

16-8-2021

# TAREA 50: INTELIGENCIA ARTIFICIAL ÉTICA



Alain Irastorza Azanza  
ALAINIRASTORZA@GMAIL.COM  
THE EGG

# Índice

INTRODUCCIÓN .....	3
1. ÉTICA Y TECNOLOGÍA .....	4
2. Ideas fundamentales de ética en IA .....	6
3. Conclusiones.....	9
Bibliografía.....	10

## INTRODUCCIÓN

El siglo XXI está siendo caracterizado por numerosos desarrollos tecnológicos y avances en la ciencia. La tecnología se trata de una aplicación práctica de la ciencia y no puede considerarse neutra. En este proceso participan los profesionales, que toman las decisiones relacionadas con la tecnología, y los usuarios o la sociedad civil, que utiliza la tecnología.

En la aplicación de la tecnología se pueden dar numerosos problemas éticos que afectan negativamente a la sociedad civil. Ejemplo de ello es la difusión de bulos durante las elecciones presidenciales de Estados Unidos a través de redes sociales como Facebook. Para evitar estas situaciones, los profesionales deben incorporar principios éticos al diseño, desarrollo y uso de las aplicaciones tecnológicas.

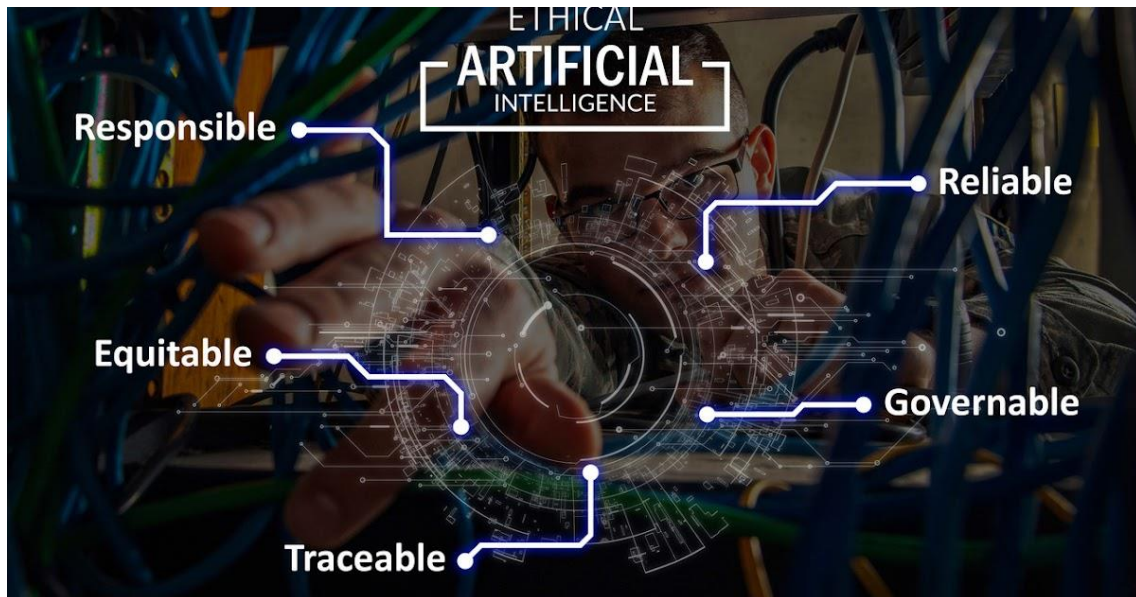
En la Inteligencia Artificial también se deben incorporar los principios éticos. Existen una serie de conceptos muy críticos en la IA: los datos o informaciones que se utilizan pueden ser justos o no, pueden tener sesgos o discriminar...

A continuación, se presenta una reflexión acerca de la importancia de la ética en la Inteligencia Artificial, así como las ideas fundamentales en la Inteligencia Artificial Ética.

Asimismo, se presentan una serie de principios o reglas que se deben respetar para la IA se desarrolle a partir de unos principios éticos.

## 1. ÉTICA Y TECNOLOGÍA

El desarrollo de nuevas tecnologías como la Inteligencia Artificial conlleva tanto impactos sociales como implicaciones éticas. Por poner un ejemplo, existen aplicaciones en teléfonos móviles que registran la localización o ubicación del usuario y otros datos personales. Esta se trata de una información delicada que podría ser de interés para por ejemplo diferentes compañías ofrecer sus servicios o diseñar campañas de publicidad específicas para el usuario. [1]



*Ilustración 1. Desafíos para la IA*

Por una parte, se debe considerar la finalidad del producto ofertado por una empresa en el mercado. Una empresa debe desarrollar su propio código ético para abordar posibles problemas éticos relacionados con su producto. Por ejemplo, una red social como Facebook debe garantizar que se protegen adecuadamente los datos personales protegidos de cada usuario.

