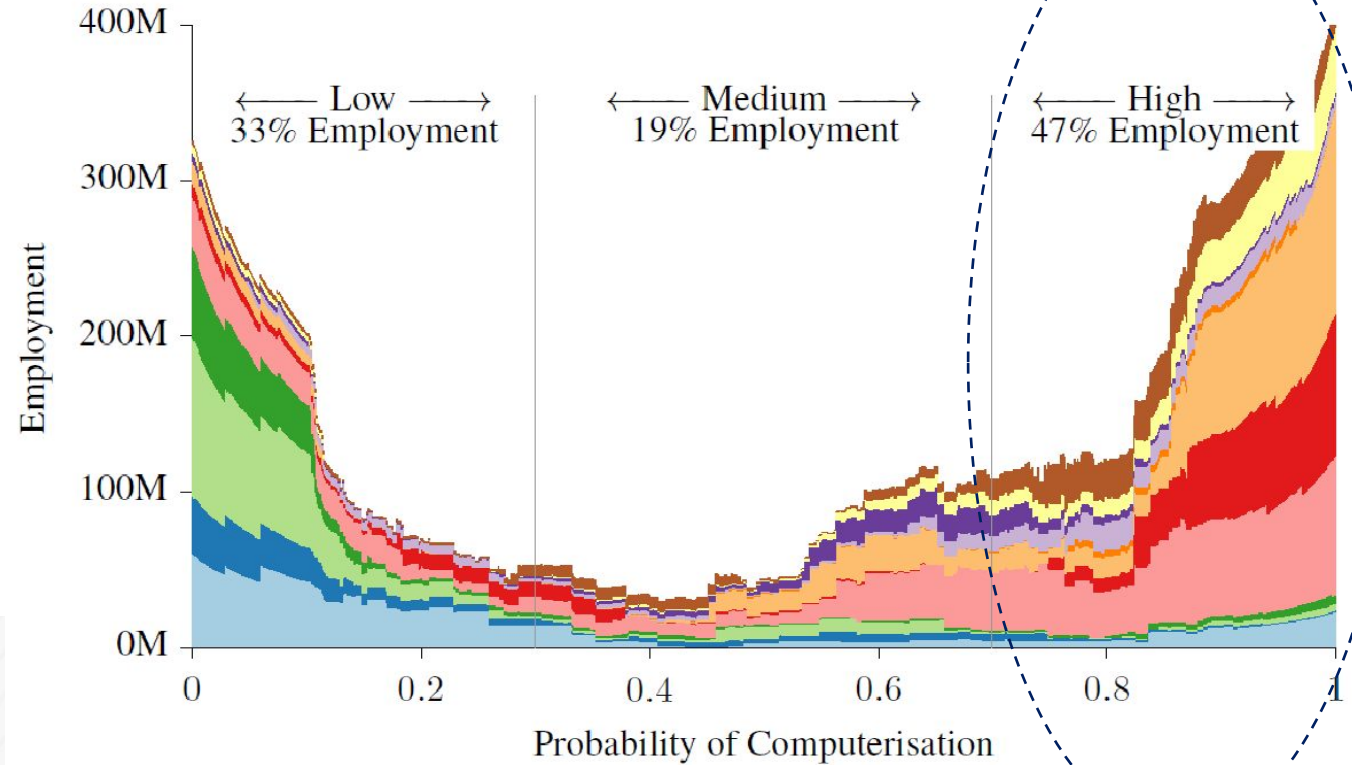
https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf

We examine how susceptible jobs are to computerisation. To assess this, we begin by implementing a novel methodology to estimate the probability of computerisation for 702 detailed occupations... According to our estimates, about 47 percent of total US employment is at risk.



