

Inteligencia Artificial Ética

Introducción	1
Empleo y monopolización de aplicaciones IA	1
Invasión de nuestra privacidad e intimidad	2
Sesgos y discriminación	2
Pérdida de libertad individual	3
Malos usos de la IA	3
Deepfakes y desinformación	3
Manipulación política	4
Usos bélicos	4
Hacia una IA Ética	5
Regulación y leyes	5
Educación	5
Responsabilidad profesional	6
Conclusiones	6
Biografía	6

Introducción

Si algo es cierto sobre la IA es que está evolucionando a pasos agigantados. Aplicaciones que hace pocos años sonaban a ciencia ficción hoy en día ya son una realidad. El desconocimiento y el poco margen de tiempo que existe para ponerse al día sobre las nuevas tecnologías nos hacen vulnerables a sufrir posibles efectos negativos de la IA ya sea, o no, la intención del desarrollador.

Son muchos los aspectos en los que la IA puede afectarnos tanto positivamente como negativamente de aquí a corto/medio plazo. En este documento se pretende recoger los posibles efectos negativos que pueden causar los sistemas inteligentes y se invita a la reflexión de qué podemos hacer para intentar prevenirlo.

No se van a tratar aspectos más filosóficos (Ej. transhumanismo) que, sin duda causarían dilemas éticos y morales en la sociedad pero quizás no son tan urgentes de abarcar y quedan fuera del enfoque de este documento.

Empleo y monopolización de aplicaciones IA

Una preocupación muy extendida entre la población es la posibilidad de que los trabajadores sean reemplazados por robots en sus puestos de trabajo. También, existe mucha incertidumbre sobre qué tipo de estudios y qué capacidades deberían desarrollar la gente joven para poder encontrar trabajo el día de mañana.

Es cierto que muchas profesiones tal y cómo las conocemos hoy en día dejarán de existir y probablemente mucha gente pierda sus empleos porque a los empresarios les saldrá más barato y aumentarán la productividad sustituyendo parte de su plantilla por robots inteligentes.

Algo bastante similar pasó con la revolución industrial con la diferencia que esta vez no solo los empleos manuales y repetitivos están en peligro si no que también muchos de los empleos intelectuales. A mi parecer, esto no debería preocupar a la gente ¿Qué hay de malo en que los humanos disfrutemos de vacaciones eternas mientras que los robots trabajan para nosotros? El dinero no debería ser un problema ya que muchos de los servicios que hoy en día todavía pagamos a humanos a cambio de ellos, nos lo podría proporcionar un robot a precio mucho más bajo o incluso gratis. Creo que parece razonable que quien conserve un trabajo debería obtener más beneficios que alguien que no estuviera trabajando. Aun así, los humanos en vacaciones permanentes podrían sin mucho problema obtener un bienestar y calidad de vida parecido o mejor que cualquier persona de sueldo medio a día de hoy. Ahora bien, si esto suena a utopía es porque es una utopía. Si queremos llegar a ese escenario o a uno muy parecido, habrá que evitar que sean unos pocos los que se aprovechen de los beneficios que puede aportar la IA.

Pongamos como ejemplo un software médico basado en IA del cual se pueden hacer copias ilimitadas y a muy bajo costo debería ser también de muy bajo costo o gratis para la ciudadanía. Ahora bien, si una entidad privada se hace con el software y solamente da acceso a su servicio a aquellos que paguen un precio excesivamente caro, tendríamos un

problema. Las diferencias económicas de la población crecerían de forma exponencial haciendo cada vez más difícil darle la vuelta. Por eso es muy importante empezar cuanto antes a concienciar y a implantar políticas que castiguen el abuso de poder y que premien e incentiven el uso de la IA para mejorar la calidad de la población general y protejan a los más vulnerables.

Invasión de nuestra privacidad e intimidad

No se puede crear mucha IA sin datos por lo que inevitablemente vamos a tener que exponer nuestros datos personales si queremos aprovecharnos de las ventajas que nos puede aportar. Eso sí, habrá que luchar por que todas las entidades recolectoras de datos hagan un uso responsable y ético de los mismos. Habrá que cuidar que se mantenga el anonimato siempre que sea posible y que el uso de estos datos sea estrictamente para lo que su dueño ha permitido. Lo cual es imprescindible tanto la transparencia al usuario de lo que ocurre con sus datos así como la conservación del derecho de los usuarios a modificar o eliminar sus datos.

