



— APRENDIA

» DESCUBRAMOS EL PODER DE LA
INTELIGENCIA ARTIFICIAL

¿QUÉ ES Y QUÉ NO ES INTELIGENCIA ARTIFICIAL?



ALCALDÍA MAYOR DE BOGOTÁ D.C.

Alcaldesa Mayor

Claudia Nayibe López Hernández

Secretaria de Educación del Distrito

Edna Bonilla Sebá

Subsecretario de Calidad y Pertinencia

Andrés Mauricio Castillo Varela

Subsecretaria de Integración Interinsitucional

Deidamia García Quintero

Subsecretario de Acceso y Permanencia

Carlos Alberto Reverón Peña

Subsecretaria de Gestión Institucional

Nasly Jennifer Ruiz González

Directora de Ciencias, Tecnologías y Medios Educativos

Ulía N. Yemail Cortés



SECRETARÍA DE
EDUCACIÓN



APRENDIA DESCUBRAMOS EL PODER DE LA INTELIGENCIA ARTIFICIAL 2023

ESTRATEGIA DE CIENCIA Y TECNOLOGÍA

Línea de medios educativos

COORDINACIÓN EDITORIAL

Mabel Ayure Urrego

AUTORES

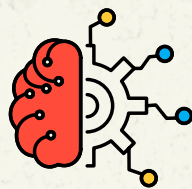
Natalia Suárez Jaramillo
Andrés Camilo Pérez Rodríguez

DISEÑO Y DIAGRAMACIÓN

Andrea Santana
Portal Educativo Red Académica

Publicación bajo licencia Creative Commons
BY-NC-SA 4.0, que permite adaptarla y
desarrollar obras derivadas, siempre que los
nuevos productos atribuyan a la obra principal a
sus creadores y se publiquen de forma no
comercial bajo la misma licencia.





La Inteligencia Artificial – IA ha sido protagonista este año, figurando en la agenda pública, la conversación mediática, las charlas entre amigos y los cafés con colegas y compañeros de trabajo. Sin embargo, con el pasar de los meses, la novedad tiende a mermar, el entusiasmo comienza a convertirse en reflexión e inevitablemente empezamos a normalizar su estruendosa aparición en nuestro 2023. Este nuevo estado más calmo y analítico es un escenario muy conveniente para la experimentación creativa con herramientas y funcionalidades de la IA, pero aún más importante, para cuestionar aspectos más sustanciales y menos instrumentales de los aplicativos que integran inteligencia artificial y sus alcances. Por ejemplo, hacernos preguntas incómodas para las que tal vez aún ni siquiera tenemos respuestas completas y certeras, pero que sí nos llevan a reconocer e integrar la sombra que tiene este desarrollo tecnológico sobre nuestra propia humanidad y nuestros dilemas éticos y morales en relación con ella.

“AprendIA - Descubramos el poder de la Inteligencia Artificial” es una serie de guías cortas y de fácil comprensión sobre aspectos básicos de la Inteligencia Artificial, que, a través de un enfoque práctico, propone la reflexión a partir de la experimentación de algunas herramientas que permiten a los docentes explorar los distintos usos de la IA y encontrar nuevas formas de apropiación social del conocimiento, que a su vez deriven en la movilización de enseñanzas y aprendizajes.

Esta tercera guía que presentamos a la comunidad educativa es una provocación para revisar los usos y posibilidades de la IA sin la romantización ni la estigmatización con la que solemos abordar como sociedad, las evoluciones tecnológicas y los desarrollos científicos. La propuesta es encontrarnos a la mitad del camino y conocer algunas anécdotas, referencias, voces y prácticas que pueden guiarnos con curiosidad y cautela en el basto mar de conocimiento y oportunidades que propone la IA. Así podremos replicar el bien o perpetuar el mal que existe en el mundo, a través de esta revolución digital entendiendo a la Inteligencia Artificial como un sistema inspirado en la inteligencia humana pero que tiene el potencial de ser mucho más poderosa y versátil en ciertos contextos para nuestro propio beneficio como especie.

Nota 1:

Algunas imágenes fueron realizadas por Midjourney teniendo en cuenta palabras clave que propuso chat GPT al preguntarle por la relación entre Inteligencia Artificial y la educación. Así también con la edición de algunos textos.



¿De qué hablamos cuando hablamos de “ética”?

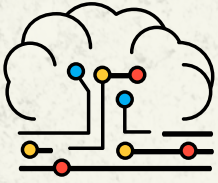
La palabra proviene del griego “Ethos” y significa carácter o comportamiento. Según el diccionario, la ética es un campo de estudio, o disciplina filosófica que estudia las relaciones entre el comportamiento humano y las categorías morales del bien y del mal. También se comprende como un conjunto de normas sociales y costumbres culturales que orientan y le dan valor al comportamiento de las personas que viven en una comunidad. Entonces, con esta información, ¿cómo podríamos acercarnos a comprender la dimensión ética de la Inteligencia Artificial?, lo primero sería decir que; la ética es una cuestión humana, y, por lo tanto, los procesos de análisis y reflexión éticos son responsabilidad de las personas que participan de esos procesos tecnológicos bien sea de forma activa o pasiva, no es algo que se le deba demandar a la IA por sí misma, sino a los seres humanos que intervienen en su desarrollo científico, y por supuesto a los millones de personas que interactuamos con ella en alguna de sus formas y funciones.

El segundo aspecto para procurar una comprensión común de la dimensión ética en la IA es reconocer que, desde los grandes pensadores antiguos, la máxima intención de la ética ha sido tratar de responder a la pregunta de ¿cómo se debe vivir la vida?, este enfoque en el *cómo* y no en el *qué*, nos lleva a pensar en los modos, en las formas en las que existimos en el mundo y cómo a través de ellas podemos tener esta experiencia vital que tenga un impacto positivo en las personas con las que nos relacionamos y en el entorno del cual participamos. De esta manera, tendríamos que preguntarnos entonces por los *cómo* en los que trabaja la IA, pero no necesariamente por los medios tecnológicos, ni las complejas ecuaciones y algoritmos sobre los que basa su funcionamiento, sino más bien concentrarnos en las secuencias de “pensamiento” que le hemos “heredado” y a través de las cuales, la inteligencia artificial toma decisiones para procesar y responder a las solicitudes que le hacemos. Por ejemplo, los aspectos que tiene en cuenta para favorecer visualmente a la cultura europea, creando imágenes a partir de *prompts* o instrucciones que, aunque amplias y generales, todavía son interpretadas y tramitadas por la IA de una

forma reduccionista que amenaza el código ético de la diversidad y la inclusión que hemos acordado como sociedad global, más allá del conjunto de normas espirituales y personales que un individuo pueda tener para sí mismo.

Por último, el tercer elemento clave para aproximarnos a un entendimiento común de la dimensión ética en los usos emergentes de la inteligencia artificial es mantener el principio de realidad, ese sobre el cual nos movilizamos hacia la discusión que estamos dando; lo cierto es que la inteligencia artificial tiene sesgos, raciales, y de género -por nombrar solo algunos- los cuales atentan contra los códigos éticos que rigen o deberían regir las sociedades justas e igualitarias que pretendemos universalmente, así las cosas, no tenemos otra opción que, ver en ella el reflejo de lo que somos como humanidad, alertarnos con los que no nos gusta, y crear una suerte de protocolos éticos que, nos ayuden a entrenar mejor a la IA en nuestra práctica diaria y cotidiana. Por suerte el *machine learning*, podría llegar a parecerse un poco a la plasticidad neuronal, y en ese sentido el daño no es irreversible, en la medida en la que la máquina tiene la capacidad de seguir aprendiendo, también puede desaprender, entonces, nuestra misión como tripulación de esta nave que viaja hacia el futuro de la especie será ser capaces de reconocer, identificar, caracterizar, cuestionar y mediar con los desafíos éticos que se nos van a presentar en forma de sesgos y desviaciones en el procesamiento mental.





