



Research Brief

Inteligência Artificial e Direitos Humanos



RESEARCH BRIEF

Inteligência artificial e direitos humanos

Ana Cláudia Albergaria



Background / Enquadramento

O projeto de investigação «Inteligência artificial e direitos humanos» insere-se no âmbito de uma Bolsa de Doutoramento FCT/IS-UP (UI/BD/150781/2020) e tem como principal objetivo o estudo dos impactos da Inteligência Artificial (IA) nos direitos humanos, através da identificação dos principais desafios e obstáculos que se têm colocado à criação e à utilização de sistemas de IA responsável. Pretende propor, com base na auscultação e participação dos principais *stakeholders*, linhas orientadoras para a ação e estratégias alternativas de prevenção e superação dos riscos.

A Inteligência Artificial (IA) surge como objeto de análise sociológica dada a sua multidimensionalidade e consequências que provoca ao criar mudanças profundas nas sociedades e nos indivíduos, a uma velocidade e profundidade incomparável com outras tecnologias que a precederam. Trata-se de um fenómeno social que exige um olhar transdisciplinar, na medida em que afeta, de forma sistémica e transversal, as diferentes dimensões da vida em sociedade. Assim, através de uma estratégia de investigação-ação, propusemo-nos conhecer as representações e disposições de atores de diferentes áreas disciplinares sobre a *IA responsável*, incluindo os peritos em IA, da área tecnológica, e outros *stakeholders* de disciplinas diversas como a filosofia, o direito, a educação, economia social e terceiro sector e política; bem como empresas de tecnologia, comunicação social, entre outros.

Problematizar este fenómeno no sentido da construção de uma *IA responsável* ou ética corresponde também ao equacionamento do papel da designada “inovação e da investigação responsáveis” (Von Schomberg, 2013) no contexto da atual revolução tecnológica e à identificação de pistas de ação para o desenvolvimento de IA que não coloque em risco os direitos humanos fundamentais.



RESEARCH BRIEF

Inteligência artificial e direitos humanos

Ana Cláudia Albergaria



Metas e Objetivos de Investigação

Este projeto, em curso desde dezembro de 2020, visa promover o conhecimento e a reflexão sobre *IA responsável*, propondo os seguintes objetivos específicos: 1) identificar os principais desafios e potencialidades da IA na relação com os Direitos Humanos; 2) identificar as estratégias da União Europeia no que concerne às políticas públicas para a promoção da *IA responsável*; 3) identificar as representações sociais sobre IA e sobre os seus impactos nas sociedades em geral e nos direitos humanos em particular; 4) identificar estratégias de prevenção de riscos decorrentes da utilização de sistemas inteligentes; 5) promover a participação ativa de grupos multidisciplinares na reflexão/ação sobre *IA responsável*; 6) disseminar o conceito e o conhecimento sobre *IA responsável*.

Beneficiadores finais / Público-alvo

Este trabalho dirige-se, desde logo, a toda a comunidade e, em particular, aos investigadores, peritos e técnicos cujas áreas de interesse passam por temáticas como era digital, sócio tecnologia, riscos sociais, desigualdades e exclusão social, direitos humanos, políticas sociais, investigação e inovação responsável, educação e cidadania. Considerando, ainda, o crescente destaque mediático que a IA vem assumindo e os interesses económicos, sociais e políticos de que a mesma é alvo, este estudo é útil para um público cada vez mais abrangente, podendo por isso considerar-se a comunicação social e o terceiro sector como potenciais beneficiários indiretos deste estudo.

Abordagem científica / Metodologia

Em termos metodológicos optamos por uma estratégia de investigação para a ação ancorada numa análise intensiva, de carácter qualitativo, desenhada em quatro eixos principais:



RESEARCH BRIEF

Inteligência artificial e direitos humanos

Ana Cláudia Albergaria

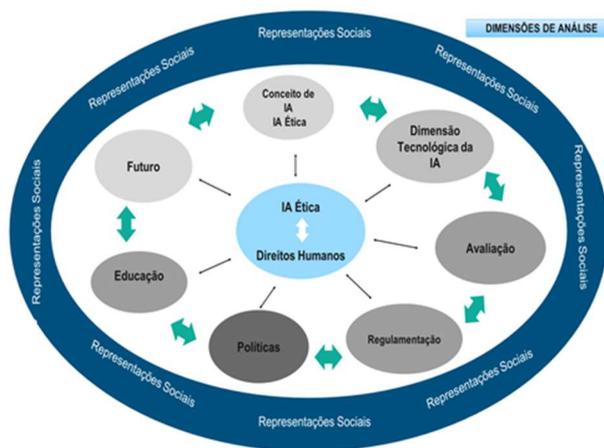


Eixo I - Fase exploratória. Iniciou-se com a análise bibliométrica e revisão da literatura sobre a temática em estudo, seguindo-se um trabalho de análise de conteúdos a documentos *on-line* sobre *IA responsável*. O tipo de documentos analisados num momento inicial recaiu sobre documentos regulamentares e políticos, tais como: pareceres, relatórios e Estratégias Nacionais, de autoria de: República Portuguesa/XXII Governo, da Comissão europeia, do Comité Europeu das Regiões do Parlamento Europeu e ainda da UNESCO, ONU, Fórum Económico Mundial e UNICEF.

Esta fase incluiu, ainda, a realização de uma mesa redonda sobre *IA responsável*, para aprofundamento de conhecimentos, definição de estratégias de superação de desafios e também para a promoção do conhecimento junto da comunidade. Este debate contou com a participação de peritos em IA de diferentes Universidades através do qual pudemos alargar a discussão a outros *stakeholders*.