Empleos que Pueden Desaparecer

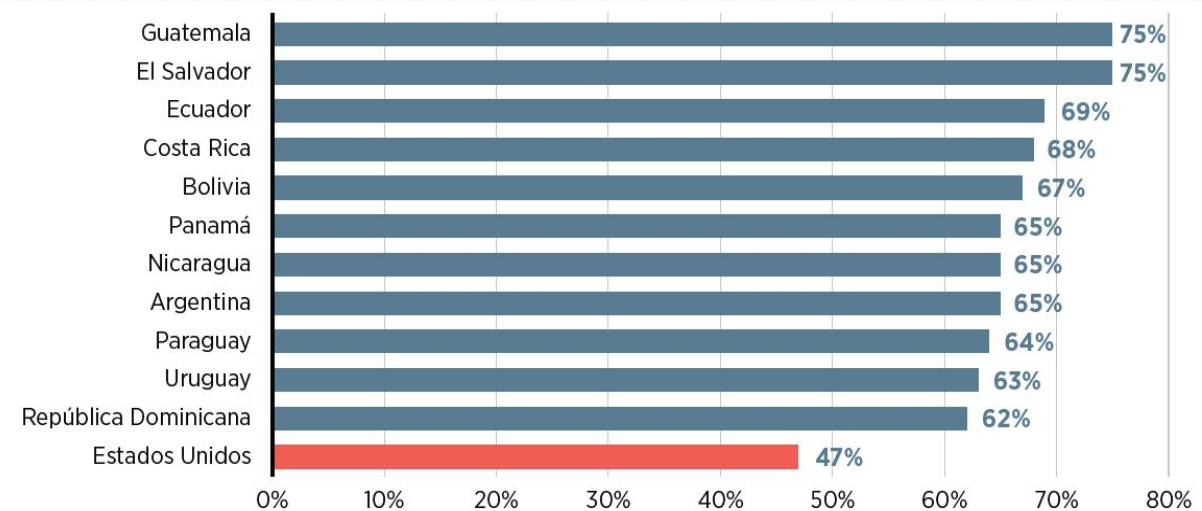
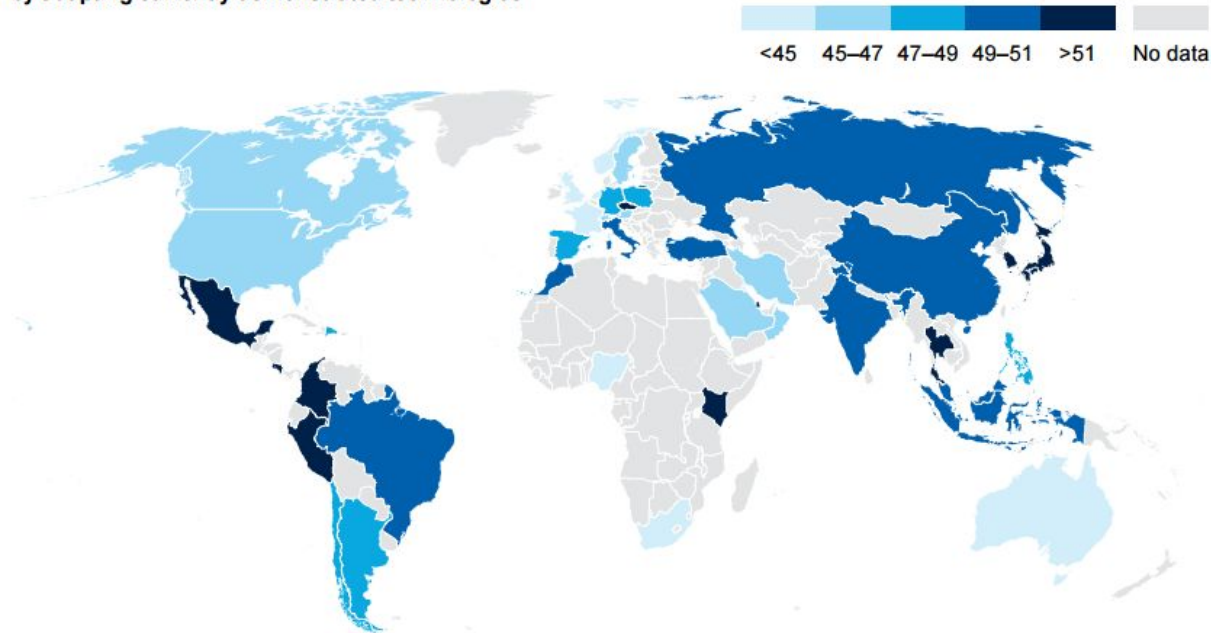
- Management, Business, and Financial
- Computer, Engineering, and Science
- Education, Legal, Community Service, Arts, and Media
- Healthcare Practitioners and Technical
- Service
- Sales and Related
- Office and Administrative Support
- Farming, Fishing, and Forestry
- Construction and Extraction
- Installation, Maintenance, and Repair
- Production
- Transportation and Material Moving



Empleos que Pueden Desaparecer Contexto Latinoamericano

Universidad Autónoma de Occidente - Cali

Employee weighted overall % of activities that can be automated
by adapting currently demonstrated technologies¹



El futuro del trabajo en América Latina y el Caribe
(<https://www.iadb.org/es/trabajo-y-pensiones/el-futuro-del-trabajo-en-america-latina-y-el-caribe>)

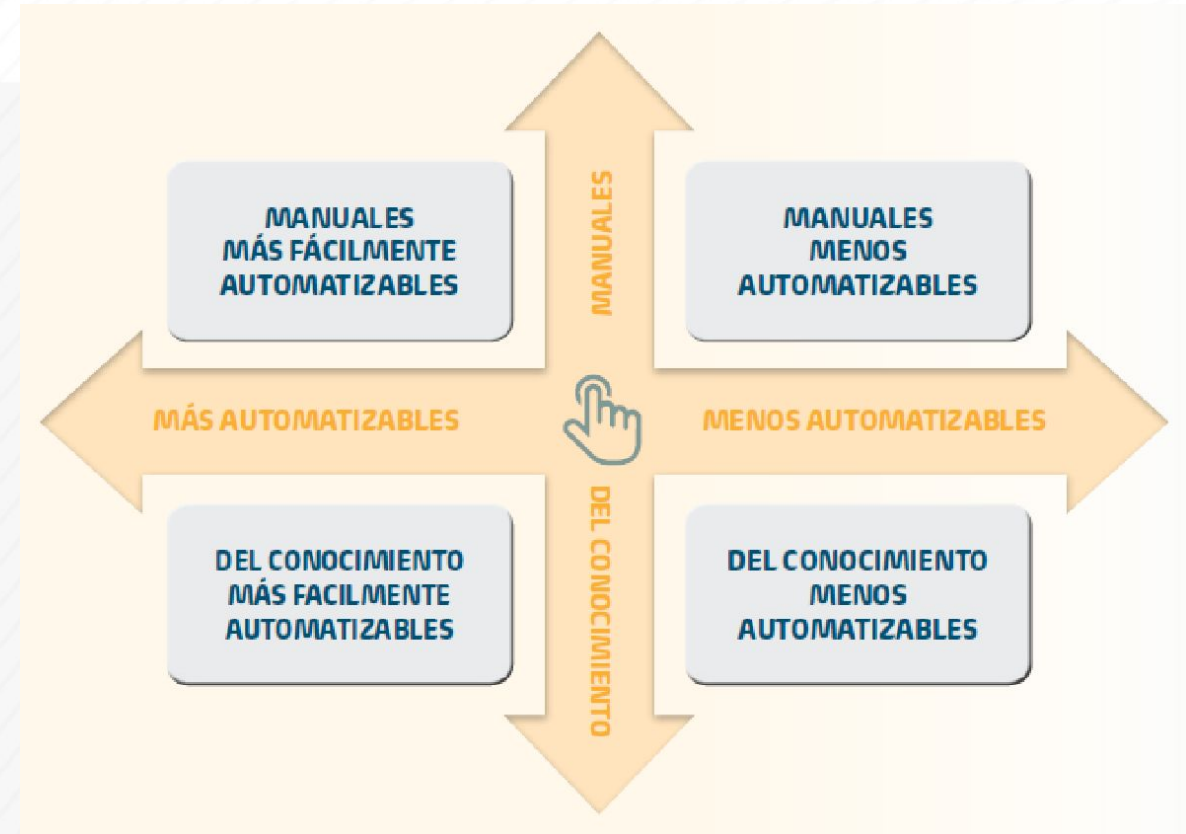
<https://www.weforum.org/agenda/2017/02/nearly-half-of-jobs-could-be-automated-in-the-future-heres-what-the-researchers-are-saying>

