

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 a IA Responsável

Dimensão: Capacidades nacionais em matéria de IA responsável

Sub-dimensão: Investimentos

Área temática: [Investigação e desenvolvimento em IA responsável](#)

Definições

A atividade de [investigação e desenvolvimento](#) (I&D) pode ser definida como uma série de iniciativas que as entidades (do setor público ou privado) realizam para inovar e introduzir novos produtos ou serviços, muitas vezes com o objetivo de melhorar as suas ofertas existentes. A I&D é normalmente a [primeira fase](#) do processo de desenvolvimento e é de natureza exploratória, na medida em que está separada das atividades operacionais de uma empresa e, por isso, não está imediatamente orientada para os lucros.

A [IA responsável](#), tal como é geralmente entendida, é uma abordagem à conceção, ao desenvolvimento e à implantação da IA de uma forma segura, fiável e ética. Os defensores da IA responsável advogam a adoção generalizada de princípios e práticas éticos de IA para garantir que o desenvolvimento de tecnologias de IA seja centrado no ser humano, interpretável e explicável, e para estabelecer sistemas que promovam a transparência, a fiabilidade e a justiça.

A I&D em IA responsável pode ser entendida, portanto, como as atividades que as entidades, incluindo as organizações dos setores público e privado, realizam para explorar novas tecnologias e inovações em matéria de IA que estejam alinhadas e em conformidade com os princípios da IA responsável.

Fundamentação

Embora as tecnologias de IA tenham a capacidade e o potencial de transformar as sociedades para melhor, é do conhecimento geral que estas tecnologias podem também introduzir novos riscos e exacerbar os danos existentes se não forem concebidas, desenvolvidas e implantadas de uma forma responsável e ética. Os criadores e utilizadores de tecnologias de IA nem sempre estão conscientes das potenciais ameaças que os sistemas e as ferramentas de IA podem representar para os indivíduos, as comunidades, as sociedades e o ambiente, razão pela qual é essencial garantir o investimento de recursos na I&D em IA responsável. Isto é necessário para garantir que as tecnologias de IA respeitem os princípios da IA responsável, nomeadamente os [cinco Princípios da IA da OCDE](#) baseados em valores e a

[Recomendação da UNESCO sobre a Ética da IA](#). Por conseguinte, o investimento em I&D no domínio da IA responsável é considerado um domínio político fundamental para os países nos seus esforços para tirar partido do potencial da IA no interesse público e para cumprir as suas agendas de desenvolvimento, como as estabelecidas nos Objetivos de Desenvolvimento Sustentável (ODS). Os [Princípios da IA da OCDE](#) apelam aos governos para que considerem o investimento público a longo prazo e incentivem o investimento do setor privado em I&D, incluindo "esforços interdisciplinares, para estimular a inovação em matéria de IA fiável que se concentre em questões técnicas difíceis e nas implicações sociais, jurídicas e éticas relacionadas com a IA e nas questões políticas¹". Além disso, os Princípios da IA da OCDE também apelam aos governos para que considerem o investimento dos setores público e privado em conjuntos de dados abertos que sejam "representativos e respeitem a privacidade e a proteção de dados", como parte da criação de um ambiente propício a uma I&D em IA responsável que possa ser acessível a várias partes interessadas, incluindo criadores, inovadores e intervenientes não estatais.

A [Recomendação da UNESCO](#) também identifica a I&D em IA responsável como um dos principais domínios políticos em que os Estados-Membros devem avançar no sentido de criar um ambiente propício a uma IA ética. Especificamente, a recomendação apela aos Estados-Membros para que apoiem a investigação em matéria de IA, em conformidade com o direito internacional e os princípios éticos, "pelo investimento em tais pesquisas ou pela criação de incentivos para que os setores público e privado invistam nessa área", uma vez que "a pesquisa contribui de forma significativa para o desenvolvimento e a melhoria das tecnologias de IA". A recomendação estipula também que os Estados-Membros "devem assegurar que avanços futuros relativos a tecnologias de IA tenham como base pesquisas científicas rigorosas e independentes" e que essa investigação deve promover abordagens interdisciplinares, recorrendo a conhecimentos especializados fora das disciplinas tradicionais da ciência, tecnologia, engenharia e matemática (STEM), e convide a contribuições dos domínios da "educação, ética, relações internacionais, direito, linguística, filosofia, ciência política, sociologia e psicologia".

