

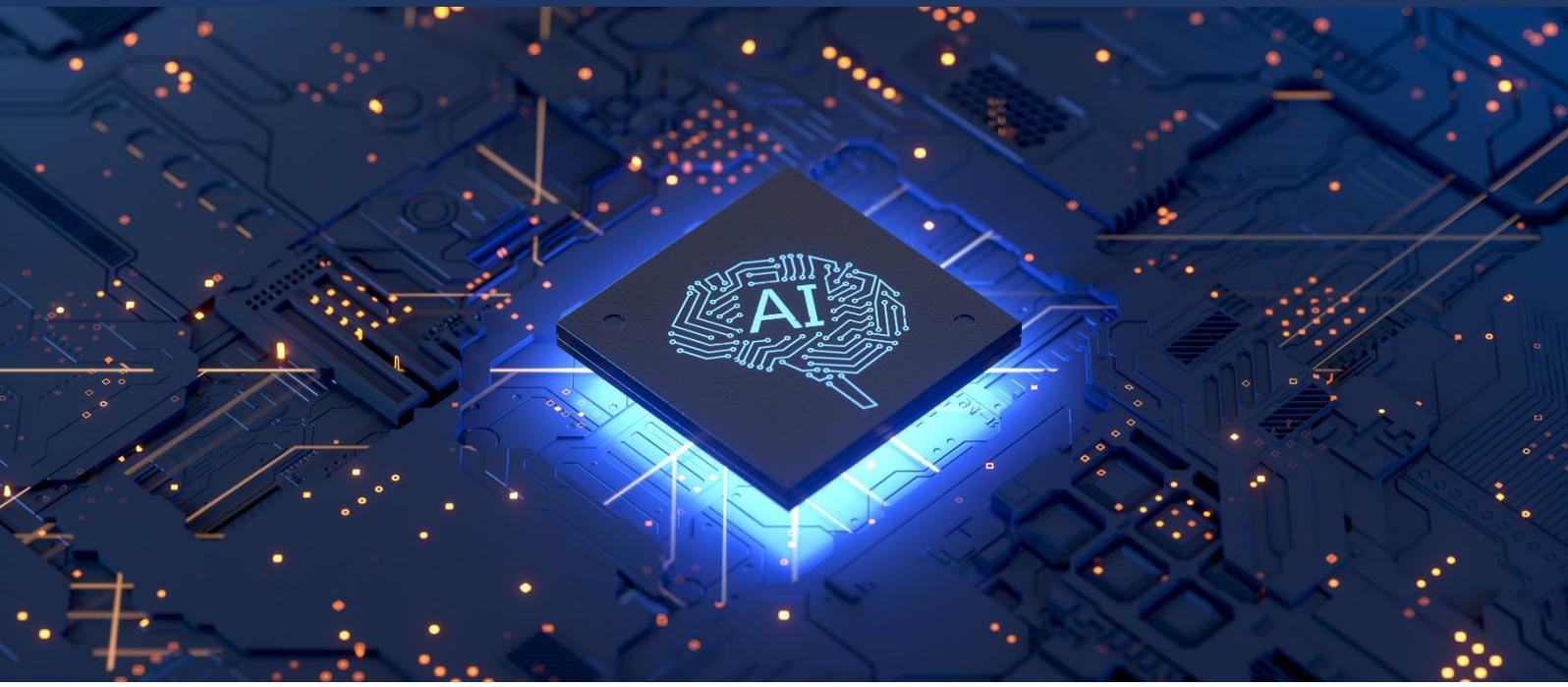


GOBIERNO
DE ESPAÑA

MINISTERIO
DE CIENCIA, INNOVACIÓN
Y UNIVERSIDADES

Comité Español de Ética de la Investigación

Recomendaciones éticas para la investigación en inteligencia artificial





El Comité Español de Ética de la Investigación es un órgano colegiado de ámbito estatal, independiente, y de carácter consultivo, adscrito al Consejo de Política Científica, Tecnológica y de Innovación, que ejerce competencias en materia de integridad científica, investigación responsable y ética en la investigación científica y técnica.