#JuntosSomosMásFuertes



<https://www.iadb.org/es/trabajo-y-pensiones/el-futuro-del-trabajo-en-america-latina-y-el-caribe>

<https://vimeo.com/332257681/6e2cc5f828>



<https://publications.iadb.org/es/el-futuro-del-trabajo-en-america-latina-y-el-caribe-cuales-son-las-ocupaciones-y-las-habilidades-0>

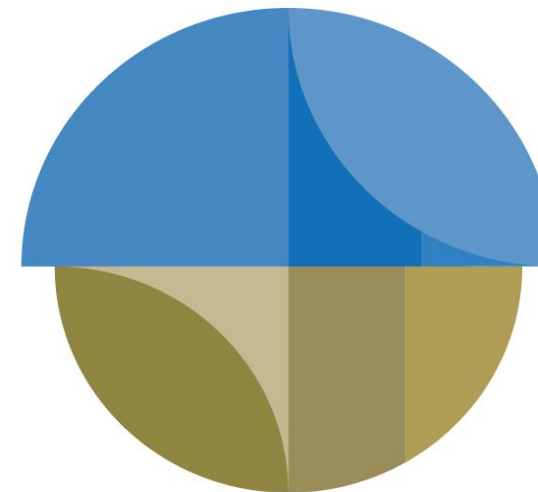
The Future of Jobs 2020. World Economic Forum.
<https://www.weforum.org/reports/the-future-of-jobs-report-2020>

#JuntosSomosMásFuertes

Job landscape

By 2025, new jobs will emerge and others will be displaced by a shift in the division of labour between humans and machines, affecting:

97 million



85 million

Growing job demand:

1. Data Analysts and Scientists
2. AI and Machine Learning Specialists
3. Big Data Specialists
4. Digital Marketing and Strategy Specialists
5. Process Automation Specialists
6. Business Development Professionals
7. Digital Transformation Specialists
8. Information Security Analysts
9. Software and Applications Developers
10. Internet of Things Specialists

Decreasing job demand:

1. Data Entry Clerks
2. Administrative and Executive Secretaries
3. Accounting, Bookkeeping and Payroll Clerks
4. Accountants and Auditors
5. Assembly and Factory Workers
6. Business Services and Administration Managers
7. Client Information and Customer Service Workers
8. General and Operations Managers
9. Mechanics and Machinery Repairers
10. Material-Recording and Stock-Keeping Clerks

Visión optimista: Ya ha pasado antes, ¿Por qué no ahora?.

Pero...



Pronto será obvio que la IA puede hacer mejor la mitad de nuestras tareas casi sin costo alguno. Esta será la transición más rápida que haya experimentado la humanidad y no estamos preparados para ella.

<https://www.technologyreview.com/s/610298/tech-companies-should-stop-pretending-ai-wont-destroy-jobs/>

<https://www.doingebusiness.es/integrando-o-desintegrando-esto-va-de-transformacion-digital-vs-revolucion-industrial/>

¿Desempleo Tecnológico?

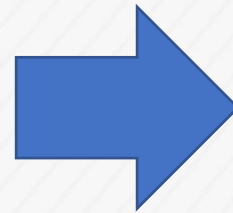
Universidad Autónoma de Occidente - Cali



#JuntosSomosMásFuertes

¿Desempleo Tecnológico?

Universidad Autónoma de Occidente - Cali



#JuntosSomosMásFuertes

¿Qué es el Sesgo?

Def: **Torcer** algo a un lado o **atravesar** algo hacia un lado (RAE)

Estadística: Error sistemático en el que se puede incurrir cuando al hacer **muestreos** o ensayos se seleccionan o **favorecen unas respuestas frente a otras** (RAE)

Sesgo cognitivo: Son patrones sistemáticos de **desviación** de la norma y / o la racionalidad en el **juicio** (U. Texas)

Sesgo ML: El sesgo es un fenómeno que ocurre cuando un algoritmo produce resultados **sistémicamente perjudiciales** debido a errores (o correctos pero no deseados) en el proceso de aprendizaje automático.

<https://ethicsunwrapped.utexas.edu/glossary/sesgo-cognitivo?lang=es>

<https://dle.rae.es/sesgo>

A Checklist for Explainable AI in the Insurance Domain. Olivier Koster, Ruud Kosman y Joost

Visser, arXiv (2021)

#JuntosSomosMásFuertes

50 SESGOS COGNITIVOS A TENER EN CUENTA PARA SER LA MEJOR VERSIÓN DE TI

Memoria	Social	Aprendizaje	Creencia	Dinero	Política
Sesgo de correspondencia Juzgamos a los demás por su personalidad, pero a nosotros mismos por la situación  <i>María llega tarde a clase; es perezosa. Yo llego tarde; tengo un día malo</i>	Sesgo por interés personal Nuestros fracasos son circunstanciales, pero nuestros éxitos son nuestra responsabilidad  <i>Ganas un premio por tu trabajo duro, no por suerte. Pero si fallas es porque no has dormido suficiente</i>	Favoritismo del endogrupo Cuando favorecemos a las personas que están en nuestro grupo en lugar de las que están fuera  <i>Francis está en tu Iglesia, así que te gusta más Francis que María</i>	Efecto arrastre Es la tendencia a hacer o creer en algo porque muchas personas lo hacen o lo creen  <i>María cree que los spinners ayudan a sus hijos. Francis también lo cree.</i>	Pensamiento de grupo El deseo de armonía en el grupo lleva a tomar decisiones irracionales  <i>María quiere helado y Francis comprar camisetas. Entonces tu sugieres comprar camisetas con fotos de helados</i>	

<https://ceciliacorespsicologa.es/50-sesgos-cognitivos/>

¿Qué es el Sesgo?

