

#50 Inteligencia Artificial Ética

Necesidad de hablar de principios éticos

Desde que en el año 1979 Hans Jonas publico “El principio de responsabilidad. Ensayo de una ética para la civilización tecnológica” [1], la necesidad de considerar el aspecto ético de la tecnología ha sido constante por parte de distintos ámbitos. Citando textualmente a Amparo Alonso [2] “el cambio tecnológico implica también un cambio social para el que no tenemos normas definitivas que hay que ir consensuando la mismo tiempo que avanzamos”. Hoy en día sin embargo, con el aumento de aplicaciones de IA en nuestro día a día, la necesidad de consensuar y crear un marco ético en relación con la IA es ya crucial y ha quedado claro que el avance en lo tecnológico tiene que ir unido a la definición del marco.

Principios éticos

De esta manera, se trata de consensuar y determinar los principios que tiene que cumplir cualquier IA y valorar cómo se consigue que así sea y definir métodos para asegurar que se cumplen estos principios. Estos principios y reglas tienen que ser tanto para la ética pública y la integridad institucional como para los profesionales, empresas y ámbitos de investigación [3].

En la mayoría de las fuentes relacionadas con el tema hay varios conceptos, principios e ideas que se repiten, aunque en muchos casos los matices no sean exactamente los mismo [1-7]. He tratado de agrupar estas ideas en los siguientes principios que toda IA tiene que respetar:

- Transparencia y explicabilidad: el código abierto facilita la transparencia de los algoritmos, en contra del enfoque de “caja negra” que puede poner en duda la ética de los sistemas. Tendría que ser posible explicar con claridad que es el sistema de IA, cómo funcionará y qué resultados obtendrá. Esto también está relacionado con la previsibilidad del sistema.
- Responsabilidad y rendición de cuentas: para que un sistema de IA sea ético tiene que quedar claro la obligación de rendir cuentas y de la responsabilidad de las acciones de la IA.
- Robustez y seguridad: el sistema de IA tiene que asegurar la seguridad, la fiabilidad y la solidez para operar de manera precisa y segura y evitar o resolver errores.
- Justicia y no discriminación, imparcialidad, no estigmatización, no sesgo, diversidad: contemplar todos los grupos de interés y garantizar la utilización justa evitando discriminaciones.
- Gobernanza de datos y privacidad: como la base de los sistemas de IA son los datos, también hay que asegurar la privacidad de datos y que se cumplan las leyes vigentes de protección de datos.
- Bienestar social y ambiental y sostenibilidad: por un lado tiene que perseguir el bienestar social (haciendo el bien) y por otro lado, relacionado con la eficiencia energética

Aunque a veces no sea fácil el consenso en lo que a principios éticos, el objetivo será asegurar el cumplimiento de los principios éticos esenciales, como lo son los que aparecen en la Carta de los Derechos Fundamentales de la Unión Europea [UE].

Cumplimiento de los principios éticos

Para asegurar que cumple con los principios listados anteriormente, toda IA debería de poder auditarse y asegurar que se cumple el marco ético. Esto se puede hacer de distintas maneras, pero en [3] se diferencian dos tipos de métodos alternativos: métodos técnicos y métodos no técnicos. Los métodos técnicos incluyen que la propia IA sea explicable y también algoritmos de control que aseguren los principios éticos de la IA a evaluar, también llamados “guardianes IA” [5] En los métodos no técnicos se incluyen la regulación, la certificación, la educación y sensibilización y la investigación.

Bibliografía

Artículos:

[1] Terrones Rodríguez, Antonio Luis (2018). Inteligencia artificial y ética de la responsabilidad. Cuestiones de Filosofía, 4 (22), 141-170.

[2] <https://www.iic.uam.es/innovacion/como-avanzamos-una-inteligencia-artificial-etica/>

[3] Marín García, Sergio (2019). Ética e inteligencia artificial. Cuadernos de la Cátedra CaixaBank de Responsabilidad Social Corporativa.

[4] UE (2000). Carta de los Derechos Fundamentales de la Unión Europea (2000/C 364/01).

[5] Cotino Hueso, Lorenzo. (2019). Ética en el diseño para el desarrollo de una inteligencia artificial, robótica. Revista Catalana de Dret Públic (58), 29-48

[6] <https://es.unesco.org/news/etica-inteligencia-artificial-paso-mas-adopcion-recomendacion-unesco>

[7] <https://www.muyinteresante.es/tecnologia/inteligencia-artificial/articulo/cuatro-puntos-clave-para-entender-la-etica-de-la-inteligencia-artificial-211604936018>