El Pleno del Comité, en su reunión de 22 de noviembre de 2023, compuesto por las personas que a continuación se expresan, aprobó estas recomendaciones éticas para la investigación en inteligencia artificial:

Jordi Camí Morell
Nuria Sebastián Gallés
Encarna Aguayo Jimenez
Rafael Yuste Rojas
M.ª Amparo Alonso Betanzos
María del Val Bolívar Oñoro
Coral Calero Muñoz
Antonio Diéguez Lucena
M.ª Carmen Fernández Criado
Pere Puigdomènech Rosell
Leonor Ruiz Sicilia

El Comité Español de Ética de la Investigación tiene los derechos sobre este documento, el cual es confidencial y no será usado para ningún otro propósito, salvo para el que fue suministrado y no será reproducido, copiado o transmitido, en todo o en parte, sin el permiso de su dueño.

The Spanish Research Ethics Committee owns the copyright of this document which is supplied in confidence and which shall not be used for any purpose other than that for which it is supplied and shall not, in whole or in part, be reproduced, copied, or communicated to any person without permission from the owner.

Secretaría del Comité Español de Ética de la Investigación

P.º de La Castellana, 162. 28046 Madrid.

oooc@ciencia.gob.es

La Inteligencia Artificial (IA) es en este momento una tecnología esencial para la economía mundial, presentando un crecimiento exponencial anual en inversión. De hecho, según se indica en el informe del I foro INDESIA, la aplicación de los datos y la inteligencia artificial en la industria española tendrá un impacto estimado en el PIB de 16.500 millones de euros en el 2025. La IA es además una disciplina transversal, y hoy en día existen desarrollos y aplicaciones en prácticamente cualquier campo de trabajo. Debido a esto, todos los países de la UE, además de otros países en el mundo como es el caso de EE. UU. China o Canadá, han publicado Estrategias en Inteligencia Artificial (IA) que tienen como objetivo convertirse en países líderes en dicha tecnología. La IA no solamente afecta a la evolución tecnológica, sino también a la economía, el empleo, la educación, y en general a la sociedad y su organización. Por ello, la ética en la investigación en IA debería ser la piedra angular de la innovación y el desarrollo tecnológico responsables.

Existen un número importante de informes en ética e IA, como el del Grupo de Expertos de Alto nivel en IA (HLEG-AI) de la UE, [1] o el de la UNESCO [2], que indican la necesidad de una reflexión ética en esta tecnología. Sin embargo, se presta menos atención a los aspectos éticos de la investigación que se lleva a cabo en el campo, ya que los sistemas pueden contener sesgos aún inadvertidos, violar la privacidad de las personas, posibilitar la manipulación y la vigilancia exhaustiva de los individuos, etc. En nuestro caso, en este documento debemos centrarnos en los desafíos especiales que pueden surgir en la **investigación en IA**, y a las cuestiones que se plantean con respecto a la responsabilidad de la investigación en este contexto.

En concreto, es de nuestro interés definir cuál es la responsabilidad de los investigadores y de las instituciones de investigación. Por esta razón, sería relevante para este Comité el poder disponer de una guía más concreta en cuanto a la **investigación en IA** (no en cuanto a la investigación con herramientas y desarrollos de IA). En este sentido, es de referencia el informe del Comité de Ética en la investigación en Ciencia y Tecnología de Noruega [3]. Por otra parte, hemos de tener en cuenta también la inclusión de los derechos presentes en la Carta de Derechos Digitales, en donde también se engloban los neuroderechos [6]. Algunas cuestiones más generales en cuanto a la investigación en tecnología están también presentes en los códigos éticos de sociedades profesionales como IEEE o ACM [4,5].

