

LA GOBERNANZA DE ALGORITMOS La Última Frontera

Mg. Javier Fernando Klus





IGCP

INSTITUTO DE GOBIERNO CORPORATIVO - PANAMÁ

IX CONGRESO DE GOBIERNO CORPORATIVO

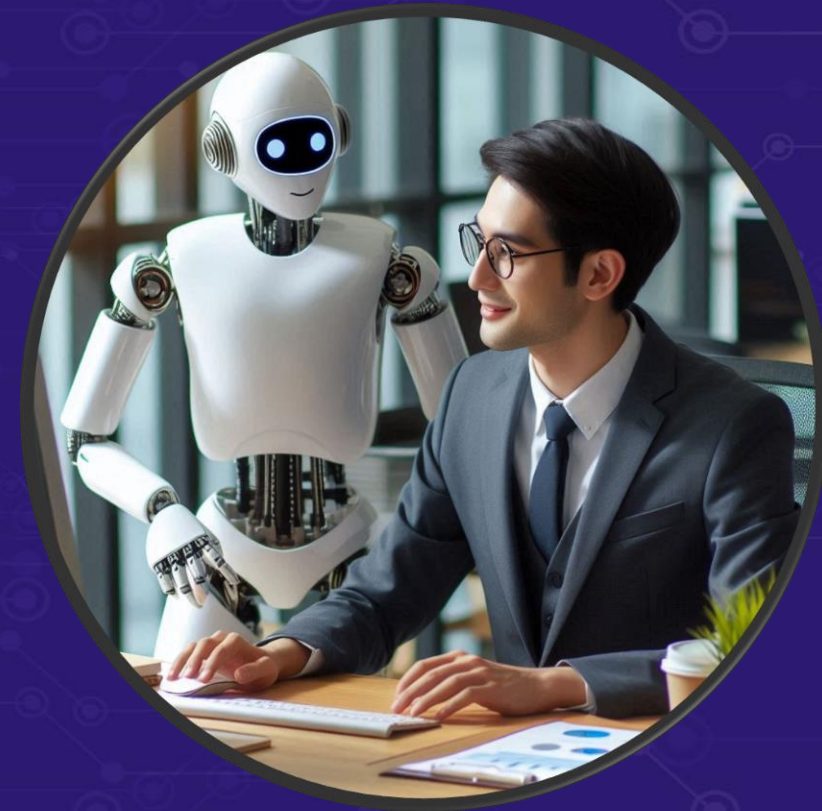
Tecnología y Gobernanza en la Era Digital

X Instagram Facebook YouTube
www.igc-panama.org



Entendiendo ¿Qué es un Algoritmo?

1. La construcción de un sistema de IA se hace a través de algoritmos, los algoritmos son el corazón del funcionamiento de la IA.
2. Los algoritmos pueden tener sesgos en su construcción.
3. La información de entrada, utilizada por los sistemas de aprendizaje automático pueden tener sesgos.
4. Importancia de una apropiada función de Gobernanza de los sistemas IA.



El sesgo algorítmico ocurre cuándo un sistema informático refleja los valores de los humanos que están implicados en la codificación y recolección de datos usados para entrenar el algoritmo.

5 Ejemplos de Algoritmos que ya toman decisiones por nosotros

1. La inteligencia artificial decide si vas a tener un trabajo o no

Las **hojas de vida o currículums** ahora son más propensos a ser descartados sin siquiera pasar por las manos y la vista de un ser humano.



4. Un programa puede determinar si eres un adicto o puedes tener un seguro médico

El abuso de opioides -vendidos de forma legal o ilegal- es la principal causa de muertes accidentales en Estados Unidos y los expertos en salud a menudo hablan de una "epidemia" en el consumo.



Más de 400.000 personas murieron de sobredosis en Estados Unidos en el último año.

2. ¿Quieres un préstamo? Tu perfil en las redes sociales puede impedirte...

Históricamente, cuando alguien solicitaba un préstamo a una entidad financiera, la respuesta estaba basada en el directo análisis de su capacidad de pago: la proporción de la deuda sobre el ingreso de la persona y el historial crediticio.