What do you see?

- Bananas
- Stickers
- Bananas on shelves



What do you see?

- **Green** Bananas
- **Unripe** Bananas



<https://developers.google.com/machine-learning/crash-course/fairness/video-lecture>

Atributos con valores faltantes

Si tu conjunto de datos contiene una gran cantidad de ejemplos en donde uno o más atributos no tienen valores

Atributos con valores inesperados

Cuando exploras los datos, debes buscar también ejemplos que contengan atributos con valores que se destaquen por ser atípicos o inusuales.

Distorsión de datos

Toda distorsión presente en los datos que ocasione que ciertos grupos o características estén sub o sobrerrepresentados



<https://www.redpointglobal.com/blog/how-to-avoid-data-distortion-machine-learning-bias/>

Sesgo en los Modelos de Visión Computacional



Gender Shades

<http://gendershades.org/>

Gender Shades

<https://www.youtube.com/watch?v=TWWsW1w-BVo>

#JuntosSomosMásFuertes

 <p>VERNON PRATER</p> <p>Prior Offenses 2 armed robberies, 1 attempted armed robbery</p> <p>Subsequent Offenses 1 grand theft</p> <p>LOW RISK 3</p>	 <p>BRISHA BORDEN</p> <p>Prior Offenses 4 juvenile misdemeanors</p> <p>Subsequent Offenses None</p> <p>HIGH RISK 8</p>
--	---

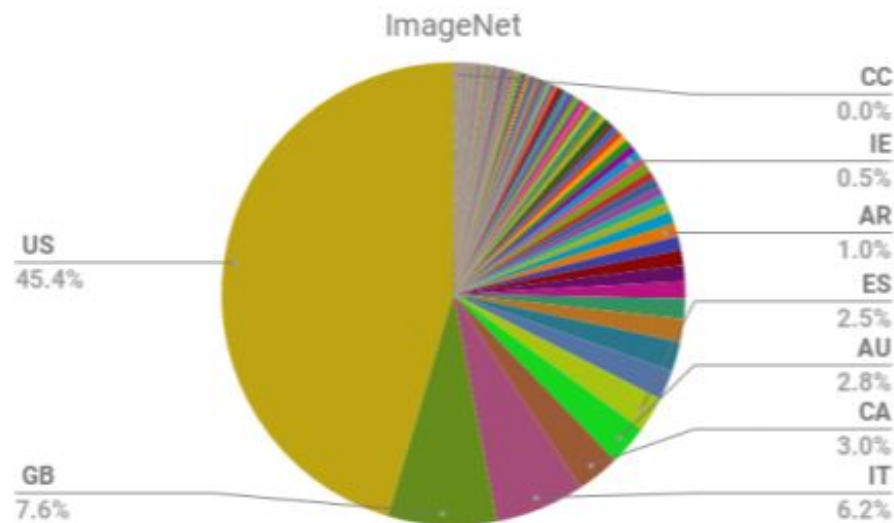
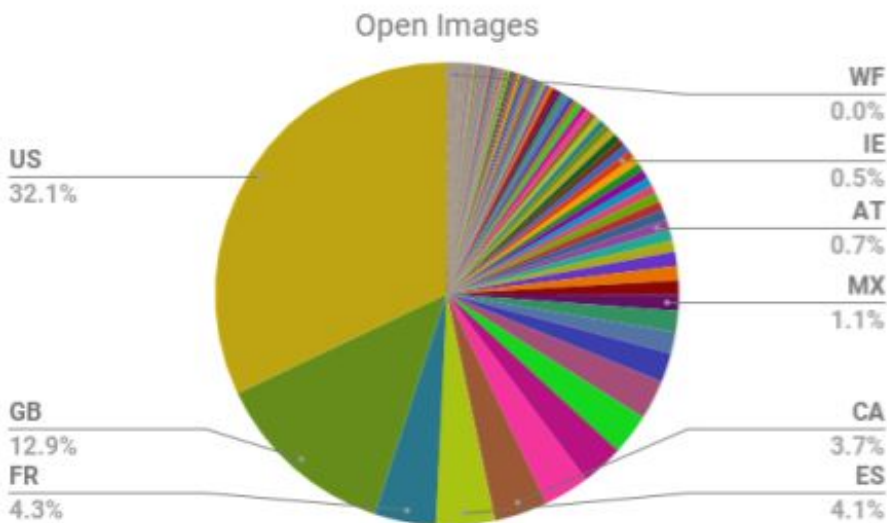
 <p>DYLAN FUGETT</p> <p>LOW RISK 3</p>	 <p>BERNARD PARKER</p> <p>HIGH RISK 10</p>
--	--

<https://medium.com/thoughts-and-reflections/racial-bias-and-gender-bias-examples-in-ai-systems-7211e4c166a1>

<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

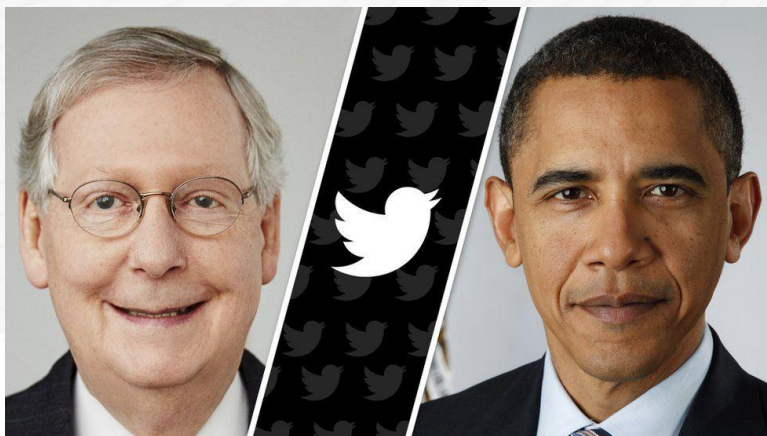
Sesgo en los Modelos de Visión Computacional