Por conseguinte, os investimentos em I&D no domínio da IA devem ser orientados por políticas que articulem um conjunto claro de prioridades que se alinhem com os princípios de uma IA responsável.

Identificações

Esta área temática avalia as medidas adotadas pelos países para investir na I&D em IA responsável. Em particular, devem ser identificados: (1) os **quadros jurídicos** relativos a políticas, programas ou diretrizes que definam o financiamento e as práticas para uma I&D em matéria de IA responsável, (2) as **ações governamentais** de apoio à I&D em IA responsável, investindo em programas centrados na integração de sistemas de IA responsável em todos os departamentos, agências e funções governamentais, criando redes de investigação, fundos ou concursos centrados na I&D em IA responsável, e (3) os **intervenientes não estatais** que trabalham para fazer avançar a I&D em IA responsável, empenhando-se em promover e apoiar o desenvolvimento de tecnologias de IA para fazer avançar os direitos humanos, criar redes de profissionais e

¹ Tradução nossa

investigadores que trabalhem sobre a I&D em IA responsável, e os esforços de sensibilização para o investimento em I&D em IA responsável são uma prioridade.

Os quadros jurídicos no país podem assumir a forma de políticas, livros brancos ou diretrizes. *As ações governamentais* podem incluir projetos de políticas ou diretrizes, a criação de entidades governamentais, como órgãos de supervisão, com a missão de fornecer recomendações políticas sobre a I&D em IA responsável e/ou garantir a alocação de financiamento para a I&D em IA responsável. Além disso, as ações podem envolver a execução de políticas destinadas a abordar a questão e/ou a promover a sensibilização ou a recolher mais dados sobre a I&D em IA responsável. *Os intervenientes não estatais* (INE) podem ser organizações não governamentais (ONG), mas também empresas multinacionais, organizações militares privadas, meios de comunicação, grupos étnicos organizados, instituições académicas, grupos de pressão, sindicatos ou movimentos sociais que trabalham para promover o investimento em I&D em IA responsável.

Exemplos:

Quadros jurídicos

A [política de investigação e inovação da Comissão Europeia em matéria de IA](#) centra-se nas seguintes quatro prioridades: (1) desenvolver e implantar soluções de IA que tenham impactos positivos na sociedade e na economia; (2) aumentar e dar prioridade aos investimentos públicos e privados (incluindo um melhor acesso e uma melhor utilização dos dados científicos); (3) promover o desenvolvimento de uma **IA de confiança**, incentivando a ética desde a conceção (ethics-by-design) nos projetos de investigação e inovação do Horizonte Europa; e (4) financiar projetos de investigação e inovação em IA que apoiem a transição industrial.

Ações governamentais

Em 2020, a Comissão Europeia [investiu 1.5 mil milhões de euros em IA](#), com o objetivo de mobilizar recursos dos setores público e privado para acelerar a implantação da IA por uma série de partes interessadas, incluindo as pequenas e médias empresas. Até ao final de 2020, a Comissão pretendia aumentar o investimento total em IA, tanto no setor público como no privado, para 20 mil milhões de euros por ano, a fim de apoiar o financiamento contínuo da I&D em IA através do [Horizonte Europa](#) (o programa de financiamento da investigação e inovação da UE) e dos programas Europa Digital em toda a região.

Intervenientes não estatais

O [Instituto Internacional de Investigação para a Paz de Estocolmo](#) (SIPRI) com sede na Suécia, é especializado em conflitos, controlo de armas e desarmamento. O SIPRI fornece dados, análises e recomendações aos decisores políticos, investigadores, meios de comunicação social e público em geral sobre as tendências emergentes, as ameaças e as inovações em matéria de segurança internacional, incluindo a evolução da IA. O SIPRI foi apoiado pela Comissão Europeia para realizar investigação sobre: (a) a forma como a IA pode ser utilizada em sistemas convencionais, cibernéticos e relacionados com forças nucleares; (b) a forma como a utilização militar da IA pode criar riscos humanitários, mas também estratégicos, e oportunidades para o controlo de armas e a verificação das exportações; e (c) a forma como os riscos colocados pela IA podem ser geridos através do direito internacional, do processo de controlo dos

armamentos e de uma [investigação e inovação responsáveis](#).