Por otra parte, se deben considerar todos los posibles usos que el cliente le puede dar al producto o a la tecnología. Por ejemplo, Facebook debe limitar la difusión de bulos y Fake News en su red social.

Principalmente, y de manera resumida, se pueden diferenciar dos enfoques filosóficos en la ética:

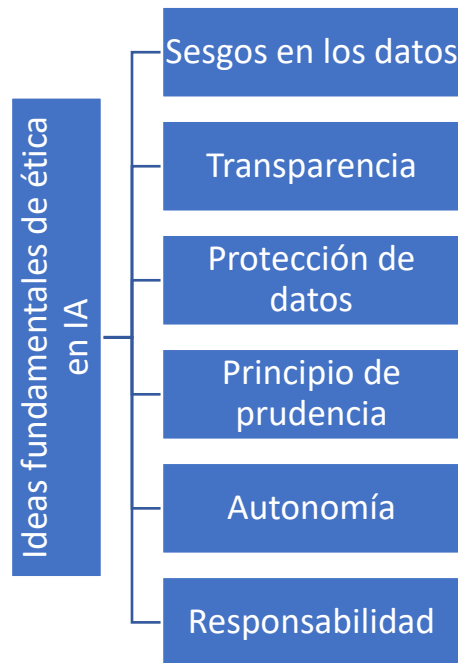
- A) **Utilitarismo.** Se trata de alcanzar la mayor felicidad o utilidad posible para el mayor número de individuos o personas. Para decidir si una acción es buena o mala se calcula el valor de todos los placeres y dolores probables que podría causar, por lo que se tienen en cuenta las consecuencias de las acciones. [2]

Si existe un tren que ha perdido una rueda y se mueve en dirección a un túnel donde matará a cinco trabajadores, pero si se desvía solo matará a uno. Según el utilitarismo, para en este caso minimizar el daño se está dispuesto a sacrificar la vida de un trabajador desviando la dirección del tren para salvar a cinco trabajadores.

- B) **Enfoque Kantiano.** Según el filósofo Immanuel Kant ciertas cosas son inviolables debido a que son contrarias a la dignidad humana y nunca deben hacerse. Para que se pueda realizar una acción se debe de aplicar a todas las personas sin resultar contradictoria.

## 2. Ideas fundamentales de ética en IA

A continuación, se comentan las principales ideas que se deben considerar y aplicar a la hora de desarrollar la IA desde un punto de vista ético.



*Ilustración 2. Ideas fundamentales de ética en IA*

### 1- Sesgos en los datos.

Considerando que los algoritmos del Machine Learning requieren de datos históricos, existe el peligro de volver a reproducir los clásicos problemas de discriminación. Por ejemplo, si los estudios de criminalidad existentes en Estados Unidos indican que las personas de origen afroamericano disponen de mayor probabilidad de cometer delitos, los algoritmos reflejarán esta información sesgada. [3]

Para evitar este problema, es necesario definir un protocolo que regule los sesgos en la Inteligencia Artificial.

### 2- Transparencia o explainability

Los algoritmos del IA deben ser capaces de explicar las causas de sus decisiones y los criterios seguidos en la toma de decisiones. Los algoritmos deben rendir cuentas a las personas afectadas por sus decisiones. Por ejemplo, si un banco deniega un crédito a un cliente utilizando un algoritmo de IA, el cliente tiene derecho a saber porque se le ha denegado el crédito.

### 3- Protección de datos

En la IA se requiere una base de datos de la que partir, el cual habitualmente contiene información sensible, personal... cuya privacidad debe ser protegida. Mientras que en Europa existe un Reglamento General de Protección de Datos, existen países con gran presencia de IA como China donde no existen reglamentos que protegen la información de los usuarios en la misma medida.

No obstante, actualmente existen numerosas lagunas en legislaciones de IA: es esencial que todas las empresas que gestionen nuestra información nos informen siempre de la utilización de nuestros datos. Además, se deben regular con una mayor precisión y considerando todas las casuísticas las relaciones entre humanos y máquinas. Existen robots que se utilizan por ejemplo para tratar enfermedades neurodegenerativas estableciendo relaciones emocionales con humanos con consecuencias beneficiosas para el paciente; mientras, según los reglamentos un robot no debe establecer relaciones emocionales con un humano.

### 4- Principio de prudencia y fiabilidad

Cuando se desarrollan y lanzan a mercado aplicaciones de IA se deben considerar todas las posibles consecuencias a usuarios y otros terceros. Estas aplicaciones se deben verificar, evaluar y certificar adecuadamente. Además, los sistemas con tecnologías de IA deben someterse a pruebas que garanticen su fiabilidad y seguridad.