En base a lo anteriormente expuesto, este Comité considera que sería conveniente que los trabajos de **investigación en IA** sigan un código ético que incluya los aspectos destacados a continuación. Siguiendo la filosofía del documento publicado por el Comité noruego, los desafíos más relevantes para la ética en la investigación en IA podrían clasificarse en tres grandes grupos:

1. RESPONSABILIDAD EN EL DESARROLLO Y EL USO DE SISTEMAS AUTÓNOMOS

Este grupo de desafíos se centraría en la capacidad que tienen los sistemas de IA para realizar toma de decisiones de forma automática, aunque sea siempre bajo supervisión humana, como recomiendan las Directrices Éticas para una IA Fiable, publicada por la Comisión Europea mediante el Grupo Independiente de Expertos de Alto nivel sobre IA en 2018, y las de la UNESCO de 2021. Los aspectos más concretos relacionados con la investigación en IA serían:

1.1.- Salvaguardar los derechos humanos y fundamentales

Los investigadores y las instituciones de investigación deben asegurarse de que los sistemas de IA están estructurados de forma que se salvaguarde la dignidad, la autodeterminación y los derechos humanos y democráticos. Los investigadores deben de considerar el impacto posible o esperado de los sistemas de IA en las personas, los seres vivos y la sociedad en general, y deben de facilitar su uso ético.

1.2. – Salvaguardar el medioambiente

Los investigadores y las instituciones de investigación deben asegurarse de que los sistemas de IA se desarrollan teniendo en cuenta su consumo energético y consiguiente impacto medioambiental. Así los investigadores deben de considerar el consumo energético y uso de otros recursos de las soluciones que desarrollan, buscando el equilibrio entre la solución suficientemente óptima y el consumo requerido por la misma.

1.3.- Responsabilidad

Los investigadores deben proporcionar guías en las decisiones y acciones de los sistemas de IA, colaborando con sus instituciones y con los actores externos implicados para asesorar el riesgo del uso de su investigación, si se plantea la comercialización de los sistemas.

1.4.- Inspeccionabilidad y trazabilidad

Es necesario poder identificar las fuentes de datos usados y/o generados por el sistema, así como la forma en la que éste toma decisiones. De esta forma se puede asegurar la confianza y la justicia de las decisiones realizadas por el mismo de forma automática. Los sistemas de IA deben dejar de ser “cajas negras” para poder ser inspeccionados, debiendo los investigadores declarar y justificar cualquier falta o dificultad en este sentido.

1.5.- Divulgación de la investigación

Actualmente es un reto asegurar una discusión balanceada sobre los riesgos y las oportunidades de la IA. Como investigadores se debe evitar la ingenuidad, siendo conscientes de los riesgos y oportunidades de la tecnología, y siendo responsables de presentar un balance equilibrado entre ambos, dado que se tiene un conocimiento más profundo e informado del nivel de desarrollo de estos.

Los investigadores por lo tanto deberían contribuir al debate público, para que la sociedad pueda construir opiniones basadas en asunciones realistas.

2. CONSECUENCIAS SOCIALES Y RESPONSABILIDAD SOCIAL DE LA INVESTIGACIÓN

La IA tiene un potencial enorme para generar cambios en prácticamente cualquier área de aplicación. Las respuestas de la tecnología varían en función de su interacción con individuos, colectivos y entornos. Por ello, existe cierta impredecibilidad, siendo difícil prever su efecto en individuos, sociedad y medioambiente.

Desde este punto de vista, en la investigación sería recomendable:

2.1.- Reconocer la incertidumbre

Es esencial asesorar y comunicar adecuadamente la incertidumbre asociada con la investigación, ya que en el caso de la IA además de incertidumbre existe impredecibilidad. En este sentido, sería deseable facilitar la investigación interdisciplinar para intentar minimizar ambas, además de recomendar la inclusión de la Ética y los Derechos Humanos como materia en la educación de los futuros desarrolladores de IA.

2.2.- Asegurar una participación amplia.

Los investigadores tienen la responsabilidad de comunicar los riesgos que se derivan a partir de los resultados de sus investigaciones. Los riesgos y oportunidades que se enfatizan pueden depender también de la perspectiva ética individual o colectiva, y de los valores e intereses en los que se basan.

Por ello, aquellos que resulten más afectados por las decisiones que se tomen deberían tener voz en los procesos de toma de decisiones. Las autoridades y las instituciones de investigación deberían facilitar la participación de los ciudadanos en los debates sobre el propósito de la investigación, la estructura de los programas de investigación y la aplicación de la investigación.