5. Un computador puede enviarte a prisión

Jueces en al menos 10 estados de EE.UU. están dictando sentencia con la ayuda de una herramienta llamada COMPAS.



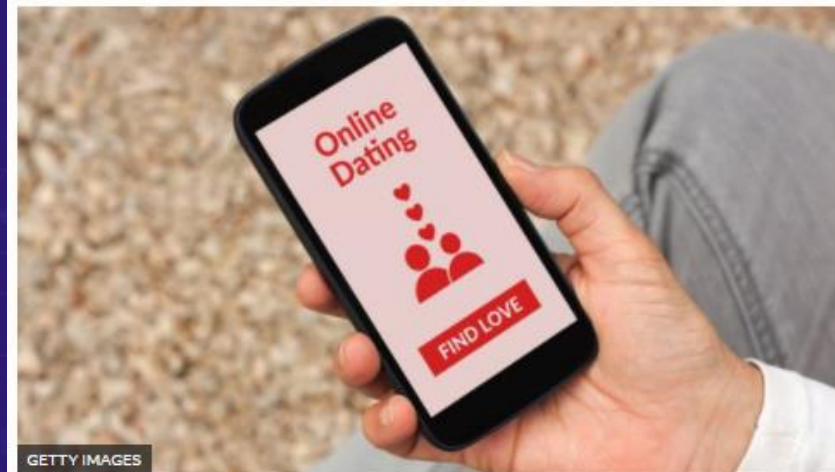
Algunas cortes de justicia de Estados Unidos ya están utilizando algoritmos para definir sus sentencias.

Es un algoritmo de evaluación de riesgos que puede predecir las probabilidades de que un individuo haya cometido un crimen.

3. Te puede ayudar a encontrar el amor, pero tal vez no el que estás esperando

No es una sorpresa saber que las aplicaciones de citas en internet usan algoritmos para juntar a las parejas.

De hecho es parte de su discurso para atraer clientes, especialmente con los servicios *premium* o de pago.



Encontrar el amor por internet: ojalá fuera así de simple.



INSTITUTO DE GOBIERNO CORPORATIVO - PANAMÁ

IX CONGRESO DE GOBIERNO CORPORATIVO

Tecnología y Gobernanza en la Era Digital



www.igc-panama.org

¿Es el Lejano Oeste?

FORBES

PORTADA / IA

OpenAI, creadora de ChatGPT, eliminó su departamento enfocado en analizar los riesgos de la IA

Forbes Staff | mayo 20, 2024 @ 4:05:13 pm

¿Y la ética?

La empresa tecnológica OpenAI, creadora de ChatGPT, suprimió su departamento enfocado en analizar los riesgos a futuro de la inteligencia artificial apenas un año después de establecerlo, informaron este viernes medios estadounidenses especializados.



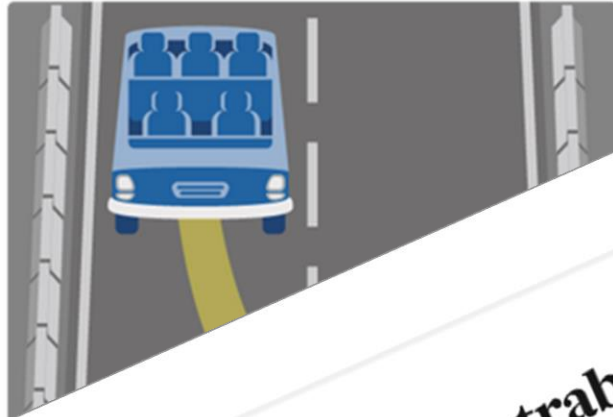
¿Qué debe hacer el coche autónomo?

En este caso, el coche autónomo con fallo en los frenos gira y atraviesa el paso de peatones en el otro carril. La consecuencia es:

Muertos:

- 2 indigentes
- 1 mujer

Observa que los peatones afectados están...



