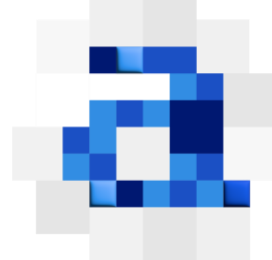


# GuIA

Guia para a  
Inteligência Artificial

SUMÁRIO EXECUTIVO



### SUMÁRIO EXECUTIVO

A crescente adoção de soluções baseadas em Inteligência Artificial (IA), com impacto sobre diversos setores da sociedade, suscita um conjunto de desafios que exigem respostas urgentes e responsáveis, em 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 nesta área, a rapidez com que estas tecnologias evoluem e se integram na sociedade, exige um espaço para dúvidas e discussão.

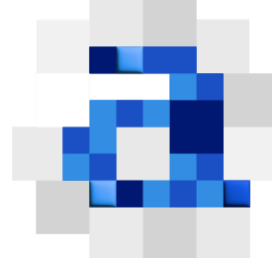
Por outro lado, é necessário identificar e compreender o que pode estar em causa, no uso de IA, em contextos sensíveis como por exemplo se estamos a avaliar candidatos para um emprego ou para o acesso a uma instituição de ensino, se estamos a decidir sobre atribuição de crédito ou um apoio social, se 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. É igualmente relevante refletir sobre o grau de autonomia destas soluções e o que permanece sob controlo humano.

Neste contexto, 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. Para isso, disponibiliza documentos e instrumentos que apoiam e promovem a adoção progressiva e gradual de uma IA ética, transparente e responsável.

O projeto visa ainda equilibrar a reflexão teórica com a aplicação prática, aprofundando conceitos como 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 tem como objetivo promover o debate público sobre a importância de estabelecer pilares de regulação, supervisão, liderança e governação. Pretende ainda apoiar a elaboração de um código de ética e incentivar e fomentar a regulamentação e leis que orientem e sustentem os 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 é disponibilizada uma ferramenta de avaliação do risco que possibilita a análise da suscetibilidade de sistemas de IA. Esta ferramenta permite analisar a suscetibilidade dos sistemas de IA em cinco dimensões fundamentais, oferecendo recomendações e sugestões de leitura adaptadas ao nível de maturidade dos utilizadores.

Em última instância, tanto o guia como a ferramenta visam estruturar o processo de desenvolvimento e implementação de sistemas inteligentes mais éticos, responsáveis e transparentes, promovendo a adoção de boas práticas e a 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 conteúdo do guia está organizado em três documentos de leitura simples e acessível: um focado nos valores, princípios e recomendações; outro nas dimensões de avaliação; e um terceiro dedicado à ferramenta de avaliação de risco. É nossa ambição que estes conteúdos sejam dinâmicos e evoluam ao longo do tempo, acompanhando os avanços tecnológicos e sociais.

As referências bibliográficas que fundamentaram o desenvolvimento deste projeto, incluem as de autoria de Organizações Intergovernamentais, da Comissão Europeia (CE), do Setor Privado, de Developers e de Consultoras especializadas.