2.3.- Asegurar la protección de datos y la consideración a los individuos

Deben seguirse los principios fundamentales de la protección de datos, recogidos en la legislación nacional y europea pertinentes, y en el futuro en la regulación europea al respecto (AI Act y Data Act). Desde el punto de vista de la ética de la investigación, el consentimiento es la regla básica cuando en la investigación se hace uso de datos personales. Pero cuando se recogen y compilan grandes cantidades de datos, hay un riesgo de que los datos personales puedan usarse de formas en las que no hemos sido conscientes, así como un riesgo de desvelado de esos datos personales, aunque estos estén anonimizados, debido a cruces con otras fuentes de datos.

Los investigadores tienen la responsabilidad de asesorar el grado de acceso público a la información, la sensibilidad de esta, la vulnerabilidad de los posibles afectados y el impacto y consecuencias en la investigación.

3. DATOS

El uso de datos, especialmente de grandes cantidades de datos, es otra de las fuentes de riesgos éticos en la investigación en IA. La calidad de los datos, la seguridad y la fiabilidad, el acceso a los mismos, la posible presencia de sesgos o la posibilidad de manipulación personal y vigilancia exhaustiva son cuestiones a tener en cuenta por investigadores e instituciones de investigación.

3.1.- Aseguramiento de la calidad

En la investigación en IA, puede haber una razón particular para realizar preguntas críticas acerca de la calidad, verdad y relevancia de los datos, ya que no siempre son conocidas sus fuentes, y los metadatos pueden ser inciertos o estar ausentes.

Materiales sesgados, herramientas de análisis con ciertas propiedades y las propias interpretaciones humanas aumentan las posibilidades de llegar a conclusiones erróneas, aumentando la incertidumbre en las interpretaciones y decisiones basadas en IA. Para asegurar la verificabilidad y la calidad, los investigadores y las instituciones de investigación deben facilitar el acceso a fuentes de datos abiertas y disponibles públicamente.

3.2.- Acceso justo a los datos

Desde el punto de vista de la ética de la investigación, es esencial asegurar que ésta, junto con los datos usados y los resultados obtenidos, está disponible abiertamente. En IA existe el riesgo de que algunas partes de la investigación evadan los requisitos de transparencia que se aplican en general a la investigación (ej. los principios FAIR), en algunos casos como pueden ser el secreto de ventaja competitiva.

Los gobiernos y las instituciones de investigación deben facilitar el acceso público a los datos, asegurando transparencia en aspectos como la propiedad de la tecnología, infraestructura y datos, priorización de áreas de investigación y justificación de esta, así como posibles beneficiarios.

REFERENCIAS

- [1] Grupo de Expertos de Alto nivel en IA, CE. Directrices éticas para una IA fiable. Junio de 2018. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- [2] UNESCO. Recommendations on the Ethics of Artificial Intelligence”, noviembre de 2021.
- [3] The Norwegian National Research Ethics Committees (NETI). The National Committee for Research Ethics in Science&Tecnology. Statement on Research Ethics in Artificial Intelligence. Noviembre de 2019. [balassets/dokumenter/4-publikasjoner-som-pdf/statement-on-research-ethics-in-artificial-intelligence.pdf](https://www.neti.no/balassets/dokumenter/4-publikasjoner-som-pdf/statement-on-research-ethics-in-artificial-intelligence.pdf)
- [4] Code of Ethics and Professional Conduct, ACM (2018). Traducción en español en <https://www.acm.org/about-acm/code-of-ethics-in-spanish>.
- [5] Code of Ethics IEEE. Traducción en español en https://edu.ieee.org/ec-ups/wp-content/uploads/sites/266/CODIGO_DE_ETICA_IEEE.pdf.
- [6] Ministerio de Asuntos Económicos y Transformación Digital. Carta de Derechos Digitales. https://portal.mineco.gob.es/RecursosArticulo/mineco/ministerio/participacion_publica/audiencia/ficheros/SEDIACartaDerechosDigitales.pdf.