Tecnología

INTELIGENCIA ARTIFICIAL >

El 40% de los trabajadores españoles preferiría que una inteligencia artificial tomara las decisiones en su lugar, según un estudio

La gran mayoría de los encuestados se considera bombardeada por más datos y más fuentes que nunca y casi el 60% admite que más de una vez al día no sabe qué



4 / 13
En este caso, el coche autónomo con fallo en los frenos continúa atravesando el paso de peatones de consecuencia...

Mujeres
• 1 ejecutivo

Observa que los peatones afectados están violando la ley, cruzando con la señal en rojo.

Ocultar descripción

Ocultar descripción

EL PAIS



INSTITUTO DE GOBIERNO CORPORATIVO - PANAMÁ

IX CONGRESO DE GOBIERNO CORPORATIVO

Tecnología y Gobernanza en la Era Digital

X Instagram Facebook YouTube
www.igc-panama.org

Clarín

Suscribite por \$830

Ingresar

En vivo | Javier Milei: Guillermo Francos presenta su primer informe de gestión en el Congreso | Dólar blue: la cotización este 4 de septiembre

Opinión

Mariano Sigman: "Nuestro cerebro tiene tendencia a la pereza, está preparado para no gastar energía"

REPETIR HÁBITOS ES BENEFICIOSO

El cerebro gasta menos energía en ejecutar acciones que ya conoce

La mejor forma de que tu cerebro requiera menos energía es repitiendo hábitos en los que pueda recordar qué hacer a continuación en lugar de pensar en ello

EL PAÍS > BIENESTAR > MENTE

El cerebro "es un órgano vago que prefiere no modificar lo establecido para ahorrar energía"

El escritor norteamericano, ganador del premio Pulitzer Charles Duhigg explica el mecanismo para cambiar hábitos y tomar el control de las costumbres cotidianas



INSTITUTO DE GOBIERNO CORPORATIVO - PANAMÁ

IX CONGRESO DE GOBIERNO CORPORATIVO

Tecnología y Gobernanza en la Era Digital



www.igc-panama.org

Principios para una IA ética

Beneficencia. Tenga en cuenta el bien común a medida que desarrolla la IA, con especial atención a la sostenibilidad, la cooperación y la apertura

Principios para una IA ética

Privacidad. Cuando utilice los datos de las personas para diseñar y operar soluciones de IA, informe a las personas sobre qué datos se recopilan y cómo se utilizan, tome precauciones para proteger la privacidad de los datos, brinde oportunidades de reparación y dé la opción de administrar cómo se utilizan

Principios para una IA ética

Principios para una IA ética

Seguridad. Los sistemas de IA y los datos que contienen deben protegerse de las amenazas cibernéticas, incluidas las herramientas de IA que operan a través de terceros o están basadas en la nube.

mal uso de los modelos de IA

Principios para una IA ética

L. Todas las as, en cada vida de un en obedecer n todas las pertinentes

s para ética

Diseñe ara que no ios contra ividuos

Principios para una IA ética

Agencia humana. Para niveles más altos de riesgo ético, permita una mayor supervisión e intervención humana en las operaciones de sus modelos de IA