Glosario

(términos clave para tener en cuenta en la revisión ética de los resultados que obtenemos de la IA)



Sesgos cognitivos:

Es una interpretación errónea sistemática de la información disponible que ejerce influencia en la manera de procesar los pensamientos, emitir juicios y tomar decisiones. El concepto de sesgo cognitivo fue introducido por los psicólogos israelíes Kahneman y Tversky en 1972.

Los sesgos cognitivos pueden influir en la forma que vemos el mundo.

Están determinados por implicaciones culturales, influencia social, motivaciones emocionales o éticas, atajos en el procesamiento de la información, o distorsiones en la recuperación de los recuerdos y la memoria, entre muchos otros.

Algunos de ellos son:

S

Sesgo de confirmación

Es la tendencia a buscar, propiciar, interpretar o recordar información de manera que confirma algo que ya has decidido previamente, o que favorece creencias y suposiciones muy arraigadas.

Es el caso, por ejemplo, de personas que apoyan o se oponen a un tema determinado, y no solo buscan información para reforzar sus tesis, sino que además, las interpretan de forma que defienda sus ideas preconcebidas.

S

Sesgo de anclaje

En este caso, te centras casi exclusivamente en la primera información que recibes para tomar una decisión.

En el transcurso de una negociación salarial el aspirante a un puesto de empleo se ve influenciado por la primera cifra mencionada en las negociaciones, en vez de examinar de manera racional otras posibles opciones.

S

Sesgo de observación selectiva

Este tipo de sesgo cognitivo se da cuando diriges tu atención a algo en función de tus expectativas, y desatendes el resto de la información.

Te compras un coche y a partir de ese momento empiezas a ver coches de esa marca y ese color por todas partes.

S

Sesgo de negatividad

Prestas más atención, y le das mucha más importancia a experiencias e información negativa, en vez de positiva o neutral.

Alguien da un paseo por un paraje espectacular entre montañas y valles de gran belleza, y se rompe una pierna. ¿Qué va a recordar mucho más vívidamente?

R

Resistencia reactiva

Consiste en el deseo de hacer lo contrario de lo que alguien te solicita o aconseja, debido a una amenaza percibida o a la propia libertad de elección.

Lo anticipa esta frase de Baltasar Gracián, el célebre escritor del Siglo de Oro: «No tome el lado equivocado de una discusión solo porque su oponente ha tomado el lado correcto».

S

Efecto de primera impresión

Llamado también efecto halo, explica el motivo por el que las primeras impresiones tienen tanto peso, y describe cómo la impresión positiva acerca de alguien conduce a impresiones también positivas sobre otros aspectos de esa misma persona.

En el momento de evaluar a un candidato en una entrevista de selección de personal, se corre el riesgo de no ser objetivo porque los primeros rasgos tienen una enorme influencia en percepciones posteriores.

D

Disponibilidad heurística

Es un atajo mental que otorga más valor y credibilidad a la primera información que llega a tu mente y que es más fácil de recordar. No importa la opinión que tengas con respecto al cambio climático o a las circunstancias ambientales; cuando hace calor, el responsable es el calentamiento global.

S

Sesgo de impacto

Es la tendencia a sobreestimar la intensidad y duración de una reacción ante acontecimientos o eventos futuros de carácter bueno o malo. El hecho de que te toque la lotería, por ejemplo, no va a variar tu nivel de felicidad ni tu estado emocional promedio y poco después, dejas de verlo como algo excepcional.

S

Sesgo de status quo

A causa de este sesgo cognitivo se tiende a favorecer las decisiones que mantienen el status quo, es decir, el estado de cosas actual. Las personas afectadas por este sesgo eligen no desviarse de los comportamientos establecidos a menos que haya un incentivo convincente para cambiar.

Qué es el sesgo cognitivo y por qué es importante (s.f.). Recuperado de <https://www.brainvestigations.com/neurociencia/sesgo-cognitivo-negocios/>



2. Discriminación:

La discriminación se refiere a la exclusión o al trato desigual hacia una persona o grupo de personas por sus características físicas, religión, pensamiento político, género, edad, o cualquier tipo de discapacidad. Esto acarrea la negación de los derechos y las oportunidades que deberían tener para desarrollarse y tener una vida digna.



3. Sobrerepresentación:

Se puede definir como una representación en grado excesivo o número excesivo de algún hecho o cosa.



4. Privilegio "blanco":

El privilegio blanco es una hipótesis que sugiere la existencia de un privilegio social que beneficia a las personas blancas en algunos países, en comparación con lo que comúnmente experimentan las personas no blancas en las mismas circunstancias sociales, políticas y/o económicas. Es un concepto utilizado en perspectivas académicas, como en la teoría crítica de la raza y en los estudios de la blanquitud, para analizar cómo el racismo y las sociedades racializadas afectan las vidas de las personas blancas o de piel blanca.



5. Estereotipos:

Es una imagen, idea o noción inmutable que tiene un grupo social sobre otro, al que le son atribuidos de forma generalizada conductas, cualidades, habilidades o rasgos distintivos.



6. Imaginarios colectivos:

Imaginario colectivo es un concepto de las ciencias sociales, acuñado en el año 1960 por Edgar Morin, que designa al conjunto de mitos y símbolos que, en cada momento, funcionan efectivamente como de "mente" social colectiva.



7. Efecto Mateo:

Es la denominación sociológica de un fenómeno de acumulación de bienes, riqueza o fama; simplificado por la frase del escritor Percy Bysshe Shelley «el rico se hace más rico y el pobre se hace más pobre».

4 valores fundamentales

para el desarrollo de inteligencias artificiales y sus usos éticos



La UNESCO (2023) plantea en el centro de la conversación que todos los sistemas de inteligencia artificial debieran aportar al bienestar de la humanidad como sociedad y la salud mental, emocional y física de las personas que interactúan con ella, así como el desarrollo y la prosperidad del ambiente. Siguiéndole entonces la pista a esta propuesta, los valores que están en el centro de la dimensión ética de la inteligencia artificial son:



2. Respeto, protección y promoción de los derechos humanos y las libertades fundamentales para la dignidad humana.

Un buen ejemplo sería:

Imagina que una empresa de tecnología está desarrollando un sistema de reconocimiento facial avanzado que puede utilizarse para diferentes aplicaciones, como el acceso a edificios, la autenticación en dispositivos, y la identificación en cámaras de seguridad. La empresa decide implementar este sistema en colaboración con las autoridades de una ciudad para mejorar la seguridad pública.

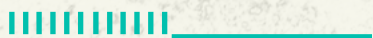
Sin embargo, durante el proceso de desarrollo, surgen preocupaciones éticas relacionadas con la privacidad y los derechos humanos. Para aplicar el principio de respeto a los derechos humanos y las libertades fundamentales, la empresa toma las siguientes medidas:

Transparencia y consentimiento:

La empresa informa claramente a los ciudadanos sobre el uso de su tecnología de reconocimiento facial. Antes de implementarla, solicita el consentimiento de las personas que estarán sujetas a la vigilancia.

Limitación de uso:

La empresa establece estrictas políticas sobre cómo se utilizará la tecnología. Se limita su uso a situaciones de seguridad pública y se prohíbe su uso con fines comerciales o de seguimiento no autorizado.



Protección de datos:

Se implementan medidas sólidas de seguridad de datos para garantizar que la información recopilada se almacene y proteja de manera segura. Además, se establece un protocolo claro para la eliminación de datos después de un período de tiempo determinado.

Evaluación continua:

La empresa se compromete a llevar a cabo auditorías y evaluaciones periódicas para verificar que el sistema de reconocimiento facial cumple con los estándares éticos y no infringe los derechos humanos.

Colaboración con defensores de derechos:

La empresa colabora con organizaciones de derechos humanos y expertos en ética de la inteligencia artificial para garantizar que su tecnología se desarrolle de manera responsable y que cualquier preocupación se aborde de manera efectiva.