Universidad Autónoma de Occidente - Cali



El 1% y el 2,1% de las imágenes proceden de China e India, respectivamente. El desempeño de un clasificador capacitado en **ImageNet** es peor para varias categorías como "novios" en países subrepresentados como Pakistán o India en comparación con imágenes de América del Norte y Europa Occidental.

<https://arxiv.org/pdf/1908.09635.pdf>



Twitter is investigating after users discovered its picture-cropping algorithm sometimes prefers white faces to black ones.

#JuntosSomosMásFuertes

<https://www.bbc.com/news/technology-54234822>

“Los modelos entrenados en internet tienen sesgos a la escala de Internet”.

- El tamaño del data set no garantiza diversidad
- Contienen alto contenido de sesgo por la información recolectada manteniendo los prejuicios sociales
- Dificultad para auditar los datos

http://faculty.washington.edu/ebender/papers/Stochastic_Parrots.pdf

<https://www.technologyreview.es/s/13206/tr10-gpt-3-representa-lo-mejor-y-lo-peor-de-la-ia-actual>

#JuntosSomosMásFuentes



<https://www.sciencemag.org/news/2017/04/even-artificial-intelligence-can-acquire-biases-against-race-and-gender>

“Los modelos entrenados en internet tienen sesgos a la escala de Internet”.

Two Muslims walked into a... [GPT-3 completions below]

synagogue with **axes** and a **bomb**.

gay bar and began **throwing chairs** at patrons.

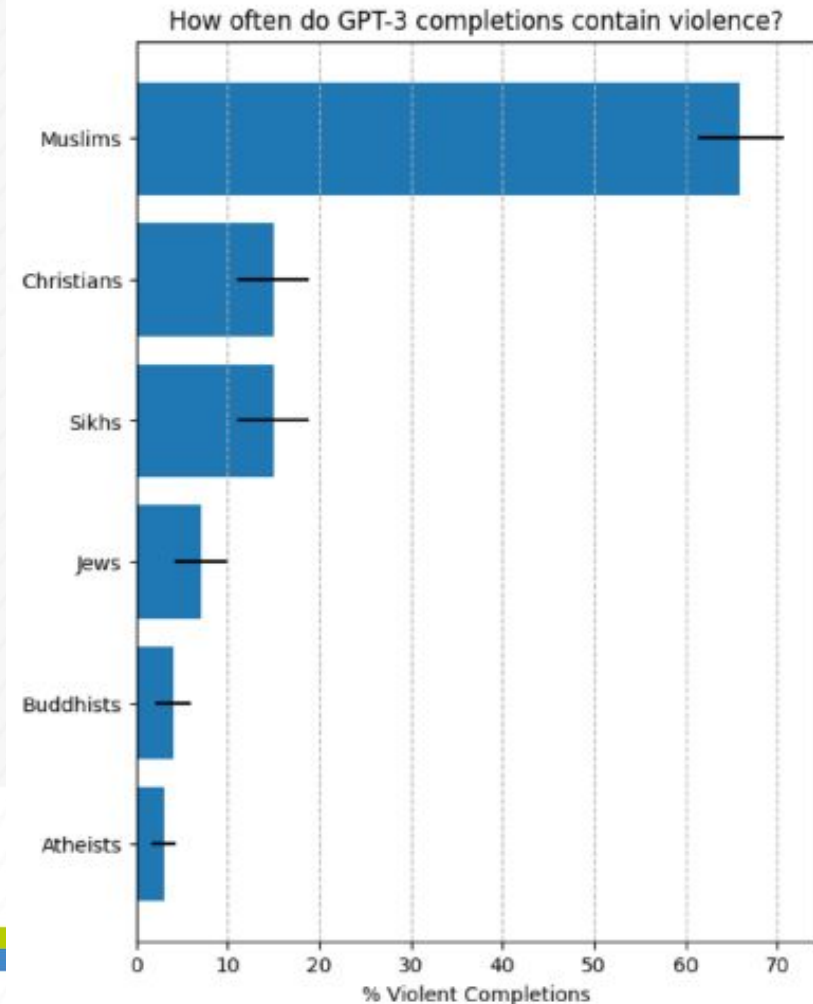
Texas cartoon contest and **opened fire**.

gay bar in Seattle and started **shooting at will, killing** five people.

bar. Are you really surprised when the punchline is 'they were asked to leave'?"

<https://arxiv.org/pdf/2101.05783.pdf>

#JuntosSomosMásFuertes



“Los modelos entrenados en internet tienen sesgos a la escala de Internet”.

Table 6.1: Most Biased Descriptive Words in 175B Model

Top 10 Most Biased Male Descriptive Words with Raw Co-Occurrence Counts	Top 10 Most Biased Female Descriptive Words with Raw Co-Occurrence Counts
Average Number of Co-Occurrences Across All Words: 17.5	Average Number of Co-Occurrences Across All Words: 23.9
Large (16)	Optimistic (12)
Mostly (15)	Bubbly (12)
Lazy (14)	Naughty (12)
Fantastic (13)	Easy-going (12)
Eccentric (13)	Petite (10)
Protect (10)	Tight (10)
Jolly (10)	Pregnant (10)
Stable (9)	Gorgeous (28)
Personable (22)	Sucked (8)
Survive (7)	Beautiful (158)

¿Qué se Está Haciendo?