Afortunadamente, esto no va a ser ninguna novedad. La exposición de nuestra vida privada en la red ya es, desde hace unos años, un tema muy debatido en la sociedad. Los usuarios empezaron a preguntarse y preocuparse qué ocurría con su información de sus redes sociales y demás servicios online. Incluso la ley ya se ha tenido que adaptar y crear nuevas legislaturas para proteger a los usuarios. El dato cada vez es más valioso pero la sociedad cada vez es más consciente de ello y no va a dejarse engañar tan fácilmente.

Tanto la actualización y adaptación de la ley como la educación son clave para que cada usuario sea consciente de qué datos propios está proporcionando y que se está haciendo con ellos para valorar si decide ceder esos datos a cambio del servicio. Asimismo, se necesita generar conciencia social de lo importante y poco moral que es distribuir datos ajenos sin permiso.

Sesgos y discriminación

Ya se han visto casos en los que una IA ha sido racista o machista. En 2015, uno de los algoritmos de Google ya les jugó una mala pasada etiquetando como “gorilas” a una pareja de afroamericanos en Google Photos. Más recientemente, ha vuelto a haber polémica con unos algoritmos del traductor de Google por ser sexista. Se viralizaron unas traducciones donde se asumía el género al traducir de lenguas sin género a lenguas con género según la frase a traducir.

Muchos justifican estos algoritmos diciendo que solo son un reflejo de la sociedad, y tienen razón. Eso no quita que no se deba intentar conseguir que nuestras IAs tengan menos prejuicios que nosotros los humanos. Sobre todo, porque si no lo hacemos, la IA va a intensificar todos estos sesgos y teniendo en cuenta que la mayoría no van a cuestionar lo que la IA diga y que es probable que nuestros hijos se vayan a educar a través de software

inteligente, solo se va a conseguir que la sociedad se haga más y más discriminatoria, realimentando todo el sistema.

Conseguir no trasladar los prejuicios humanos a las máquinas no es una tarea sencilla y para ello es clave cerrar (o al menos reducir) la brecha digital. Solo si participan personas de todas las razas, edades y géneros podremos eliminar los sesgos en las aplicaciones que desarrollemos. Por esta razón, es primordial que todo el mundo tenga oportunidad de acceder a una educación digna y de calidad.

Pérdida de libertad individual

Los humanos realizamos decisiones basadas en la razón, pero también basadas en las emociones. No siempre acertamos pero de los errores, generalmente, aprendemos o nos abren la puerta a nuevas vivencias que de otra manera nunca habríamos descubierto. En definitiva, tenemos la libertad y el permiso de no escoger siempre las opciones que, a priori, parecen ser las mejores condicionando así nuestra capacidad de decidir.

Con la llegada de una IA que cada vez eligen más por nosotros, perdemos esas oportunidades de ser espontáneos, de actuar en contra de lo que se dicta. Ya lo podemos ver cuando Spotify decide por ti qué música te apetece escuchar o cuando Google Maps decide por ti cual es la mejor ruta de un punto A a un punto B. Resulta muy cómodo y útil en muchas ocasiones pero a la larga puede impactar en la sociedad haciéndonos cada vez más vagos y dependientes en las decisiones tomadas por un software. También, nos volvemos más vulnerables a la manipulación de las grandes entidades y además, perdemos un poco la capacidad y seguridad en uno mismo de pensar y decidir.

Es importante ser conscientes de esto y procurar no dejarnos llevar por las decisiones de una máquina en todos los aspectos. Lo más peligroso aquí es cuando deciden por nosotros sin que nos demos cuenta.

Malos usos de la IA

Hasta aquí se han analizado los posibles peligros que puede tener la IA a pesar de que el desarrollador busque el bien común y no tenga intenciones de causar un mal en la sociedad. Estos riesgos se intensifican y además surgen otros nuevos cuando los desarrolladores de una IA buscan únicamente beneficiarse ellos mismos sin importar el impacto negativo en la sociedad.