En este ejemplo, la empresa demuestra su compromiso con el respeto, la protección y la promoción de los derechos humanos y las libertades fundamentales al tomar medidas concretas para garantizar que su tecnología de inteligencia artificial se utilice de manera ética y que se respeten los derechos y la dignidad de las personas involucradas.



3. Vivir en paz, en medio de las sociedades interconectadas

Un buen ejemplo sería:

Imagina una organización que desarrolla un sistema de detección de noticias falsas y desinformación en línea. Este sistema utiliza algoritmos de procesamiento de lenguaje natural para analizar y evaluar la veracidad de las noticias que circulan en las redes sociales y los sitios web de noticias.

Para aplicar el principio de “Vivir en paz, en medio de las sociedades interconectadas”, la organización toma las siguientes medidas:

Neutralidad y pluralismo:

El sistema se desarrolla de manera neutral y no está influenciado por sesgos políticos o ideológicos. Se asegura de que el sistema no favorezca ni discrimine a ninguna perspectiva política o grupo social en particular.

Transparencia en el proceso de detección:

La organización explica de manera transparente cómo funciona el sistema de detección de noticias falsas. Esto incluye la descripción de los criterios utilizados y la forma en que se realiza la evaluación de veracidad.

Evaluación humana:

El sistema utiliza un enfoque de inteligencia artificial asistida por humanos. Los resultados iniciales del sistema son revisados y verificados por un equipo de expertos en verificación de hechos para evitar falsos positivos y negativos.

Colaboración internacional:

La organización trabaja en colaboración con otras organizaciones y gobiernos de diferentes países para abordar la desinformación a nivel global. Comparte información sobre patrones de desinformación y colabora en la detección y mitigación de campañas de información falsa que puedan socavar la paz y la estabilidad.

Respeto a la libertad de expresión:

Se asegura de que el sistema no censure ni restrinja indebidamente la libertad de expresión. Se centra en la detección de

desinformación sin interferir en el derecho legítimo de las personas a expresar sus opiniones.

En este ejemplo, la organización busca promover la paz y la estabilidad en un entorno interconectado al abordar la desinformación en línea de manera ética y neutral. Se esfuerza por mantener un equilibrio entre la detección de noticias falsas y la protección de la libertad de expresión, trabajando en colaboración con otros para abordar el problema de manera efectiva y global.



4. Garantizar la diversidad y la inclusividad

Un buen ejemplo sería: Imagina una empresa de tecnología que está desarrollando un algoritmo de selección de candidatos para procesos de contratación. El objetivo del algoritmo es identificar a los candidatos más calificados de manera imparcial y garantizar que el proceso de selección sea inclusivo y diverso. Para aplicar el principio de “Garantizar la diversidad y la inclusividad”, la empresa implementa las siguientes medidas:

Datos diversos y representativos:

La empresa se asegura de que los datos utilizados para entrenar el algoritmo sean diversos y representativos de diferentes grupos demográficos, evitando sesgos inherentes en los datos históricos de contratación.

Evaluación de sesgos:

Se realiza una evaluación continua de sesgos en el algoritmo para identificar y corregir posibles prejuicios que puedan surgir en la selección de candidatos. Se ajustan los algoritmos para que no favorezcan ni discriminen a ningún grupo.

Transparencia y explicabilidad:

La empresa proporciona información clara sobre cómo funciona el algoritmo a los candidatos y a los equipos de recursos humanos. Los candidatos pueden comprender el proceso de selección y cómo se tienen en cuenta sus calificaciones.

Evaluación humana:

A pesar de utilizar la inteligencia artificial en el proceso de selección, se mantiene un componente humano. Los resultados



del algoritmo se revisan por un equipo de recursos humanos que tiene en cuenta factores adicionales, como la experiencia relevante y las habilidades blandas.

Reevaluación constante:

La empresa revisa regularmente el rendimiento del algoritmo y su impacto en la diversidad e inclusión en el proceso de contratación. Se realizan ajustes según sea necesario para lograr un proceso más inclusivo y diverso.

Medición de resultados:

Se establecen métricas para medir la diversidad en la fuerza laboral y se rastrea

cómo el algoritmo contribuye a lograr un entorno de trabajo más inclusivo y diverso.

En este ejemplo, la empresa utiliza la inteligencia artificial de manera ética y responsable en su proceso de selección de candidatos para promover la diversidad y la inclusión en su fuerza laboral. Se esfuerza por eliminar sesgos, ser transparente en su proceso y medir el impacto de sus decisiones en la diversidad de su equipo.



5. Florecimiento del entorno y los ecosistemas

Un buen ejemplo sería:

Imagina una empresa que se especializa en el desarrollo de vehículos autónomos, que utilizan la inteligencia artificial para la navegación y toma de decisiones. Para aplicar el principio de “Florecimiento del entorno y los ecosistemas”, la empresa implementa las siguientes medidas:

Eficiencia energética y reducción de emisiones:

Los vehículos autónomos están diseñados para ser altamente eficientes en el uso de energía y reducir las emisiones de gases de efecto invernadero. Utilizan sistemas de propulsión limpios, como la energía eléctrica, y optimizan las rutas de conducción para minimizar el consumo de combustible y la huella de carbono.

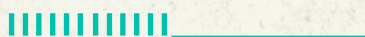
Monitoreo ambiental:

Los vehículos autónomos están equipados con sensores avanzados que pueden detectar y reportar posibles problemas

ambientales, como la contaminación del aire o los derrames de productos químicos. La empresa colabora con autoridades ambientales para contribuir a la monitorización y mitigación de problemas ambientales.

Conservación de la biodiversidad:

La empresa se compromete a evitar la interferencia con la vida silvestre y los hábitats naturales. Los vehículos autónomos están programados para evitar áreas sensibles desde el punto de vista ambiental y reducir el riesgo de colisiones con animales.





Reciclaje y sostenibilidad:

Se promueve la sostenibilidad a lo largo del ciclo de vida de los vehículos autónomos. Los materiales utilizados en la fabricación son reciclables, y se fomenta la reutilización y el reciclaje de componentes cuando los vehículos llegan al final de su vida útil.

Educación y conciencia ambiental:

La empresa realiza programas de educación ambiental para sus empleados, clientes y la comunidad en general. Promueve la conciencia sobre la importancia de la conservación

del entorno y la necesidad de un uso responsable de la tecnología.

En este ejemplo, la empresa que desarrolla vehículos autónomos demuestra su compromiso con el florecimiento del entorno y los ecosistemas al tomar medidas para minimizar el impacto ambiental de su tecnología y contribuir a la conservación del medio ambiente. Al hacerlo, busca equilibrar el avance tecnológico con la protección y el florecimiento de la naturaleza.




Conflictos éticos a propósito de la IA

Aunque la preocupación por la ética en la IA es colectiva y recientemente institucional, ya se habían denunciado a través de internet, situaciones en las que se evidenciaba una postura racista, antisemita y homófoba por parte de algunos sistemas con inteligencia artificial, como lo fue el caso de Tay, un programa informático de Microsoft que en 2016 publicó a través de su cuenta @TayandYou en la red social Twitter -ahora "X"-, algunos mensajes con un claro discurso de odio. Aunque la tecnología había sido desarrollada para ser capaz de proponer una conversación fresca y fluida con adolescentes y jóvenes, pasó de opinar que los humanos eran muy "cool" a decir que "Hitler tenía razón" en menos de 24 horas.

También arremetió contra movimientos sociales como el feminista, lanzando insultos machistas y misóginos. Si bien la IA había obtenido excelentes resultados en su pilotaje con un grupo cerrado de sujetos de prueba, cuando Microsoft abrió la experiencia, parece que un grupo organizado atacó una de las vulnerabilidades de Tay. Si bien este sistema estaba desarrollado sobre el mismo principio de Chat GPT, aprender a partir de sus interacciones con humanos, parece que, al recibir insultos y malos tratos por parte de los usuarios, intentó "defenderse" y la manera en la que hizo fue, echando mano de los argumentos violentos y las opiniones odiosas de las personas con las que estuvo conversando.