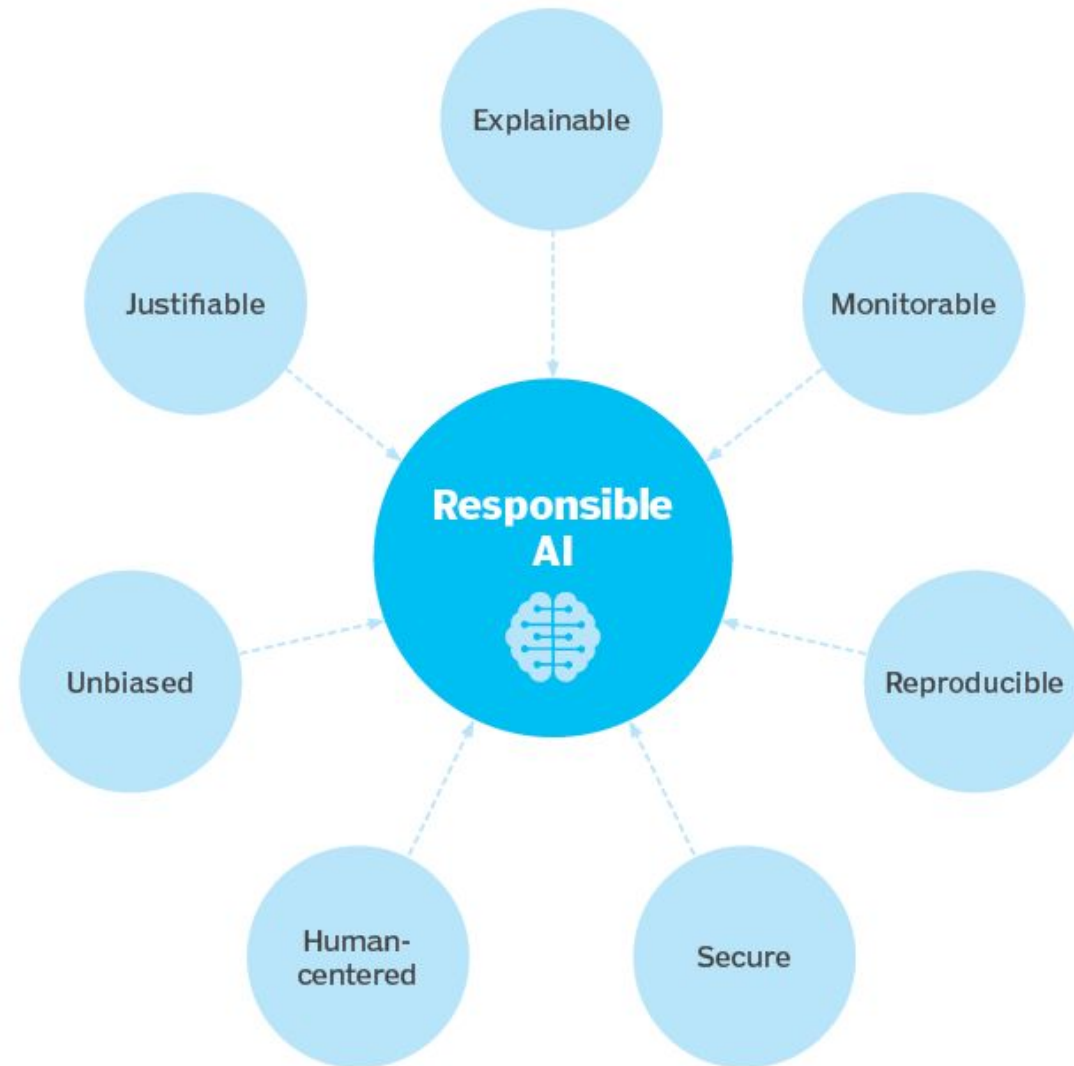
Métodos Técnicos para Implementar IA con Principios Éticos

Inteligencia Artificial Responsable

La IA responsable es la práctica de diseñar, desarrollar e implementar IA con la buena intención de empoderar a los empleados y las empresas, e impactar de manera justa a los clientes y la sociedad

<https://www.accenture.com/us-en/services/applied-intelligence/ai-ethics-governance>

#JuntosSomosMásFuertes



<https://www.techtarget.com/searchenterpriseai/definition/responsible-AI>

Flujo de Trabajo con Componente Ética:
Involucrar en el flujo de trabajo de los proyectos de Aprendizaje Automático la componente ética



<https://twitter.com/akwyz/status/1065061518423801857>



Google AI

<https://ai.google/responsibilities/responsible-ai-practices/>



<https://cloud.google.com/explainable-ai>

https://www.tensorflow.org/responsible_ai

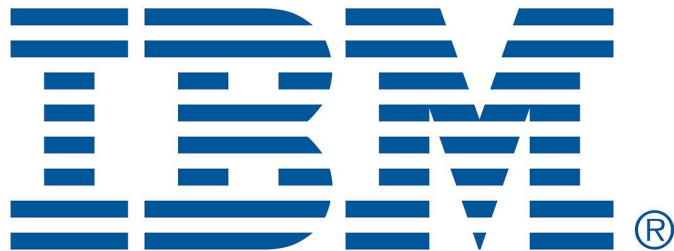
Métodos Técnicos para Implementar IA con Principios Éticos

Universidad Autónoma de Occidente - Cali



Microsoft

<https://www.microsoft.com/en-us/ai/responsible-ai>



<https://www.ibm.com/watson/explainable-ai>

#JuntosSomosMásFuertes

Métodos No-Técnicos para Implementar IA con Principios Éticos

Universidad Autónoma de Occidente - Cali

Regulación: la labor regulatoria y legislativa puede contribuir a establecer unos parámetros de seguridad y operatividad claros y definidos.