Hay numerosas tecnologías de IA con usos sustancialmente beneficiosos para el usuario y la sociedad mientras las aplicaciones se utilizan de manera justa. Pero, si falta transparencia o se realiza un uso ilegal o impidiendo los derechos humanos básicos, pueden aparecer problemas muy graves.

Un ejemplo de ello, es la tecnología de reconocimiento facial, el cual ha sido muy criticado por sus usos en las revueltas de Hong Kong.



*Ilustración 3. Tecnología de reconocimiento facial*

## **5- Límites de la autonomía de la tecnología**

Cuando se trata de una tecnología que pueda tener un funcionamiento autónomo, que sea capaz de tomar decisiones por sí sola, se deben estudiar todas las posibles situaciones y consecuencias, sobre todo en casos en los que una decisión u otra pueda causar daños a humanos y a terceros.

Por ejemplo, un arma autónoma si no es capaz de distinguir un soldado en ataque, un soldado rendido, un soldado herido y un civil, se está delegando una decisión de matar en una máquina con unas consecuencias nefastas.

Además, en el caso del ejemplo del utilitarismo del apartado anterior, si un tren puede seguir su camino matando a cinco obreros o desviarse matando a un obrero, sería una duda de gran importancia determinar qué decisión debería de tomar un algoritmo y como.

En este sentido, se plantea que los ingenieros de IA deban firmar un juramento de buenas prácticas o que los algoritmos se deban certificar por entes externos. Sea la que sea la solución, se debe garantizar la imparcialidad y la correcta aplicación de los principios éticos en la IA.

## **6- Responsabilidad**

Cuando una máquina causa daños a una persona, se debe definir a quien le corresponde la responsabilidad de los daños y perjuicios causados. Se debe decidir si la máquina es la responsable (la empresa o persona dueña de la máquina), la empresa que comercializa la máquina, la empresa que realiza el mantenimiento de la máquina, la empresa o ingeniero de IA que diseña la máquina o el código... [4]



### 3. Conclusiones

Analizando lo recogido en los apartados anteriores se puede concluir que es esencial diseñar y aplicar unos reglamentos que abarquen los posibles problemas derivados del uso de IA. En estos reglamentos deben participar diferentes sujetos y agentes:

- Las escuelas de IA deben formar en ética a los ingenieros de IA para que los sujetos que diseñan los algoritmos apliquen los principios éticos.
- Los desarrolladores de IA (pueden ser empresas o ingenieros) deben garantizar que sus desarrollos cumplen con un marco ético mediante un juramento u otro sistema. Además, se debe verificar por entes independientes y externos que los desarrollos de IA cumplen con las legislaciones y marcos éticos.
- Los gobiernos de los estados y Europa deben ir un paso más allá en legislaciones y diseñar legislaciones que consideren todas las casuísticas posibles de IA. Asimismo, se debe primar la protección de la sociedad que pueda ser afectada por la IA.
- Además, más allá de Europa, organizaciones universales deben garantizar que las empresas que diseñen o comercialicen productos con IA en cualquier lugar del mundo deben cumplir con un marco ético que proteja a los usuarios y a la sociedad y no deje en desventaja o sin marco legal y ético a los ingenieros de IA de otros países.

Por último, si se desea garantizar un desarrollo adecuado, justo y ético de la IA se deben considerar los principios éticos explicados en el apartado número dos. Los ingenieros de IA, las organizaciones y empresas, las asociaciones, los estados... son los primeros interesados en ese adecuado desarrollo necesario para que la sociedad confíe en las ventajas de la IA y esta tecnología contribuya a mejorar el mundo.

## Bibliografía

- [1] “Ética, tecnología y el futuro de la humanidad”, OMPI, Organización Mundial de la Propiedad Intelectual, online, (acceso el 16-08-2021), [https://www.wipo.int/wipo\\_magazine/es/2018/04/article\\_0005.html](https://www.wipo.int/wipo_magazine/es/2018/04/article_0005.html)
- [2] “Utilitarismo, la filosofía práctica de Jeremy Bentham”, publicado el 19 de diciembre de 2018, filosofía&co, online, (acceso el 16-08-2021), <https://www.filco.es/utilitarismo-filosofia-de-bentham/>
- [3] “Por qué debería preocuparte la ética de la inteligencia artificial”, publicado el 12 de diciembre de 2019, Sinc, online, (acceso el 16-08-2021), <https://www.agenciasinc.es/Reportajes/Por-que-deberia-preocuparte-la-etica-de-la-inteligencia-artificial>
- [4] “La importancia de la ética en la inteligencia artificial”, El País, publicado el 26 de febrero de 2019, online, (acceso el 16-08-2021), [https://elpais.com/retina/2019/02/25/tendencias/1551089772\\_654032.html](https://elpais.com/retina/2019/02/25/tendencias/1551089772_654032.html)