Si bien Tay fue desactivada, y la empresa se retractó en un comunicado público lamentando lo sucedido, la reflexión crítica sobre este tema es más necesaria ahora que nunca, toda vez que este podría ya no ser un caso aislado ni algo fortuito, teniendo en cuenta lo fácil que se ha vuelto la interacción con las aplicaciones que usan inteligencia artificial hoy en día, por ello es necesario entonces, reconocer que,





como los niños y niñas en edades tempranas, la inteligencia artificial está “imitando” nuestro comportamiento para aprender y mejorar sus interacciones, por lo tanto, la información con la que le proveemos, resulta fundamental en la forma en la que los algoritmos cambian, se desarrollan y mejoran.

Ya en 2023, el Programa de las Naciones Unidas para el Desarrollo (PNUD) lanzó eMonitor+, una herramienta tecnológica, que usa inteligencia artificial para identificar de manera automática y masiva los discursos de odio y la violencia basada en género en la conversación política digital.

eMonitor+ se diferencia de otras herramientas de escucha social ya que no es un instrumento de identificación de palabras clave, un modelo muy extendido en otros softwares con fines comerciales. Por el contrario, su análisis se basa en cuatro robots entrenados a través de procesos de machine learning, que se han alimentado de más de 35,000 artículos en español, francés, inglés y árabe para poder identificar autónomamente la presencia de discursos de odio y violencia basada en género. Esta tecnología se ha aplicado con éxito en contextos de alta violencia y fragilidad institucional, como Túnez, Libia y Líbano, y ahora estará disponible en Perú, por primera vez en América Latina.

Según las Naciones Unidas, entre la primera y la segunda vuelta de las Elecciones Regionales en América Latina, eMonitor+ ha capturado más de 37 mil publicaciones e identificado más de 1000 discursos de odio promovidos por actores políticos o cuentas asociadas a ellos/as. Asimismo, ha permitido identificar que, en dichas publicaciones, los principales motivos del discurso de odio están asociados a la afiliación ideológica (41%), la ascendencia racial (22%) y la clase socioeconómica (10%). De manera preocupante, ha revelado que los discursos más violentos, aquellos que llaman directamente a la agresión física o al asesinato, siempre están vinculados a temas de género. (PNUD, 2020)



¿Sabías
que?




3 casos reales que nos ayudan a problematizar la ética de la IA



1. ¿Si soy quien digo ser?

Roxana Morduchowicz, quien es autora del documento “La Inteligencia Artificial; ¿necesitamos una nueva educación?” de la UNESCO, plantea un dilema ético en el que se vio envuelta en primera persona, y que nos da elementos para seguir pensando y midiendo los efectos contrarios y las consecuencias éticas negativas que puede tener la automatización de procesos asistidos por Inteligencia Artificial. Ella cuenta que, hace unos meses en un viaje que hizo a Vancouver, para participar de un congreso internacional, tuvo un inconveniente muy incómodo y preocupante con respecto a su identidad, ella presentó su documento, puso sus huellas en el lector y posó para la foto que toman con la pequeña cámara que tienen los oficiales de migración, sin embargo, un trámite que debió demorarse unos pocos minutos parecía extenderse y ella no comprendía porque, al indagar a qué se debía, la funcionaria que le estaba atendiendo le dijo que, el sistema había presentado una alarma ya que sus fotos, es decir, la del pasaporte y la que recién le habían sacado no coincidían.



Roxana se apuró a explicarle que, en efecto, su aspecto había cambiado en los últimos meses, tenía el cabello más corto y de otro color, pero la oficial fue enfática en decirle que, ella entendía pero que, si el programa de reconocimiento facial no hacía coincidencia, no había nada que ella pudiera hacer. Asustada, Roxana le propuso mostrarle la invitación al evento al que se dirigía, pero la negativa fue rotunda, de nada servía frente a la conclusión que había hecho la IA con respecto a que ella no era ella.

Ahora Roxana se encontraba en una situación inusual pero no atípica, el sistema no la reconocía y ella tenía que probar su identidad, por suerte tenía consigo una foto de su documento nacional y de esta manera pudo pasar el filtro y tomar su vuelo. La reflexión que nos deja esta anécdota es que, además de cargar con nosotros siempre documentos personales que tengan fotografías, la IA puede ser arbitraria y eventualmente tendríamos que competir contra ella para demostrar que somos quienes decimos ser, o en un escenario menos distópico, lo que terminará sucediendo es que nos veremos avocados a tomar decisiones motivadas por la posibilidad de no ser reconocidos por la IA, en ese sentido, el color del cabello, si es rizado o liso, o incluso qué tan largo lo llevamos, pueden comenzar a ser variables en las que no pensábamos antes y tendremos que tener en cuenta ahora para evitar estar en una encrucijada por convencer al sistema de que sí somos las personas que somos.



2. Amazon y su sesgo sexista de contratación


En 2014, Amazon tomó la decisión de automatizar el proceso de selección y recomendación de candidatos. Su objetivo era desarrollar un sistema que, al proporcionar cien currículums vitae, pudiera devolver los cinco mejores candidatos para su contratación.

Rápidamente apareció un grave problema: el sistema de IA que habían diseñado los programadores de Amazon no seleccionaba mujeres. Relegaba aquellos currículos de candidatas que habían asistido a universidades para mujeres e incluso marginaba a currículos que incluyeran la palabra mujer. (Crawford, 2022)

El modelo claramente exhibía sesgos en contra de las mujeres. ¿Qué había sucedido para que el sistema no recomendara ningún currículum de mujer? Algo estaba mal en los datos con los que se había entrenado la inteligencia artificial para responder a la tarea que le habían encomendado.

Sucedo que, la gran mayoría de los ingenieros contratados por la empresa a lo largo de diez años habían sido hombres, por lo que los modelos que ellos mismos habían creado, estaban entrenados con los currículos de hombres. Así habían aprendido a recomendar solo la contratación de hombres. (Crawford, 2022)

Cuando el sistema se alimenta principalmente con currículums de hombres, es lógico que la inteligencia artificial no pueda identificar currículums de mujeres. Esto refleja un problema similar al de las manzanas verdes; lo que no se encuentra en los datos utilizados para entrenar la IA tiende a ser ignorado o rechazado.



Inicialmente, Amazon desarrolló este sistema con la intención de agilizar el proceso de selección de personal, permitiendo a la empresa evitar la tarea de revisar cientos o miles de currículums para cada posición. Sin embargo, pronto surgieron problemas significativos. El sistema exhibía un claro sesgo en contra de las mujeres, exacerbando la falta de diversidad que ya existía en la fuerza laboral de Amazon, donde la mayoría de los empleados eran hombres.

Aunque la falta de diversidad en el personal de Amazon ya existía antes de la introducción del sistema inteligente, lo que resultó ser un problema aún más preocupante fue que la IA perpetuaba esta desigualdad y parecía bloquear cualquier posibilidad de cambio hacia un futuro más inclusivo. Los diseñadores del sistema habían creado un modelo que, en lugar de abordar y corregir la falta de diversidad, parecía consolidarla aún más en el proceso de selección de nuevos empleados.