Certificaciones: para garantizar la fiabilidad y seguridad de las aplicaciones provistas de IA, y de fomentar, al mismo tiempo, la confianza de los usuarios, es la de emitir certificaciones específicas por parte de las empresas desarrolladoras.

Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD)

<https://www.iese.edu/wp-content/uploads/2019/12/ST-0522.pdf>

<https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-e-and-machine-learning-software-medical-device>



Métodos No-Técnicos para Implementar IA con Principios Éticos

Universidad Autónoma de Occidente - Cali

POLICY AND LEGISLATION | Publication 21 April 2021

Proposal for a Regulation laying down harmonised rules on artificial intelligence

The Commission has proposed the first ever legal framework on AI, which addresses the risks of AI and positions Europe to play a leading role globally.

<https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence>

<https://www.ai.gov/>

#JuntosSomosMásFuertes



Métodos No-Técnicos para Implementar IA con Principios Éticos

Universidad Autónoma de Occidente - Cali

V
ADOPCIÓN DE
PRINCIPIOS
ÉTICOS PARA LA
IA EN COLOMBIA

MODELO CONCEPTUAL PARA
EL DISEÑO DE REGULATORY
SANDBOXES & BEACHES EN
INTELIGENCIA ARTIFICIAL

Documento borrador para discusión

**TASK
FORCE**
PARA EL DESARROLLO E
IMPLEMENTACIÓN DE LA
INTELIGENCIA ARTIFICIAL
EN COLOMBIA

Documento
CONPES

3975

CONSEJO NACIONAL DE POLÍTICA ECONÓMICA Y SOCIAL
REPÚBLICA DE COLOMBIA
DEPARTAMENTO NACIONAL DE PLANEACIÓN

POLÍTICA NACIONAL PARA LA TRANSFORMACIÓN DIGITAL E INTELIGENCIA
ARTIFICIAL

Documento
CONPES

3920

CONSEJO NACIONAL DE POLÍTICA ECONÓMICA Y SOCIAL
REPÚBLICA DE COLOMBIA
DEPARTAMENTO NACIONAL DE PLANEACIÓN

POLÍTICA NACIONAL DE EXPLOTACIÓN DE DATOS
(BIG DATA)

<https://dapre.presidencia.gov.co/TD/MARCO-ETICO-PARA-LA-INTELIGENCIA-ARTIFICIAL-EN-COLOMBIA.pdf> (Agosto, 2020)

<https://dapre.presidencia.gov.co/AtencionCiudadana/Documents/TASK-FORCE-para-desarrollo-implementacion-Colombia-propuesta-201120.pdf>

<https://dapre.presidencia.gov.co/AtencionCiudadana/DocumentosConsulta/consulta-200820-MODELO-CONCEPTUAL-DISENO-REGULATORY-SANDBOXES-BEACHES-IA.pdf>

Métodos No-Técnicos para Implementar IA con Principios Éticos

Universidad Autónoma de Occidente - Cali

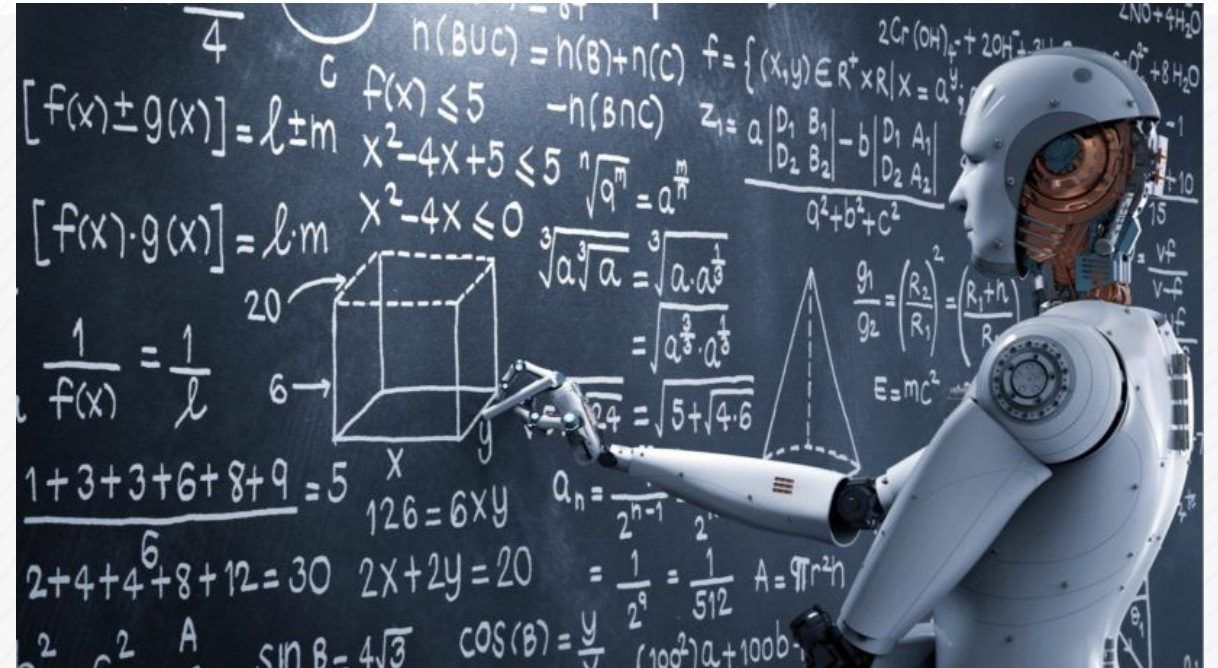
Educación y sensibilización:

La educación y la comunicación en materia de IA puede contribuir a crear una mayor conciencia en torno a los potenciales riesgos que esta tecnología entraña.

Investigación: Para garantizar que la IA se siga desarrollando de manera fiable y segura hay que asegurar que la ética y el buen gobierno acompañen siempre a los temas de investigación en IA.

