

# Guia

Guia para a Inteligência Artificial

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

## FERRAMENTA DE AVALIAÇÃO DO RISCO ÉTICO



Este documento resume conteúdos que integram o Guia para a Inteligência Artificial Ética, Transparente e Responsável na Administração Pública, elaborado pela AMA e visa constituir uma leitura mais simples e focada na

### **Ferramenta de Avaliação do Risco Ético 3/3**

A Ferramenta de Avaliação de Risco traduz os **valores e princípios de IA Responsável**, detalhados ao longo do Guia. A utilização desta Ferramenta é indispensável à antecipação e mitigação de riscos em sistemas com IA de forma global e nas cinco dimensões: **Responsabilização, Transparência, Explicabilidade, Justiça e Ética**.

A Ferramenta de Avaliação de Risco (Ferramenta) tem o intuito de:

- **Analisar a suscetibilidade** de projetos de IA, de sistemas inteligentes ou de algoritmos relativamente às cinco dimensões subjacentes a uma IA Responsável;
- **Comparar os resultados** obtidos com as avaliações nacional e setorial de referência;
- **Recomendar ações** em função do nível de maturidade de IA auferido.

As dimensões consideradas transpõem os cinco princípios de IA Responsável adotados:

- **Responsabilização** (responsabilidade e possibilidade de auditoria/inspeção);
- **Transparência** (acesso às componentes e procedimentos);
- **Explicabilidade** (explicação do funcionamento);
- **Justiça** (proteção e garantias para os utilizadores e beneficiários);
- **Ética** (mecanismos efetivos de mitigação de vieses inesperados).

A Ferramenta **destina-se a todas as pessoas e entidades** que pretendam avaliar riscos em projetos de IA, sistemas inteligentes ou algoritmos que se encontrem numa das seguintes etapas:

- Conceção;
- Planeamento;
- Desenvolvimento inicial;
- Desenvolvimento avançado;
- Testes;
- Protótipo;
- Validação; e
- Produção.

Procura-se, deste modo, garantir a **avaliação ao longo do ciclo do projeto**, quer na fase anterior à implementação (*by design*), quer na fase posterior (*by evolution*).

A Ferramenta pode ser preenchida por qualquer pessoa, mesmo que não esteja associada a uma entidade ou a uma equipa específica do projeto. Esta pode inclusive ser utilizada por diferentes pessoas da mesma entidade. Como destinatários incluem-se:

- Pessoas externas à entidade;
- Utilizadores;
- Programadores;
- Analistas/ Engenheiros;
- Consultores de IA.

## BENEFÍCIOS

A Ferramenta auxilia utilizadores e desenvolvedores na construção de sistemas inteligentes mais responsáveis por via da **compreensão/assimilação de conceitos e da mudança comportamental**.

A Ferramenta contribui significativamente para a:

- Eliminação do efeito de *black box*;
- Compreensão de como a aprendizagem de máquina pode ser incorporada nas entidades;
- Redução de vieses;
- Proteção de pessoas vulneráveis;
- Não discriminação;
- Interdisciplinaridade;
- Identificação de riscos e impactos nos utilizadores/ beneficiários;
- Monitorização de resultados;
- Melhoria contínua do desempenho dos sistemas, através da aprendizagem;
- Melhoria das políticas e dos mecanismos de mitigação de riscos;
- Eficácia dos processos de inspeção/ auditoria dos sistemas;
- Segurança, qualidade e proteção dos sistemas;
- Compreensão dos resultados no contexto nacional e setorial;
- Sustentabilidade ambiental, social e económica.

## ARQUITETURA

A Ferramenta está estruturada em:

- Conjunto de **perguntas** do tipo binário, likert ou escolha múltipla, validado anualmente;
- **Ponderações** atribuídas pela AMA e pelo utilizador a cada uma das cinco dimensões, validadas anualmente;
- **Pontuação** de avaliação;
- E matriz de **recomendações** associada ao nível de maturidade em que se encontra a entidade.

## UTILIZAÇÃO

A utilização da Ferramenta tem início com a **autenticação do utilizador** através de Chave Móvel Digital ou do Cartão de Cidadão.

Após a autenticação, procede-se ao **registo** do utilizador, da entidade e do projeto, solicitando-se as informações explícitas na figura 1.

Todas as perguntas são de **resposta obrigatória**.

Após preenchimento das respostas, o utilizador solicita o Relatório de Avaliação, o qual pode ser impresso ou arquivado em formato PDF.

No final da avaliação e, após consentimento prévio, o utilizador pode submeter os seus contactos pessoais para iniciativas desenvolvidas no âmbito da IA Responsável.