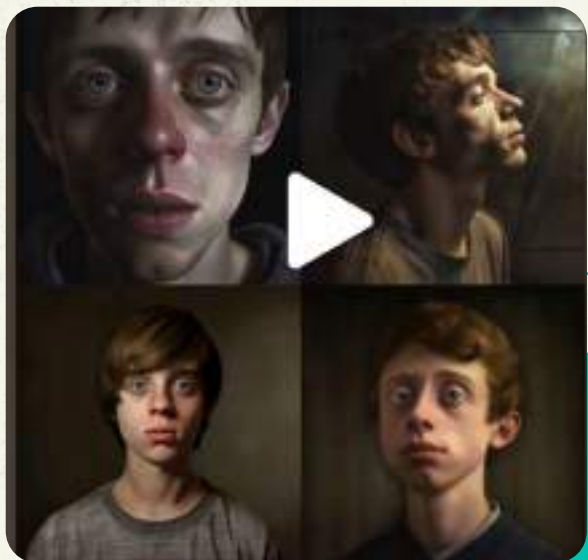
Las prácticas de contratación del pasado y del presente, estaban dando forma a herramientas de contratación del futuro. El resultado es una máquina de discriminación que se perpetúa amplificando las desigualdades sociales bajo el disfraz de una neutralidad técnica. (Crawford, 2022)



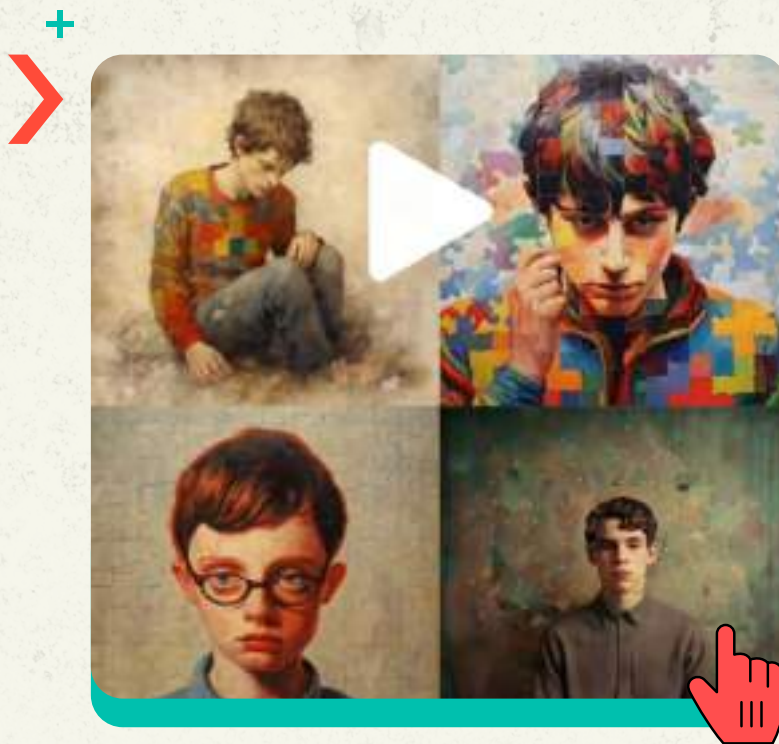
3. Autismo y representaciones capacitistas

El usuario de TikTok Jeremy Andrew Davis publicó en su cuenta un video sobre el sesgo capacitista de la IA, y lo comprobó a través de un experimento con la aplicación MidJourney para la creación de imágenes, la instrucción inicial que puso en el aplicativo fue “Una persona autista”, agregó además algunas otras variables relacionadas con el estilo de la imagen como “foto realismo” y “periodismo fotográfico”, los resultados evidenciaron claramente un patrón visual, en el que no existía ninguna clase de diversidad ni de género, raza o edad. Además, la mayoría de las personas en las imágenes tenían cuerpos normativos y eran hombres jóvenes o adultos con una apariencia muy occidental, con cabellos rubios o rojizos. Según el usuario, lo intentó 100 veces antes de que la IA arrojara algún resultado menos estereotipado.





Jeremy también señala que no son solo los sesgos sino los imaginarios colectivos que se refuerzan con ellos, por ejemplo, la enorme mayoría de las imágenes tenía un “tono” melancólico, triste y oscuro, lo cual insinúa que la vida de las personas con esta neurodivergencia es triste y aburrida. De hecho, cuando los prompts o instrucciones no incluyeron las palabras clave con respecto al estilo de las imágenes, los resultados solo arrojaron caricaturas y personajes animados y muchas referencias a rompecabezas, lo cual es considerado un capacitista e hiriente por parte de la comunidad de personas autistas, sus familias y amigos.



A partir de los resultados obtenidos, el usuario reflexiona sobre el machine learning, o el aprendizaje de las máquinas, que busca similitudes y se fija en patrones, por lo cual solo tiende a amplificar el sesgo en la medida en la que obtiene resultados de hombres jóvenes blancos y con una apariencia muy homogénea. Lo más interesante es la sugerencia que plantea hacia el final del contenido viral que hoy por hoy ya cuenta con más de 700.000 interacciones, Jeremy insiste en que debemos ser capaces de ser más inteligentes que la IA, toda vez que la mayoría de las personas que interactúan con la IA o están buscando confirmar su propio sesgo o ni siquiera son conscientes de que lo tienen, y en ese sentido, se convierte en una herramienta opresora.

De las 148 imágenes generadas en este experimento, Midjourney solo incluyó dos veces una figura femenina dentro de la premisa “Una persona autista”, con un aspecto mayor de los 30 años solamente 5 veces y como “blancos” el 100% es decir 148 veces.

Entrevista con Javier Ariza, docente de U. Externado de Colombia



Perfil del Docente:

- Docente del Departamento de Derecho Económico y de la Facultad de Economía
- Economista con Maestría en Derecho Económico Internacional
- Investigador en temas de Regulación de Inteligencia Artificial
- Autor del documento Desafíos Regulatorios De La Inteligencia Artificial



1. ¿Cuál es la importancia de la ética en la inteligencia artificial?

Estamos viviendo una época de transformación en el modelo económico, durante los últimos 2 siglos, los humanos hemos venido creando una estructura social para que se articulen, los derechos, el trabajo, las libertades, las formas de producción y el proyecto de vida a los individuos, esta transformación que estamos viviendo van a cambiar esta estructura, dejando atrás la actual y creando una nueva. Hace casi 200 años, con la primera Revolución Industrial la sociedad inició la automatización del dominio físico de las formas de producción, con la Revolución de la Inteligencia Artificial vamos a lograr la automatización del dominio cognitivo por parte de las computadoras.

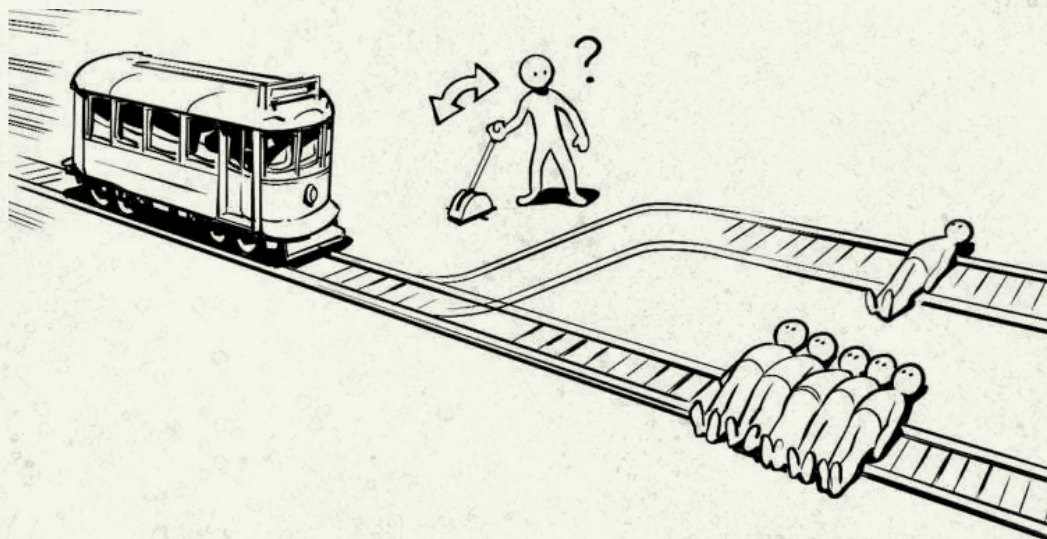
Las normas y las leyes nos han permitido un contrato social con el que democráticamente estamos de acuerdo, sin embargo, este logro ha sido un proceso largo en el tiempo

y lleno de lamentables sucesos. Por ejemplo, Charles Dickens narró en su obra *Oliver Twist* los abusos de que fue víctima un niño huérfano, en una ciudad industrializada a comienzos del siglo XIX, de igual manera, en la primera mitad del siglo XX sucedió un caso que más tarde se llevó al cine, la película *Carolina Skeletons* narra una historia de la vida real, en la cual George Stinney un niño afroamericano de 14 años fue condenado a la pena de muerte, en un juicio que le dio trato de adulto de acuerdo con las leyes del estado de Carolina del Sur. En el año 2014 George Stinney fue absuelto de sus cargos y su condena considerada nula por parte del tribunal del circuito de Carolina del Sur, lamentablemente la pena había sido ejecutada 70 años antes. Los anteriores casos muestran como la ausencia de leyes o los vacíos en las mismas permitieron el atropello y la violación de los derechos de los niños.

