

# Gui*i*A

Guia para a Inteligência Artificial

GUIA PARA UMA INTELIGÊNCIA ARTIFICIAL  
ÉTICA, TRANSPARENTE E RESPONSÁVEL  
NA ADMINISTRAÇÃO PÚBLICA

## SUMÁRIO EXECUTIVO



# SUMÁRIO EXECUTIVO

A crescente escolha de soluções de Inteligência Artificial (IA), com elevado impacto sobre inúmeros setores da sociedade, suscita um conjunto de desafios aos quais urge dar resposta. Neste âmbito, enfatizam-se as questões associadas à ética, justiça, transparência, responsabilidade e explicabilidade destes sistemas.

Embora a evolução dos sistemas de IA se faça acompanhar de um significativo aumento de investigação académica e científica, a rapidez com que estas tecnologias proliferam e o modo como se relacionam com a sociedade, exige um espaço para dúvidas e discussão.

Por outro lado, é necessário identificar o que pode estar em causa, no uso de IA, por exemplo se estamos a avaliar candidatos para um emprego ou para o acesso a uma instituição de ensino, se estamos a decidir sobre atribuição de crédito ou um apoio social, se estamos perante diagnósticos médicos ou decisões judiciais, ou veículos autónomos, todos estes casos constituem alguns exemplos onde a questão do impacto ganha uma outra dimensão de discussão e importância.

É importante lembrar que a IA é concebida por pessoas e deve centrar-se na criação de benefícios para as pessoas, e que as questões éticas associadas aos sistemas de IA estão intrinsecamente relacionadas com todos aqueles que estão envolvidos na sua conceção e utilização, desde aqueles que desenvolvem os algoritmos, aos decisores, e aos governos. Em igual medida, é pertinente considerar o que está sob o controlo humano e o grau de autonomia destas soluções.

A Agência para a Modernização Administrativa (AMA) vem uma vez mais assumir o seu papel enquanto instituição pública responsável pela promoção e desenvolvimento da modernização administrativa em Portugal propondo, através do projeto IA Responsável, a exploração de caminhos para se conseguir uma IA que considere estas questões. Esta disponibiliza documentos e instrumentos com o intuito de dar suporte e de promover a adoção progressiva e gradual de uma IA ética, transparente e responsável.

Faz igualmente parte da visão deste projeto, contribuir para o equilíbrio entre uma discussão filosófica/teórica e a discussão prática, aprofundando os conceitos de responsabilização, transparência, explicabilidade, justiça e ética. Conceitos estes que contêm, a título de exemplo, a problemática do viés, muito associada aos algoritmos com impacto social. A partir de uma reflexão que considere ambas as discussões, pretende-se garantir a proteção da democracia, do Estado de direito e dos direitos fundamentais, com a materialização destes conceitos no modo como os serviços de IA são pensados, desenhados e providenciados, quer no setor público quer no setor privado.

Na sua estrutura, o guia procura definir o que é a IA e o seu funcionamento, mostrar como está presente na sociedade e identificar alguns dos efeitos decorrentes da sua utilização. Procura ainda informar sobre o enquadramento da IA no Mundo e em Portugal, e identifica

o ecossistema e os principais atores no contexto nacional. Por outro lado, e considerando a importância dos dados para o desenvolvimento e alimentação destes sistemas, faz também referência ao ecossistema de dados na Administração Pública (AP) e aos princípios que lhe devem estar subjacentes.

Este guia pretende trazer para a discussão pública a necessidade de estabelecer os pilares de regulação, supervisão, liderança e governação, para a elaboração de um código de ética, fomentar regulamentação e leis que forneçam orientações e suporte aos desenvolvimentos tecnológicos.

O guia enuncia também um conjunto de valores e princípios em linha com a lista de direitos humanos, e explora o tema da inclusão, da igualdade, do desenvolvimento sustentável e do bem-estar.

Em contraponto, são abordados os efeitos perniciosos associados a sistemas de IA, com alguns exemplos, e é reforçada a importância de se criarem mecanismos rigorosos de monitorização, auditoria, proteção e segurança.

Este trabalho fornece também recomendações do ponto de vista mais amplo e genérico, e identifica um conjunto de barreiras e desafios que devem ser considerados aquando da construção e implementação de sistemas de IA Responsáveis.

Na dimensão prática deste projeto foi desenvolvida uma ferramenta de avaliação do risco, construída em linha com as orientações do guia. Este instrumento possibilita a análise da suscetibilidade de sistemas de IA, associada às cinco dimensões subjacentes a uma IA Responsável, referidas no guia, e dá recomendações de ações e sugestões de leituras, em função do nível de maturidade dos atores.

Em última instância, o guia e a ferramenta permitem estruturar o processo de construção e de implementação de sistemas inteligentes para que sejam mais responsáveis, éticos e transparentes, por via da compreensão e adoção de conceitos, e da mudança comportamental. Desse modo, ambos constituem um recurso importante na antecipação e mitigação de riscos em sistemas de IA nas cinco dimensões para uma IA Responsável, proporcionando-se soluções mais éticas tanto na Administração Pública como no Setor Privado.

O guia é sintetizado em três documentos de leitura simples e rápida, focados: nos valores, princípios e recomendações; nas dimensões de avaliação; e na ferramenta de avaliação do risco. É nossa ambição que as versões referentes aos conteúdos teóricos e à ferramenta, não sejam estáticas e possam sofrer atualizações ao longo do tempo.

Nas referências bibliográficas que fundamentaram o desenvolvimento deste projeto, incluíram-se as de autoria de Organizações Intergovernamentais, da Comissão Europeia, do Setor Privado, de Developers e de Consultoras.