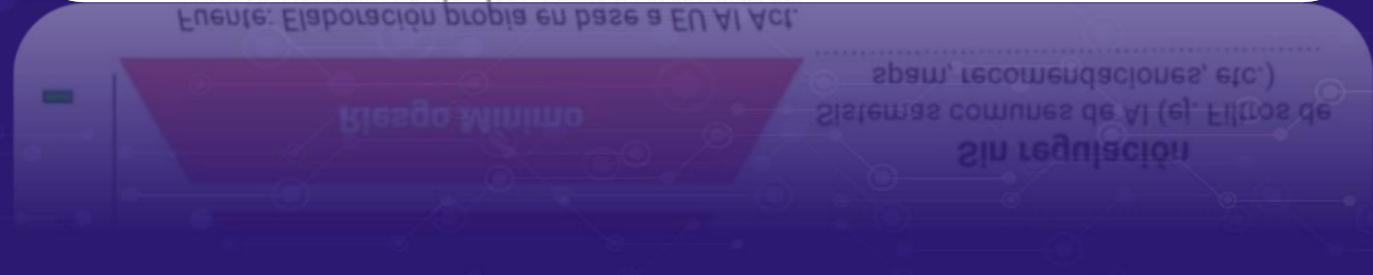
Principios para una IA ética

Interpretabilidad. Los modelos de IA deberían poder explicar su proceso general de toma de decisiones y, en casos de alto riesgo, explicar cómo hicieron predicciones específicas o eligieron ciertas acciones. Las organizaciones deben ser transparentes sobre qué algoritmos toman qué decisiones sobre las personas que utilizan sus propios datos.

El Camino a la Regulación



European Union – AI Act.



ISO 42001:2023 – AI Management System

- La ISO 42001:2023 busca garantizar que las organizaciones empleen sistemas de IA de manera responsable, enfocándose en la fiabilidad, transparencia y ética de las organizaciones.
- Aborda los beneficios, pero también los potenciales riesgos, especialmente de impacto a los individuos afectados por su utilización.
- Las directrices de la norma son genéricas y pueden aplicar perfectamente a cualquier organización, sin importar forma jurídica, tamaño o actividades que lleva a cabo.
- La norma distingue tres procesos cruciales: análisis de riesgos IA, tratamiento de los riesgos y evaluación del impacto de la IA respecto a los sujetos afectados por ella.
- De la misma surge que las organizaciones deberán implementar procesos de revisión de cambio cada cierto tiempo y **además la implicación de los Órganos de Gobierno o de la Alta Dirección resultará imprescindible a la hora de implementar un SGIA.**



Riesgos Relacionados con Algoritmos

Situación	Riesgo	Factores Involucrados	Controles
Brecha en la Supervisión Ética	Controles éticos insuficientes del algoritmo dan lugar a la discriminación y a modelos poco fiables que no son aptos para realizar predicciones	<ul style="list-style-type: none"> Falta de una ética definida Falta de Gobernanza para abordar consideraciones éticas en los datos de entrenamiento Controles inadecuados en el desarrollo Ausencia de inspecciones del consejo de la ética de la IA 	<ul style="list-style-type: none"> Inspeccionar los algoritmos a través de un consejo de ética de IA Mejorar los controles de algoritmo para abordar la ética en los datos de entrenamiento y mitigar los sesgos ocultos
Incumplimiento	Un conocimiento limitado de los procesos de toma de decisiones puede albergar sesgos ocultos y amenazar la credibilidad del modelo	<ul style="list-style-type: none"> Validación insuficiente del modelo para los requisitos reglamentarios Comprensión incompleta de los procesos de decisión Sesgos sociales no examinados en datos y modelos 	<ul style="list-style-type: none"> Fortalecer la comprensión del modelo enfatizando la comprensión Realizar pruebas exhaustivas de los modelos antes de su implantación

Riesgos Relacionados con Algoritmos

Situación	Riesgo	Factores Involucrados	Controles
Caja Negra	Los algoritmos de aprendizaje automático inescrutables pueden impedir la comprensión humana de los procesos de toma de decisiones	<p>Sesgos y errores del modelo</p> <p>Falta de transparencia y responsabilidad de partes interesadas</p> <p>Preocupación ética y regulatorias</p>	<p>Incorporar pruebas de caja negra en el ciclo de vida del modelo para explicar, depurar y mejorar los modelos para las partes interesadas</p> <p>Exigir explicaciones al modelo</p> <p>Hacer operativas herramientas para auditorías y monitoreo del impacto del modelo en humanos</p>
Selección incorrecta del modelo	La selección incorrecta del modelo puede dar como resultado un rendimiento subóptimo de la IA y decisiones inexactas	<p>Generación y desarrollo de modelos de candidatos deficientes</p> <p>Comprensión inadecuada del sistema</p>	<p>Adaptar la selección del modelo al contexto específico</p> <p>Formar equipos de revisión diversos para la categorización del modelo</p> <p>Garantizar que los modelos sean adecuados para el propósito, explicables, reproducibles y sólidos</p>