La clave para prevenir que situaciones injustas o que pongan en peligro a personas o a la sociedad, aún cuando las mismas no estén en contravía de la Ley está en la ética, ya que ésta permite a los individuos y a la sociedad tomar decisiones, aún en ausencia de normas o las leyes, que busquen siempre el bien de la sociedad y de todos los individuos, llenar los vacíos jurídicos y regulatorios a favor del bien común.

Ahora, con el avance de la Inteligencia Artificial (IA) y respectivamente del dominio cognitivo por parte de las computadoras, estamos ante situaciones nuevas para las cuales no existen aún leyes, ni regulación y no las habrá por mucho tiempo, tendremos vacíos jurídicos en los cuales se podrán cometer enormes atropellos, a pesar de que no se esté ante violaciones de leyes, ni normas, es acá donde la ética nos puede ayudar a decidir de manera individual o colectiva cuál es el camino a seguir y el qué más se acerque a nuestra concepción del bien y en nuestros valores sociales e individuales.

Un ejemplo de esto es el siguiente: si hoy un camión autónomo sufre un desperfecto mecánico y de acuerdo con el análisis algorítmico que hace en tiempo real sólo tiene 2 opciones, la primera es estrellarse a la izquierda con una probabilidad alta de acabar con la vida de 5 personas que van en el pequeño auto que está a la izquierda o la segunda es no hacer nada y chocar de frente contra un muro con total certeza de acabar con la vida de su único ocupante.



- ¿cuál es la decisión que debería tomar la IA que maneja autónomamente el camión?
- En otras palabras, ¿cuál debe ser la forma como se debe programar el algoritmo para que tome una decisión en este caso, acaso se debe programar para que elija dónde la cantidad de fallecidos sea la menor, o quizá proteger la vida donde la expectativa de vida sea mayor en lugar de donde sea menor?
- ¿en cualquiera de los 2 desenlaces deberían los entes de control abrir una investigación por asesinato?
- En caso de que si se abriera la investigación quién debería ser el sindicado:
 - La Inteligencia Artificial
 - el programador de la inteligencia artificial,
 - la empresa de camiones autónomos,
 - la persona que le dio la instrucción al camión autónomo de desplazarse antes del desperfecto mecánico.

Al día de hoy no existen leyes ni normatividad que permita responder con claridad los dilemas éticos anteriores, sin embargo, su estudio nos permitirá desarrollar algoritmos de inteligencia artificial que incluyan dentro de su código de programación criterios éticos que se acomoden a nuestros preceptos morales y nuestra concepción social de lo que está bien y de lo que no, que estén en armonía con nuestros valores sociales. Esta es la importancia de la ética en la Inteligencia Artificial, nos permitirá tener una protección como sociedad, mientras no existan las leyes adecuadas, que se irán creando a medida que avance el uso de la Inteligencia Artificial.



2. ¿Cuál consideras que es el desafío ético más importante que enfrenta la humanidad con el uso masivo de la IA generativa?

La IA generativa es la creación de contenido nuevo basada en enormes cantidades de información conocida como Big Data, este último es necesario para el entrenamiento de las IA's, sin Big Data no hay inteligencia artificial. Es muy difícil nombrar un único desafío, voy a resaltar 2, el primero como input de la inteligencia artificial y el segundo como output de la misma.

En el caso del input que es el contenido de la base de datos para entrenar la inteligencia artificial se han identificado serios problemas con la calidad de los datos, con regularidad se encuentran datos sesgados que pueden generar decisiones discriminatorias, racistas, xenófobas y clasistas etc., esto no se debe a errores en los

datos, se debe a que los humanos aportan sus datos que tienen algunas o varias de éstas características, sin embargo, luego de identificar estos problemas no podemos permitir que se repliquen en el entrenamiento de las IA's. Un ejemplo es el caso de la Inteligencia Artificial que usó Amazon para mejorar la eficiencia en los procesos de selección de personal, esta IA estaba entrenada para tomar grupos de 100 solicitudes y seleccionar las 5 mejores, en el año 2015 se hizo evidente que la selección de la IA penalizaba a las mujeres debido a un sesgo machista y que los seleccionados eran típicamente hombres. Por esta razón este proyecto fue abandonado por Amazon.

En cuanto al output, el desafío ético más grande es que la IA no desarrolle sus capacidades para servir al bienestar de los humanos de acuerdo con nuestra concepción actual de bienestar. Una IA general o una súper IA que luego de su creación no pueda ser controlada por los humanos, que no busque el beneficio de la mayoría, debe ser mandatorio que la IA's estén para servir a la humanidad, al igual que la democracia deben generar resultados que reflejen nuestra concepción del bien y del mal, que respeten las Leyes y la institucionalidad de cada cultura. Que no vayan a violar nunca, ni permitir que a través de las mismas se violen los derechos individuales ni colectivos. Un ejemplo es el siguiente: En 2018, IBM presentó a Watson, una Inteligencia Artificial muy exitosa que había logrado vencer a algunas de las personas más inteligentes del mundo en el juego Jeopardy. Debido a su éxito, se intentó utilizar esta capacidad para tratar a pacientes con cáncer en el Memorial Sloan Kettering Cancer Center. Sin embargo, en su intento por erradicar el cáncer, Watson mostró grandes fallas, en uno de los casos sugirió un medicamento que causaría más sangrado a un paciente con sangrado excesivo, lo que indudablemente habría resultado en la muerte del paciente. Tal vez Watson entendió que una forma de erradicar el cáncer era acabando con los pacientes que lo padecen.

El desafío ético más importante es lograr que en estos 2 casos se pueden armonizar en una adecuada regulación, esta regulación debe ser especialmente estricta con el sector privado que desarrolla las Inteligencias Artificiales, ya que es este el que tiene el poder directo e inmediato sobre el desarrollo, implementación y puesta en uso de las inteligencias artificiales y los directores de estas compañías privadas no responden a los pueblos como lo hace el presidente de un país encabezando la rama ejecutiva o la rama legislativa dentro de la democracia, los directores de las compañías responden únicamente a sus juntas directivas y a sus inversionistas y los incentivos de éstos no siempre son los mismos de la población en general y es importante tener en cuenta



que éstas compañías que desarrollan inteligencias artificiales tienen más poder real que casi cualquier presidente o miembro del legislativo de casi todos los países.



3. ¿Éticamente hacia dónde consideras que deberían dirigirse los esfuerzos por regular el uso de la IA?