UTILIZADOR	ENTIDADE	PROJETO
<ul style="list-style-type: none"> <li>• Habilitações académicas</li> <li>• Área profissional/ académica</li> <li>• Função</li> </ul>	<ul style="list-style-type: none"> <li>• Setor</li> <li>• Validação do registo associado à entidade</li> <li>• <i>Upload</i> do documento da entidade para efeitos de credenciação do utilizador</li> </ul>	<ul style="list-style-type: none"> <li>• Designação</li> <li>• Setor de aplicação</li> <li>• Contexto em que se desenvolve o projeto de IA</li> <li>• Objetivos do projeto</li> <li>• Resultados esperados</li> <li>• Aplicação</li> <li>• Técnica</li> <li>• Etapa do projeto</li> <li>• Grupo-alvo do projeto</li> <li>• Impacto esperado</li> <li>• <i>Framework</i>, Ambiente de Desenvolvimento Integrado, Simulador, Linguagem e Biblioteca utilizados no projeto</li> <li>• Tecnologias emergentes aplicadas</li> </ul>

FIGURA 1: ELEMENTOS DE REGISTO DO UTILIZADOR, ENTIDADE E PROJETO.

### NÍVEL DE MATURIDADE

O nível de maturidade da entidade está associado ao projeto consignado.

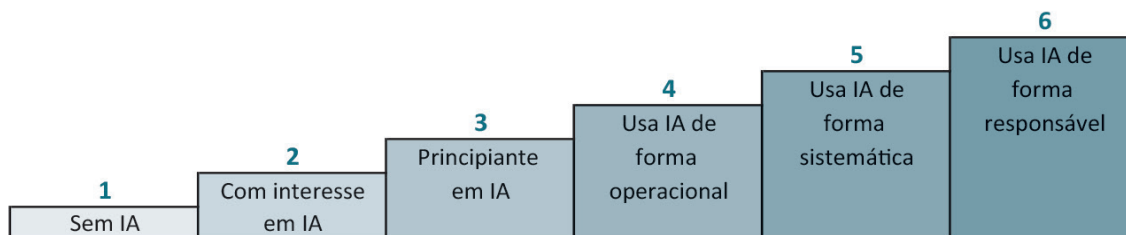


FIGURA 2: NÍVEIS DE MATURIDADE/ ESTÁGIOS DE MATURIDADE.

Este está relacionado com diferentes **estágios de maturidade**, ilustrados na figura 2:

1. **Uma entidade sem IA**, que não possui qualquer tipo de sistema baseado em técnicas adaptativas;
2. **Uma entidade com interesse em IA**, que tem pesquisado sobre o tema e discute internamente alguns assuntos, mas não concretizou nenhuma iniciativa com vista ao desenvolvimento de sistemas baseados em técnicas adaptativas;

3. **Uma entidade principiante em IA**, que tem ideias e já iniciou, de forma embrionária, alguns projetos de IA. Os membros da equipa são normalmente pluridisciplinares e estão orientados para a experimentação e visualização de resultados, na tentativa de perceber se o algoritmo tem interesse e cumpre a função pretendida;
4. **Uma entidade que usa IA de forma operacional**, que tem normalmente processos de ML implementados e equipas mais diversificadas em termos de conhecimento. Os algoritmos desenvolvidos são aplicados a casos de utilização concretos e apoiam de forma efetiva os processos de decisão;
5. **Uma entidade que usa IA de forma sistemática**, que aplica os pressupostos anteriores, suportando-os por princípios de IA, frameworks e standards internacionalmente reconhecidos;
6. **Uma entidade que usa IA de forma responsável**, que resulta do desenvolvimento do estágio antecedente, que segue processos e procedimentos que contribuem, no mínimo, para a transparência e para a mitigação dos riscos éticos dos sistemas.

## RECOMENDAÇÕES

Cada pergunta é alvo de uma recomendação gerada em função de um limiar de respostas definido.

A matriz de recomendações considera as seguintes possibilidades:

	ABAIXO DO LIMIAR	ACIMA DO LIMIAR
Sem IA	Sugerimos ler sobre o tema	Sugerimos ler sobre o tema
Interesse em IA	Sugerimos ler sobre o tema	Excelente para a fase em que se encontra
Principiante em IA	Recomendamos ler e planear	Ótimo para a fase em que se encontra
Usa IA de forma operacional	Recomendamos planear e atuar	Bom para a fase em que se encontra
Usa IA de forma sistemática	É essencial atuar	Razoável para a fase em que se encontra
Usa IA de forma responsável	Já deveria ter atuado sobre este aspeto	Siga assim, está no bom caminho

FIGURA 3: MATRIZ DE RECOMENDAÇÕES.

Todas as recomendações são complementadas com sugestões de leituras acessíveis através da internet.

## RELATÓRIO DE AVALIAÇÃO

O relatório de avaliação apresenta, por projeto, os seguintes resultados:

- Pontuação relativa ao projeto;
- Pontuação nacional;
- Pontuação setorial;
- Pontuação por dimensão de avaliação;
- Respostas dadas a cada pergunta;
- Recomendações por pergunta.

Pretende-se que as entidades e os projetos a partir da Responsabilização, da Transparência e da Explicabilidade evoluam para sistemas cada vez mais justos e éticos.