Eixo II – Fase analítica. Corresponde fundamentalmente à realização de 29 entrevistas semiestruturadas a atores-chave de oito áreas científicas e de atuação profissional diferenciadas, com as quais pretendemos identificar as representações sociais e as práticas desses atores no que diz respeito a sete dimensões de análise principais: conceito de IA ética, ou responsável; dimensão tecnológica da IA; avaliação e monitorização dos sistemas de IA; regulamentação da IA; políticas para a IA responsável; educação para a literacia digital e, por último, a perspetiva destes atores sobre o futuro. Todas estas dimensões de análise convergem para a problematização dos impactos da IA nos direitos humanos.





RESEARCH BRIEF

Inteligência artificial e direitos humanos

Ana Cláudia Albergaria



A seleção dos entrevistados foi construída recorrendo à técnica de bola de neve, o que nos permitiu um grupo de entrevistados de áreas científicas diversificadas, com conhecimentos e responsabilidades no âmbito dos direitos humanos, mas também, e principalmente, com alto nível de experiência em IA, correspondendo maioritariamente a docentes universitários e investigadores em centros de investigação de universidades e fundações. Entrevistámos 19 homens e 10 mulheres, entre os 20 e 65 anos, maioritariamente de Portugal, mas também do Brasil, Alemanha e Espanha.

Tabela 1. Caracterização sociodemográfica dos entrevistados

| | ni | % |
|----------------------------------|----|-------|
| Sexo | | |
| Masculino | 19 | 65,5 |
| Feminino | 10 | 34,5 |
| Total | 29 | 100,0 |
| Área Científica | | |
| Perito em IA | 11 | 38 |
| Direito | 7 | 24,2 |
| Filosofia | 2 | 6,9 |
| Comunicação social | 2 | 6,9 |
| Políticas sociais | 1 | 3,4 |
| Empresas de tecnologia | 2 | 6,9 |
| Humanidades digitais | 1 | 3,4 |
| Ciências sociais | 2 | 6,9 |
| Instituição do terceiro setor | 1 | 3,4 |
| Total | 29 | 100,0 |
| Idade | | |
| 25 a 35 anos | 12 | 41,4 |
| 36 a 45 anos | 4 | 13,8 |
| 46 a 55 anos | 4 | 13,8 |
| Mais de 55 anos | 9 | 31 |
| Total | 29 | 100,0 |
| País de residência | | |
| Portugal | 24 | 82,8 |
| Brasil | 3 | 10,4 |
| Alemanha | 1 | 3,4 |
| Espanha | 1 | 3,4 |
| Total | 29 | 100,0 |
| Anos de experiência em IA | | |
| 30 anos | 4 | 13,8 |
| 20 anos | 5 | 17,2 |
| 15 anos | 2 | 6,9 |
| 10 anos | 13 | 44,8 |
| Menos de 8 anos | 2 | 6,9 |
| Não se aplica | 3 | 10,4 |
| Total | 29 | 100,0 |



RESEARCH BRIEF

Inteligência artificial e direitos humanos

Ana Cláudia Albergaria



Eixo III – Mobilização para a Participação. Corresponde à vertente de investigação-ação deste estudo, em que pretendemos promover a participação da comunidade em geral, mas sobretudo dos *stakeholders* principais, através do desenvolvimento de Grupos Ativos de Reflexão (GAR). Em maio de 2023 realizou-se o primeiro GAR, no Porto, no qual participou um grupo heterogéneo de 25 atores-chave, com o intuito de devolver os resultados parciais do estudo e de promover a reflexão sobre eles, bem como recolher contributos para a construção de um documento com linhas de ação para a promoção de *IA responsável*. Ao longo do projeto pretendemos desenvolver outros GAR, diferenciados de acordo com as fases e resultados *ongoing* obtidos.

Eixo IV – Divulgação. Através do GAR e da participação em congressos, seminários, simpósios, programas televisivos, debates junto das comunidades escolares; participação em grupos de trabalho e de investigação em parceria com outras universidades e institutos superiores. Organizaremos, também, um seminário, em 2024, para apresentação dos resultados finais do projeto, bem como redigir vários documentos de cariz diversos para a divulgação dos resultados.

Recomendações / Implicações para a prática

Um documento com linhas orientadoras para a promoção de IA responsável e com estratégias de prevenção e superação dos riscos será disseminado em 2024.

Referências

Von Schomberg, Rene (2013). "A vision of responsible innovation". In: R. Owen, M. Heintz and J Bessant (eds.) Responsible Innovation. London: John Wiley, forthcoming.



RESEARCH BRIEF

Inteligência artificial e direitos humanos
Ana Cláudia Albergaria



Entidade(s) financiadora(s)



Fundação
para a Ciência
e a Tecnologia

(ao abrigo do protocolo de colaboração FCT e IS-UP, Ref. UI/BD/150781/2020)

Entidades promotora / Instituição de acolhimento



Contactos

Investigador(a) responsável:

Ana Cláudia Albergaria
Bolseira de Doutoramento FCT/IS-UP (UI/BD/150781/2020)
E-mail: homemdealbergaria@hotmail.com

Equipa de orientação científica:

Critina Parente
Orientador
E-mail: cparente@letras.up.pt

Maria Goreti Carvalho Marreiros
Co-orientadora
E-mail: cparente@letras.up.pt

António Vieira de Castro
Co-orientador
E-mail: avc@isep.ipp.pt

Instituto de Sociologia

Instituto de Sociologia
Faculdade de Letras da Universidade do Porto
Via Panorâmica, s/n
4150-564 Porto
PORTUGAL
Telefone: 00 351 226 077 132 (ext. 3364)
E-mail: isociologia@letras.up.pt
URL: <http://isociologia.up.pt>