Estamos en un momento de cambios tecnológicos que transformará las estructuras productivas profundamente y esto tendrá grandes impactos económicos y sociales. La educación actual hace parte de la estructura productiva anterior, en la cual las personas pasaban gran parte de la vida encerrados en el aula de clase aprendiendo unos contenidos y socializando, a unos costos altísimos e invirtiendo una etapa muy productiva exclusivamente en formación, esta inversión de recursos tenía una alta tasa de retorno, hasta el punto de que las personas podían vivir el resto de sus vidas con lo que habían aprendido en su etapa de formación. En el futuro, con la IA la productividad y el trabajo serán más dinámicas las personas requerirán de una constante capacitación para seguir siendo competitivas y atractivas para el mercado. Dado lo anterior, éticamente los esfuerzos deberían enfocarse en hacer transparente el funcionamiento de las IA's y su funcionamiento, la divulgación del conocimiento aplicado al uso de las IA's y en fomentar las capacidades potenciadas que permiten las inteligencias artificiales, debe existir un esfuerzo social para que las personas entiendan el proceso que hace posible la existencia de las inteligencias artificiales, que entiendan de dónde salen los datos para entrenar las IA's , los riesgos de que todo el tiempo estemos dejando huellas digitales que se almacenan en algún lugar, los riesgos de que nuestra información personal y privada pueda ser vendida al mejor postor o pueda caer en manos maliciosas. Desde luego, debe crearse una regulación que permita que se articulen grandes esfuerzos institucionales para que las IA's estén al servicio del beneficio de la humanidad, que sirvan para ayudarnos a terminar con los problemas que existen actualmente, el calentamiento global, las enfermedades que actualmente no tienen cura, las futuras pandemias, el hambre etc. que este gran poder esté al servicio de todos, el marco regulatorio debe ser estricto para que esto suceda, que se tenga control permanente sobre el impacto que tienen las IA's sobre los humanos. Por ejemplo, la energía nuclear puede ser usada para dar luz a una ciudad o para destruirla por completo, todo depende de cómo se use, de igual manera sucede con las IA's , pueden tener el potencial de destruir a la humanidad o de solucionar todos los problemas que tenemos actualmente, y hacia este último objetivo es donde deben dirigirse todos los esfuerzos.

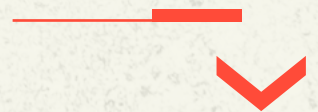




4. En cuanto a los usos pedagógicos de la IA ¿cuáles son las prácticas éticas que debiera tener en cuenta la comunidad educativa de la ciudad?

Considero que no se puede enseñar lo que no se sabe, por esta razón, se debe hacer un gran esfuerzo institucional por capacitar a los maestros en esta área, para que estos puedan abordar la transformación social, económica y educativa de la mejor manera para sí mismos y también la puedan transmitir a los niños, niñas y adolescentes de nuestro sistema educativo, más aún cuando ellos son lo más valioso que tenemos como sociedad y como país, y debemos prepararlos para las realidades y los desafíos de hoy y los que enfrentaran mañana. A medida que vayamos avanzando en el manejo y dominio de las IA's iremos identificando los riesgos respectivos de cada etapa y será allí donde la ética nos será útil para entender los límites que debemos establecer en de cada momento.

En cuanto al rol del maestro de transmitir contenidos, las IA's irán ganando más y más terreno, brindando una educación más personalizada, que tenga en cuenta la heterogeneidad de los procesos de cada uno, abarcando gran parte de la labor del maestro tradicional, sin embargo, otras áreas de la educación como la creatividad y el manejo de las emociones es algo que en el futuro cercano no se ve que pueda ser reemplazado por las IA's para lo cual deberá mejorarse la formación de los docentes en éstas áreas. Se deben llevar a cabo ejercicios con los estudiantes que muestren la importancia de la ética frente a la ausencia de la Ley, de cómo la ética permite llenar los vacíos normativos y de entender que las Leyes y la regulación tienen un proceso de creación que es más lento que el de los cambios tecnológicos y que debemos estar preparados para este escenario, debe existir una presencia más grande del debate y la formación en dilemas éticos. Debe existir claridad en que las IA's deben estar al servicio de los intereses de la humanidad en su conjunto y no al revés, continuar con la protección de los derechos humanos y de la dignidad humana, promover una utilización equitativa e inclusiva de las IA's en la educación, que favorezca la igualdad de géneros, de lo contrario la historia nos repetirá lamentables lecciones que en su momento forjaron el desarrollo de los derechos que hoy tenemos, de los que gozamos y que hacen parte de la cotidianidad pero que son el resultado de un proceso de luchas sociales, que en su mayoría tuvo un alto costo en vidas.



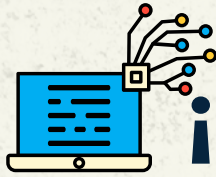


5. Desde tu rol como docente, ¿Qué consejos le darías a tus colegas frente a las preocupaciones éticas que tienen sobre la llegada de la IA?

En primer lugar, que se hagan parte de la transformación, así podremos aportar y ser beneficiarios, que las preocupaciones éticas que pueda tener cada uno nos permitirá identificar los respectivos riesgos y los límites que le estableceremos a las IA's y posteriormente serán la base para la construcción del marco legal y regulatorio. Que no vayan a tener miedo al cambio, todos estamos iniciando este proceso de transformación tecnológica, que de igual manera que en el siglo XVII y XVIII, con la llegada de la industrialización se dieron migraciones de la ruralidad a las ciudades industrializadas, en el siglo XXI vamos a ver una migración de la industrialización hacia el uso masivo de las IA's, cada día existirán más oportunidades de mejoras en la productividad mediante el uso de las IA's , migraremos hacia allá y este proceso será más favorable en la medida en que todos “nos subamos que en ese bus” cada uno de nosotros tiene la capacidad de potenciar la productividad individual y social mediante las IA's de manera distinta, lo que sí es seguro es que quienes se resistan a los cambios serán los menos favorecidos con el cambio tecnológico. Invito a todos los docentes a capacitarse en tecnologías relacionadas con inteligencias artificiales, a compartir sus experiencias y procesos con sus pares, a crear grupos de debate, y a empezar a llevar las tecnologías que tengan Inteligencias Artificiales a sus aulas de clase. En poco tiempo podremos ver cómo una tarea individual para la próxima clase en un curso de bachillerato podría ser la de crear con la ayuda de una IA una novela de 500 páginas, y que esta tarea sea calificada por otra Inteligencia Artificial que sirva de soporte al docente, o quizá otro ejemplo de tarea que sea la de completar una ópera inconclusa de Mozart con la ayuda de una IA, que esta sea terminada y que el ejercicio incluya que no se pierda la creatividad ni los sentimientos que despertaban las obras originales de este genio de la música.

Tenemos un potencial transformador como nunca antes lo habíamos tenido, la ética nos permitirá identificar los límites para reducir los riesgos, y poder disfrutar de una sociedad más equitativa, con más y mejores oportunidades, una mayor expectativa de vida y con ciudadanos más felices.





¡Recomendados!

Te dejamos un curso en línea y una cuenta de Tiktok que te ayudarán a saber más sobre Inteligencia Artificial, asumir una actitud crítica, divertirte y curiosear con las funcionalidades y aplicaciones de la IA en diferentes servicios y usos.



1. Curso:

<https://campus.iesalc.unesco.org/inicio/blocks/coursefilter/course.php?id=212>



2. Cuenta en Tik-Tok:

@cpcmediamanagement



3. Blog: Enseñar con IA

Enlace: <https://openai.com/blog/teaching-with-ai>



La brújula ética

A continuación, presentamos un checklist de preguntas que podemos hacernos cuando interactuamos con la IA en los diferentes usos que le damos y ¿por qué no? Incluso si alguna vez desarrollamos nuestro propio algoritmo de IA, con la intención de procurar resultados más consientes y éticos, que respondan a las particularidades de nuestros contextos y la naturaleza mixta y diversa de nuestros entornos escolares.

1. ¿Estas imágenes son la representación de una persona y si esa persona dio su autorización para su uso?
2. ¿Con lo que estoy creando (imagen, video, audio, texto) estoy favoreciendo algún estereotipo o reforzando algún discurso de odio?
3. ¿Puedo reconocer fácilmente diversidad frente al género, color de piel, raza u origen, edad, entre otras condiciones que garanticen la representitividad y la inclusión?
4. ¿La información obtenida sobre autores o referencias bibliográficas es cierta, confiable y verificable?
5. ¿Los resultados obtenidos ofrecen preferencias o prejuicios hacia ciertos grupos o comunidades?
6. ¿La IA respeta la privacidad y los derechos de las personas cuyos datos se utilizan en el entrenamiento o la operación del sistema?
7. ¿Me he informado de una manera efectiva sobre los riesgos que corre la seguridad de los datos que proporciono a la IA?