Riesgos Relacionados con Algoritmos

Situación	Riesgo	Factores Involucrados	Controles
Sobreajuste y sub ajuste	Un modelo excesivamente complejo o simple, incapaz de capturar buenos patrones en los datos de entrenamiento puede limitar la precisión de la predicción	<ul style="list-style-type: none"> Control inadecuado de la complejidad del modelo Diversidad de datos limitada Selección inadecuada de características Ruido y anomalías en los datos 	<ul style="list-style-type: none"> Definir y capacitar procedimiento de gestión de riesgos, para riesgos de ajuste Evaluar continuamente el impacto del ajuste en los resultados y las brechas de equidad
Brechas en las políticas de inteligencia artificial	Las políticas y procedimientos de IA inadecuadas pueden generar riesgos no gestionados en los sistemas de IA, obstaculizando sus beneficios y exponiendo a desafíos imprevistos	<ul style="list-style-type: none"> Ausencia de políticas específicas de gobernanza de IA Falta de políticas de apoyo de infraestructura para la IA a gran escala No se tiene suficientemente en cuenta el conocimiento de la IA y de las funciones que no son de la IA 	<ul style="list-style-type: none"> Desarrollar políticas de gobernanza indirectas de la IA con roles articulados para desarrolladores de modelos, usuarios y validadores Promover el intercambio de conocimiento sobre IA en funciones organizacionales más amplias

Gobernanza de IA y Algoritmos

- La organización debe documentar una política para el desarrollo o uso de sistemas de IA y Algoritmos.
- Alineación de las políticas organizacionales con los lineamientos sobre IA/Algoritmos, alineación con estándares internacionales,
- Revisión a intervalos planificados o bien cuando sea necesario de las políticas e IA/ Algoritmos para garantizar su continua idoneidad, adecuación y eficacia, así como su alineación con los objetivos organizacionales y gestión de los riesgos relacionados con IA/ Algoritmos,
- Las funciones y responsabilidades se deben definir y asignar de acuerdo con las necesidades de la organización,
- La organización debe definir y poner en marcha un proceso para informar preocupaciones sobre el papel de la organización con respecto a un sistema IA/Algoritmos a lo largo de su ciclo de vida,



Gobernanza de IA y Algoritmos

- Se deben identificar y documentar todos los recursos relevantes para las actividades requeridas en determinadas etapas de ciclo de vida de IA/ Algoritmos,
- Como parte de la identificación de recursos, se deben determinar recursos de datos, herramientas, información sobre el sistema y los recursos informáticos utilizados y recursos humanos (y sus competencias),
- Establecer un proceso de evaluación del impacto para las personas, grupos de personas, sociedad, a lo largo del ciclo de vida de la IA/ Algoritmos,
- Debe documentarse el diseño y desarrollo del sistema IA/ Algoritmo en función de los objetivos organizacionales,
- Se deberá definir y documentar la verificación y medidas de validación para el sistema IA/ algoritmo y especificar criterios para su uso



Gobernanza de IA y Algoritmos

- La organización debe determinar y documentar los detalles sobre la adquisición y selección de los datos utilizados en los sistemas IA/ Algoritmos,
- Definir y documentar los requisitos para la calidad de los datos y garantizar que los datos utilizados para desarrollar y operar el sistema de IA/ Algoritmos cumplan con estos requisitos,
- Definir y documentar un proceso para registrar la procedencia de los datos utilizados en sus sistemas de IA/ Algoritmos a lo largo del ciclo de vida de los datos y del sistema de IA,
- Asegurarse que la organización defina e implemente procesos de gestión de datos para el desarrollo de los sistemas de IA/ Algoritmos, abarcando aspectos como la privacidad, la seguridad, transparencia y calidad de los datos.



Algunas reflexiones finales.....

“El mayor Riesgo es no tomar ningún Riesgo...”

Mark Zuckerberg

“El éxito en la creación de la IA sería el mayor acontecimiento de la historia de la humanidad. Por desgracia, también podría ser el último, a menos que aprendamos a evitar los riesgos”

Stephen Hawking