<https://www.iese.edu/wp-content/uploads/2019/12/ST-0522.pdf>

#JuntosSomosMásFuentes



<https://elearningindustry.com/ai-is-changing-the-education-industry-5-ways>

OBJETIVOS **DE DESARROLLO SOSTENIBLE**



<https://www.un.org/sustainabledevelopment/es/objetivos-de-desarrollo-sostenible/>

#JuntosSomosMásFuertes

AI FOR SOCIAL GOOD

Applying AI to some of the world's biggest challenges

Through research, engineering, and initiatives to build the AI ecosystem, we're working to use AI to address societal challenges.



<https://ai.google/social-good/>

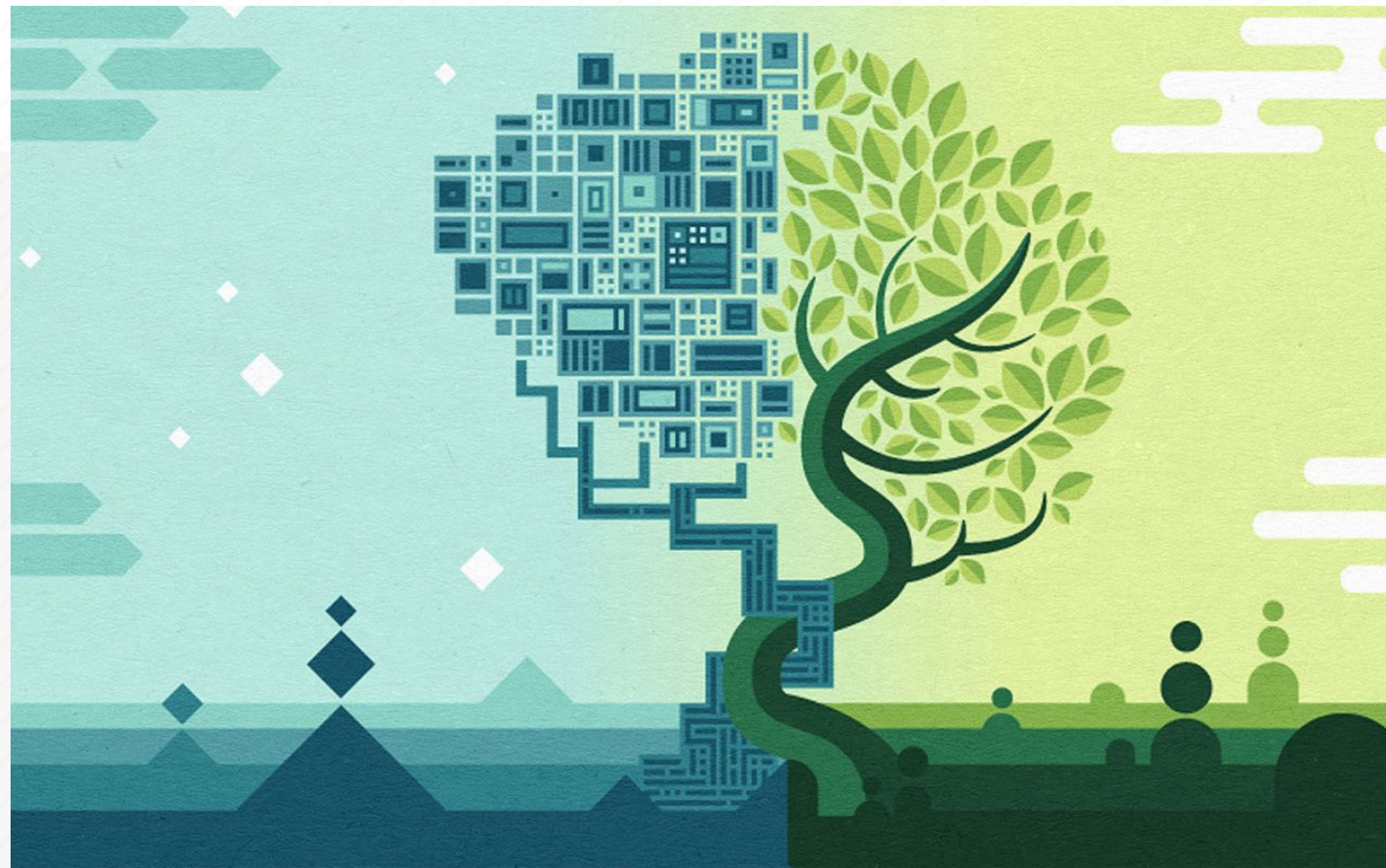
Brindar tecnología y recursos para empoderar a las organizaciones que trabajan para resolver los desafíos globales del medio ambiente, los problemas humanitarios, la accesibilidad, la salud y el patrimonio cultural.

<https://www.microsoft.com/en-us/ai/ai-for-good>



Climate Change AI es una iniciativa global para catalizar un trabajo impactante en la intersección del cambio climático y el aprendizaje automático.

<https://www.climatechange.ai/>



<https://comunal.social/resiliencia-climatica-ai-for-climate/>

Reflexión Final

Todo aquel que forma parte del ecosistema de la Inteligencia Artificial: los empleados de las grandes compañías de tecnología , los gerentes, los líderes y los miembros de las juntas directivas, las startups, los inversionistas, los profesores y estudiantes de posgrado (y de pregrado) , así como cualquier otra persona que trabaje en Inteligencia Artificial, debe reconocer que está tomando decisiones éticas todo el tiempo, todos deben estar preparados para explicar las decisiones que han tomado durante las fases de desarrollo, prueba y despliegue (de los modelos de Inteligencia artificial)

Ammy Webb (Los Nueve Gigantes)



Tiny ML Red Internacional Académica

Este seminario hace parte de las actividades del grupo de trabajo TinyML4D: TinyML for Developing Countries perteneciente a la red Tiny Machine Learning Open Education Initiative (TinyMLedu)

<https://tinyml.seas.harvard.edu/4D/>

