

RESPONSIBLE AI R&D THEMATIC AREA NARRATIVE IN ENGLISH ARABIC FRENCH PORTUGUESE SPANISH

Rachel Adams , Kelly Stone

Rachel Adams , Kelly Stone

©2025, RACHEL ADAMS , KELLY STONE



This work is licensed under the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/legalcode>), which permits unrestricted use, distribution, and reproduction, provided the original work is properly credited. Cette œuvre est mise à disposition selon les termes de la licence Creative Commons Attribution (<https://creativecommons.org/licenses/by/4.0/legalcode>), qui permet l'utilisation, la distribution et la reproduction sans restriction, pourvu que le mérite de la création originale soit adéquatement reconnu.

IDRC GRANT / SUBVENTION DU CRDI : - GLOBAL INDEX ON RESPONSIBLE ARTIFICIAL INTELLIGENCE

Índice Global sobre IA Responsable

Dimensión: Capacidades nacionales para una IA responsable

Subdimensión: Inversiones

Área temática: I+D en IA responsable

Definiciones

El concepto [investigación y desarrollo](#) (I+D) se puede definir como una serie de iniciativas que las organizaciones (del sector público o privado) llevan a cabo para innovar e introducir nuevos productos o servicios, a menudo con el objetivo de mejorar su oferta. La I+D suele representar la [primera fase](#) del proceso de desarrollo y es de carácter exploratorio, en el sentido de que es independiente de las actividades operativas de una empresa y, por eso, no tiene una orientación inmediatamente lucrativa.

La [IA responsable](#), tal y como generalmente se entiende, se centra en el diseño, desarrollo y despliegue de la IA de una forma segura, fiable y ética. Los defensores de la IA responsable abogan por la adopción generalizada de principios y prácticas éticos de IA, para garantizar que el desarrollo de las tecnologías de IA esté centrado en el ser humano, sea interpretable y explicable, y para establecer sistemas que promuevan la transparencia, la fiabilidad y la equidad.

Por consiguiente, la I+D en IA responsable puede entenderse como las actividades que las entidades, incluidas las organizaciones del sector público y privado, llevan a cabo para explorar nuevas tecnologías e innovaciones en IA, alineadas y en consonancia con los principios de la IA responsable.

Justificaciones

Aunque las tecnologías de la IA tienen la capacidad y el potencial de transformar las sociedades para mejor, es de todos sabido que estas tecnologías también pueden introducir nuevos riesgos y acentuar los daños existentes si no se diseñan, desarrollan y despliegan de forma responsable y ética. Los creadores y usuarios de las tecnologías de IA no siempre son conscientes de las potenciales amenazas que los sistemas y herramientas de IA pueden representar para las personas, las comunidades, las sociedades y el medio ambiente, por lo que es esencial garantizar que se inviertan recursos en I+D para una IA responsable. Esto es necesario para garantizar que las tecnologías de IA respeten los principios de la IA responsable, es decir, los [cinco](#)

[Principios de la IA de la OCDE](#) basados en valores y la [Recomendación de la UNESCO sobre la Ética de la IA](#) (Recomendación). Por ello, la inversión en I+D responsable, en materia de IA, se considera un área política fundamental para los países, en sus esfuerzos por aprovechar el potencial de la IA en el interés público y cumplir con sus agendas de desarrollo, como las establecidas en los Objetivos de Desarrollo Sostenible (ODS). Los [Principios de la IA de la OCDE](#) piden a los gobiernos que consideren la inversión pública a largo plazo e incentiven la inversión privada en I+D, incluidos los "esfuerzos interdisciplinarios, para estimular la innovación en una IA confiable, que se centre en problemas técnicos desafiantes y en implicaciones sociales, legales y éticas relacionadas con la IA y asuntos de política".¹ Además, los Principios de la IA de la OCDE también piden a los gobiernos que consideren la inversión pública y privada en conjuntos de datos (datasets) abiertos que sean "representativos y que respeten la privacidad y la protección de datos",² como parte de la creación de un entorno propicio para la I+D en IA responsable, que pueda ser accesible para las distintas partes interesadas, incluidos desarrolladores, innovadores y agentes no estatales.

Además, la [Recomendación de la UNESCO](#) también identifica la I+D en IA responsable como una de las principales áreas políticas en la que los Estados Miembros deben avanzar para crear un entorno propicio para una IA ética. En concreto, la recomendación insta a los Estados Miembros a que apoyen la investigación en materia de IA, de conformidad con el derecho internacional y los principios éticos, "mediante inversiones en este tipo de investigaciones o creando incentivos para que los sectores público y privado inviertan en este ámbito", dado que "las investigaciones contribuyen de manera significativa al desarrollo y la mejora ulteriores de las tecnologías de la IA". La recomendación también estipula que los Estados Miembros "deben velar por que cualquier evolución futura relacionada con las tecnologías de la IA se base en investigaciones científicas rigurosas e independientes" y promover la investigación interdisciplinaria en IA, mediante la inclusión de disciplinas distintas de la ciencia, la tecnología, la ingeniería y las matemáticas (CTIM), como los estudios culturales, "la educación, la ética, las relaciones internacionales, el derecho, la lingüística, la filosofía, las ciencias políticas, la sociología y la psicología".