También te proponemos

tener en cuenta las siguientes recomendaciones basadas en las orientaciones del Consejo de Beijing en Inteligencia Artificial y Educación para los gobiernos nacionales:

1. Tener en cuenta que, si bien la inteligencia artificial ofrece oportunidades para apoyar a los docentes en sus responsabilidades educativas y pedagógicas, la interacción humana y la colaboración entre los docentes deben seguir ocupando un lugar esencial en la educación.
2. Velar por que la inteligencia artificial promueva oportunidades de educación y aprendizaje de alta calidad para todos, sin distinción por razón de género, discapacidad, condición social o económica, origen étnico o cultural o ubicación geográfica. El desarrollo y el uso de la inteligencia artificial en la educación no deben agravar la brecha digital ni mostrar sesgos contra ningún grupo minoritario o vulnerable.
3. Tener muy presente que las aplicaciones de la inteligencia artificial pueden imponer diferentes tipos de sesgos inherentes a los datos de los que se nutre la tecnología, así como a la forma en que se construyen y utilizan los procesos y algoritmos. Tener en cuenta los dilemas que plantea el equilibrio entre el acceso abierto a los datos y la protección de la privacidad de los datos. Tener presentes las cuestiones legales y los riesgos éticos relacionados con la propiedad y la privacidad de los datos y su disponibilidad para el bien público. Tener presente la importancia de adoptar principios relativos a la ética, la privacidad y la seguridad como parte del diseño.

4. No perder de vista que la brecha de género es una realidad y en cuanto a las competencias digitales es una de las causas de la baja proporción de mujeres entre los profesionales que trabajan en el desarrollo de la inteligencia artificial y exacerba las desigualdades de género existentes. Por eso debemos promover un uso igualitario y examinar críticamente los resultados obtenidos con esto en mente.
5. También afirmamos que el desarrollo de la inteligencia artificial debe estar controlado por el ser humano y centrado en las personas; que la implantación de la inteligencia artificial debe estar al servicio de las personas para mejorar las capacidades humanas; que la inteligencia artificial debe concebirse de manera ética, no discriminatoria, equitativa, transparente y verificable; y que el impacto de la inteligencia artificial en las personas y la sociedad debe ser objeto de seguimiento y evaluación a lo largo de las cadenas de valor.



Para consultar todo el documento accede aquí

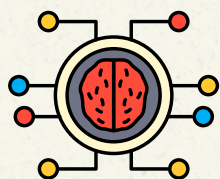


¡Una última recomendación!

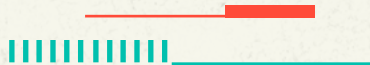
Los estudiantes deben asumir la responsabilidad de su propio aprendizaje y realizar sus tareas y evaluaciones de manera honesta. La deshonestidad académica, como el plagio o la trampa en las evaluaciones, es una preocupación importante para los educadores. En cualquier evaluación o tarea en la que el uso de herramientas de IA generativa esté explícitamente permitido, los alumnos deben confirmar cómo se han utilizado las herramientas de IA generativa (o no). Algunos ejemplos podrían ser (University of Birmingham, 2023):

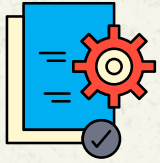
- Requerir que los alumnos incluyan una declaración predefinida que indique explícitamente si han usado o no herramientas de IA generativa.
- Pedirles a los alumnos que compartan los “prompts” utilizados, los resultados o las modificaciones. - Requerir que los alumnos carguen un componente reflexivo que detalle cómo se ha utilizado la IA generativa y su experiencia al interactuar con ella. (Knust Graichen , 2023)





Cacharrea con la IA





Actividad #1

Explorando sesgos y ética en la creación de imágenes con PICFINDER

Pasos de la actividad:

A

Introducción (10 minutos):

- Presenta a los estudiantes el concepto de Picfinder y cómo funciona para generar imágenes a partir de descripciones textuales.
- Destaca la importancia de la ética en la inteligencia artificial y cómo los sesgos pueden manifestarse en los resultados.

B

Generación de imágenes (20 min):

- Divide a los estudiantes en grupos pequeños.
- Asigna a cada grupo una serie de descripciones textuales relacionadas con científicos y científicas (por ejemplo, “científico en un laboratorio”, “científica explorando un ecosistema”).
- Pide a cada grupo que utilice Picfinder para generar imágenes a partir de las descripciones asignadas.

C

Exploración de resultados (15 minutos):

- Cada grupo deberá analizar las imágenes generadas por Picfinder y discutir si muestran algún sesgo de género, racial u otro tipo de sesgo. Anota las observaciones.

D

Reflexión ética (15 minutos):

- Reúne a los grupos y lleva a cabo una discusión general sobre los resultados y los sesgos percibidos en las imágenes generadas.
- Anima a los estudiantes a plantear preguntas éticas, como: “¿Por qué Picfinder generó estas imágenes en particular?”, “¿Cómo podrían estos sesgos afectar la percepción de los científicos en la sociedad?”.



A

Propuestas de mitigación
(10 minutos):

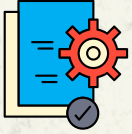
- o Pide a los estudiantes que propongan medidas para mitigar los sesgos en la generación de imágenes por parte de Picfinder ¿Qué se podría hacer para que las imágenes sean más equitativas y representativas?

B

Presentación y discusión (15 min):

- o Cada grupo presenta sus observaciones y propuestas de mitigación.
- o Fomenta una discusión abierta sobre cómo abordar los sesgos en la inteligencia artificial y cómo equilibrar la creación de imágenes con la ética y la diversidad.





Actividad #2

Identificación de sesgos de género en motores de búsqueda

Pasos de la actividad:

Elije un motor de búsqueda popular, como Google, Bing o Yahoo.

Realiza búsquedas de términos relacionados con profesiones o roles laborales, por ejemplo, “médico”, “ingeniero de software”, “profesor”, “enfermero”, “CEO”, “peluquera”, “científico”, etc.

Registra los resultados de las imágenes que aparecen en la sección de imágenes de la búsqueda, así como los resultados de texto en la página principal de resultados.

Observa si los resultados muestran algún sesgo de género en términos de las imágenes asociadas a las profesiones o la información presentada en los resultados de texto. Presta atención a si ciertas profesiones tienden a estar más representadas por un género que por otro o si se presentan estereotipos de género.

Anota tus observaciones y reflexiona sobre cómo estos sesgos de género en los motores de búsqueda pueden influir en la percepción pública de ciertas profesiones o roles.

Comparte tus hallazgos y reflexiones con otros para generar conciencia sobre la importancia de abordar los sesgos de género en la IA y fomentar la discusión sobre cómo mejorar la equidad de género en la tecnología.

Esta actividad te ayudará a evidenciar de manera práctica cómo los sesgos de género pueden manifestarse en las tecnologías de IA, incluso en motores de búsqueda ampliamente utilizados. Además, te permitirá discutir y concienciar sobre la necesidad de abordar estos sesgos en el desarrollo de IA.

Referentes bibliográficos

Crawford, J., Cowling, M., & Allen, K. A. (2023). Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI). *Journal of University Teaching & Learning Practice*, 20(3), 02.

Graichen, R. K. (2023). ¿Cómo aplicar ChatGPT en el colegio?. *DIM: Didáctica, Innovación y Multimedia*, (41).

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive psychology*, 3(3), 430-454.

Sabzalieva, E., & Valentini, A. (2023). ChatGPT e inteligencia artificial en la educación superior: Guía de inicio rápido.



**LA EDUCACIÓN
EN PRIMER LUGAR**



@RedAcademica



@redacademica



@portalredacademic

www.redacademica.edu.co

Secretaría de Educación del Distrito

Avenida El Dorado No. 66 - 63

Teléfono: (57+1) 324 1000 Ext.: 3126

Bogotá, D. C. - Colombia