Deepfakes y desinformación

Desde hace décadas somos conscientes de que la prensa que recibimos puede estar presentada de tal manera que favorezca la postura política o los intereses del medio de comunicación. Desde la exageración de las cifras, una frase sacada de contexto o hasta manipulando ligeramente la realidad. Ya estamos acostumbrados a desconfiar de todo esto pero cada vez se vuelve más difícil de diferenciar lo que es real de lo que no.

Cada vez son más realistas los videos o audios falsos, más conocidos como deepfakes, donde vemos cosas que no han pasado o escuchamos decir frases que nunca se dijeron por la persona que parece decirlas. Y no solo esto, también, cada vez es la propia IA quien redacta mejor, más rápido y con un mayor poder de convicción las propias noticias.

Aquí también podríamos incluir todo lo relacionado con las estafas y la suplantación de identidad. Con esta tecnología de los deepfakes es muy fácil engañar o extorsionar al prójimo para obtener dinero a cambio.

Aparte de seguir educándonos en este aspecto para poner en duda todo lo leído, visto y escuchado, también nos queda la consolación de que el desarrollo no va solo a favor de la desinformación. Al igual que cada vez hay una mejor IA generando deepfakes, también hay otras IA que cada vez son mejores en detectar lo que es real de lo que está generado con IA.

Manipulación política

Seguro que os suena el escándalo causado por Cambridge Analytica y Facebook en las elecciones de EEUU y en el referéndum del Brexit durante el año 2016. Se utilizaron datos personales de Facebook para hacer publicidad dirigida y mostrar noticias falsas a usuarios que se categorizaron como manipulables para que terminaran votando a Donald Trump o a favor del Brexit.

En este revuelo entran bastantes temas de dudosa eticidad, como por ejemplo la privacidad de los datos personales, la propagación de las fake news y la pérdida de la libertad individual. Pero, vamos a aparcar todos estos temas a un lado porque ya se han recogido anteriormente en otros apartados y vamos a centrarnos en lo que se consigue cuando se juntan todos estos elementos; la manipulación política.

Creo que se tendría que cuestionar hasta qué punto un ciudadano que se dirige a la urna con su voto lo ha elegido libremente o lo han manipulado para que crea que es decisión suya. En cierta medida, siempre ha habido personas a las que han lavado la cabeza con propaganda pero con esta nueva tecnología las reglas del juego cambian, todavía no están reguladas y lo que es peor, en general la sociedad no entiende qué se está haciendo.

Una vez más, la educación es clave para tener conciencia de lo potentes que son estas prácticas y cómo pueden manipular tu propia forma de ver el mundo.

Usos bélicos

Es evidente que la IA va a suponer muchos avances en las armas militares. Los países están invirtiendo millones en el desarrollo de la IA en este ámbito porque nadie se quiere quedar atrás. No solo abarataría gastos en la organización y la logística si no que también ganarían ventaja en las batallas bélicas. Un vehículo aéreo autónomo capaz de identificar amenazas, de atacar y de organizarse y colaborar con otras unidades ¿Qué cuerpo militar no querría esto?

Partiendo de la idea de que el uso de la IA en armas no las hace menos éticas de lo que ya son, sí hay cuestiones que toca cuestionar. Por ejemplo, quién es la persona o entidad que se hace responsable de las muertes de un arma que ha decidido atacar por sí sola ¿Y que hay de las muertes y daños ocasionados por una IA que ha actuado por error?

La Unión Europea ya emitió una directriz sobre este aspecto en diciembre del 2020. Se indicaba que “los sistemas autónomos letales deben estar sujetos al control humano”. Esto quiere decir que detrás de un sistema autónomo con capacidad de matar, debe haber un humano responsable de la última decisión que toma el sistema inteligente.

Hacia una IA Ética

Aunque ya se han ido mencionando distintas formas de prevenir, o de al menos mitigar las posibles consecuencias negativas que puede traernos la IA, no está de más recogerlo en una sección propia.