Por lo tanto, las inversiones en I+D de IA deben guiarse por políticas que articulen un conjunto claro de prioridades que se ajusten a los principios de una IA responsable.

Identificaciones

Esta área temática examina las medidas adoptadas por los diferentes países para invertir en I+D en IA responsable. En concreto, el análisis debe tener en cuenta: (1) los **marcos jurídicos** relativos a políticas, programas o directrices que definan la financiación y las prácticas para una I+D responsable en materia de IA, (2) las **acciones gubernamentales** de apoyo a la I+D en IA responsable, que invierten en programas centrados en la integración de sistemas de IA responsable, en todos los departamentos, agencias y organismos gubernamentales, en la creación de redes de investigación, de fondos o concursos centrados en I+D en IA responsable, y (3) los **agentes no estatales** que trabajan para hacer avanzar la I+D en IA responsable,

¹ Traducción no oficial

² Traducción no oficial

esforzándose en promover el desarrollo de tecnologías de IA para hacer avanzar los derechos humanos, crear redes de profesionales e investigadores que trabajen sobre I+D en IA responsable, y por concienciar sobre la inversión en I+D en IA responsable.

Los *marcos jurídicos* del país pueden asumir la forma de políticas adoptadas, libros blancos o directrices. Las *acciones gubernamentales* pueden incluir proyectos de políticas o directrices, la creación de entidades gubernamentales, como organismos de supervisión, encargados de proporcionar recomendaciones políticas sobre I+D responsable en IA y/o garantizar la asignación de fondos para la I+D responsable en IA. Además, las acciones pueden implicar la aplicación de políticas destinadas a abordar la cuestión y/o a sensibilizar o recopilar más datos sobre la I+D responsable en materia de IA. Los *Agentes No Estatales* (ANE) pueden ser organizaciones no gubernamentales (ONG), pero también empresas multinacionales, organizaciones militares privadas, medios de comunicación, grupos étnicos organizados, instituciones académicas, grupos de presión, sindicatos o movimientos sociales que trabajan para promover la inversión en I+D en materia de IA responsable.

Ejemplos:

Marcos jurídicos

La [política de investigación e innovación de la Comisión Europea, en materia de IA](#) se centra en estas cuatro prioridades: (1) desarrollar y desplegar soluciones de IA que tengan un impacto positivo en la sociedad y en la economía; (2) aumentar y dar prioridad a la inversión pública y privada (**incluyendo un mejor acceso y uso de los datos científicos**); (3) promover el desarrollo de una **IA de confianza**, incentivando la ética desde el diseño (ethics-by-design) en los proyectos de investigación e innovación del programa Horizonte Europa; y (4) financiar proyectos de investigación e innovación sobre IA que apoyen la transición industrial.

Acciones gubernamentales

En 2020, la Comisión Europea [invirtió 1.500 millones de euros en IA](#), con el objetivo de movilizar los recursos del sector público y privado para acelerar el despliegue de la IA por parte de diversas partes interesadas, incluidas las pequeñas y medianas empresas. Para finales de 2020, la Comisión pretendía aumentar la inversión total en IA del sector público y del privado, a 20.000 millones de euros anuales, a fin de apoyar la financiación continua de I+D en IA, a través de [Horizonte Europa](#) (el programa de financiación de I+D de la UE) y los programas Europa Digital, en toda la región.

Agentes no estatales

El [Instituto Internacional de Estocolmo para la Investigación de la Paz](#) (SIPRI) es un instituto independiente de investigación internacional, con sede en Suecia, especializado en conflictos, control de armamento y desarme. El SIPRI proporciona datos, análisis y recomendaciones a responsables políticos, investigadores, medios de comunicación social y público en general sobre tendencias emergentes, amenazas e innovación en materia de seguridad internacional, incluyendo la evolución de la IA. El SIPRI recibió el apoyo de la Comisión Europea para realizar investigaciones sobre: (a) cómo puede utilizarse la IA en sistemas convencionales, cibernéticos y nucleares; (b) cómo el uso militar de la IA puede crear riesgos humanitarios, pero también estratégicos, y oportunidades para control de armamento y verificación de las exportaciones; y (c) cómo pueden gestionarse los riesgos que plantea la IA a través del

derecho internacional, el proceso de control del armamento y la [investigación e innovación responsables](#).