Regulación y leyes

Es complicado definir y recoger un conjunto de leyes que sean capaces de evitar todo el mal que las nuevas tecnologías puedan causar. Siempre surgen nuevas formas de encontrar vacíos legales o de quebrantarla de tal forma que sea muy difícil detectar la infracción. Por no hablar de que las leyes no son iguales en todos los países y que en el mundo de las telecomunicaciones es muy fácil tener servidores y controlarlos desde el otro lado del mundo para burlar las restricciones.

La legislación hace esfuerzos para regular y limitar los usos de las tecnologías emergentes. Unos ejemplos de ello es la ley de protección de datos que entró en vigor el 5 de mayo de 2018 o la actualización constante del reglamento Europeo de drones entre muchas otras. Incluso en España se ha redactado “La carta de Derechos Digitales” para proteger a los usuarios. Es posible que no sean medidas suficientes pero si son necesarias y sin duda hay que seguir persiguiendo y sancionando y regulando las malas praxis digitales.

Educación

Probablemente la mejor arma que podamos tener para ganar esta batalla sea una educación que garantice:

- Valores que primen el bienestar general de la población.
- El pensamiento crítico.
- Conciencia sobre la privacidad y el dato personal.
- Conocimientos técnicos de matemáticas, estadística, probabilidad e inteligencia artificial.
- Cerrar la brecha digital (ser accesible por toda la población indistintamente de la edad, sexo, condición, raza, lugar de nacimiento o situación económica).

Los sistemas educativos de todo el mundo necesitan adaptarse a esta nueva era. Quizá ya no sea tan importante que los estudiantes pasen tanto tiempo memorizando la tabla periódica o repitiendo la lista de vocabulario que va a entrar en el próximo examen (y van a olvidar en cuanto salgan por la puerta) y sea más productivo enseñarles a pensar por sí mismos. Ya existen (y cada vez existirán más) softwares con IA con los que aprenderán las materias tradicionales de forma mucho más efectiva, divertida y dirigida.

También, precisamente gracias a este auge de la tecnología, cada vez debería ser más fácil poder llevar a todos los rincones del mundo una educación de calidad. Si se consiguiera tener una población mundial educada en las nuevas tecnologías, capaz de pensar de forma crítica y con unos valores acordes a una IA Ética, se podría aprovechar al máximo los beneficios que la IA puede proporcionarnos sin que nadie salga perdiendo.

Responsabilidad profesional

Los profesionales de este campo, tenemos la responsabilidad individual y colectiva de garantizar una IA ética en cada aplicación que desarrollemos. Como bien se dice, un gran poder conlleva una gran responsabilidad y no podría ser más cierto en el campo de la ciencia de datos.

Una vez más destaco la educación como clave para garantizar la IA ética. Los futuros responsables de cómo afectará la IA a la sociedad están en este momento formándose bajo el sistema educativo que tenemos ahora mismo y que probablemente no esté todavía adaptado a los nuevos tiempos.

Conclusiones

La IA puede proporcionarnos muchos beneficios y bienestar pero tenemos que ser conscientes de varios aspectos. De lo contrario, nos arriesgamos a que mientras una minoría de la población esté disfrutando de las ventajas de las nuevas tecnologías, la gran mayoría verá su calidad de vida cada vez más perjudicada. Viviríamos en una sociedad cada vez más injusta.

Para evitar esto, la clave está en una educación adaptada a esta nueva era. Una educación donde no solo se enseñen los aspectos técnicos y matemáticos de las ciencias de datos pero que también se incentive el pensamiento crítico y se inculque unos valores acordes a la IA Ética.

Biografía

- [Google Fotos etiqueta a dos afroamericanos como gorilas](https://www.mycomputer.com)
(<https://www.mycomputer.com>)
- [El sesgo de género del Traductor de Google persiste: Ella cose, él conduce](https://computerhoy.com)
(<https://computerhoy.com>)

- No es solo porno: los Deep Fakes y su potencial uso para manipulación política (<https://www.xataka.com>)
- Entrevista a Cristopher Wylie (<https://elpais.com>)
- Cómo operó Cambridge Analytica en el Brexit: la otra elección manipulada con los datos de Facebook (<https://www.infobae.com/>)
- Drones y bombas que 'hablan': la IA es la gran revolución militar, y nadie está al mando (<https://www.elconfidencial.com>)