MINISTÉRIO PÚBLICO
PORTUGAL

EM DEFESA DA
LEGALIDADE
DEMOCRÁTICA

SINOPSE

Carta para a utilização ética da IA pelo Ministério Público e Política de Auditoria e Monitorização Técnica

MINISTÉRIO PÚBLICO
PORTUGAL
EM DEFESA DA LEGALIDADE DEMOCRÁTICA

Índice

1. Carta para a utilização ética da Inteligência Artificial pelo Ministério Público

1.1. Preâmbulo

O desenvolvimento da IA , nomeadamente na sua vertente generativa, e a disseminação de ferramentas baseadas em IA enquanto aplicações acessíveis à comunidade, têm já um impacto transformador na sociedade em geral, que terá naturais reflexos no âmbito da Justiça.

A IA é o campo da ciência da computação que se dedica a criar sistemas capazes de realizar tarefas que, normalmente, requerem inteligência humana, como reconhecimento de padrões, tomada de decisões e resolução de problemas, envolvendo um conjunto amplo de técnicas e algoritmos para construir sistemas que podem simular um comportamento inteligente em diversos contextos, seja ao nível de *Machine Learning* (ML), subárea da IA que se foca em ensinar máquinas a aprender a partir de dados, em vez de programá-las diretamente com regras rígidas, seja em sistemas de *Deep Learning* (DL), técnica que utiliza redes neurais artificiais profundas, compostas por múltiplas camadas de processamento e particularmente eficaz para lidar com grandes volumes de dados complexos, como imagens, áudio e texto, permitindo reconhecimento de padrões de forma mais hierárquica e abstrata.

Pese embora a extraordinária utilidade destas vertentes, a disseminação desta tecnologia ao público em geral ocorre essencialmente pelas aplicações de IA Generativa (GenIA) ou seja, aos sistemas baseados em IA que podem criar conteúdo novo, seja ele textual ou multimédia, em vez de apenas analisar ou classificar dados existentes, e que permitem simular a criação humana em diferentes domínios, particularmente na linguagem.

Em qualquer uma destas vertentes, existe um amplo campo de aplicação da IA ao sistema formal de justiça e, em particular, à atividade do Ministério Público.

A utilização de sistemas de IA pelo Ministério Público pode assim contribuir para a melhoria da eficiência, da qualidade dos serviços e da capacidade de resposta aos desafios desta magistratura.

No entanto, a utilização de sistemas de IA pelo Ministério Público deve ser sempre efetuada em conformidade com o Estado de Direito, os princípios democráticos e com respeito pelos direitos, liberdades e garantias fundamentais, sendo essencial que, para uma utilização legítima de ferramentas de IA, esta se baseie em princípios éticos e jurídicos claros.

Não se olvida que já hoje, através de aplicações comerciais de IA, principalmente de IA generativa, se faça uso de IA na administração da Justiça, nomeadamente na vertente de compilação de referentes doutrinários e jurisprudenciais e no auxílio à elaboração de peças processuais, o que torna assim premente a definição de critérios éticos e de transparência para tal utilização.

Com efeito, são vários os riscos que, particularmente no caso da IA generativa, decorrem da sua utilização, como sejam:

a) Risco de discriminação ou amplificação da discriminação, particularmente devido à utilização de conjuntos de dados enviesados;

b) Risco de violação de direitos fundamentais ou desequilíbrio na conciliação de direitos fundamentais conflituantes;

c) Risco de utilização ou divulgação indevida de dados pessoais, dados sensíveis, dados sujeitos a segredos de justiça, segredo profissional ou segredo de Estado;

d) Risco de redução da responsabilidade e da autonomia dos magistrados devido à utilização de modelos de IA não explicáveis;

e) Risco de decisões fundamentadas em dados factualmente inexatos (respostas falsas, «alucinações» e enviesamentos);

f) Risco de criação e utilização de disposições legais e fundamentos jurisprudenciais inexistentes ou descontextualizados através de IA generativa;

g) Risco de omissão de referências para os dados e informações fornecidas e potencial violação da propriedade intelectual e dos direitos de autor;

h) Risco de as ferramentas de IA não refletirem com precisão o raciocínio jurídico;

i) Risco de cristalização da jurisprudência;

j) Risco de criação de perfis de magistrados, propiciando situações de *forum shopping*;

k) Risco de viés de automação – confiança no resultado de IA (particularmente sério na IA generativa generalista).

Tendo presentes estes riscos, pretende-se estabelecer as orientações éticas e os princípios de conduta para a utilização, mas igualmente para a conceção, desenvolvimento e avaliação de sistemas de IA pelo Ministério Público, integrando os princípios contidos nas principais referências internacionais, europeias e nacionais nesta matéria, enquanto enquadramento jurídico, mas igualmente numa vertente ética, assumindo-se como um verdadeiro compromisso de cidadania e responsabilidade por todos os magistrados do Ministério Público.

Para a implementação e utilização de ferramentas de IA no Ministério Público, importa ter presente o enquadramento legal aplicável, particularmente incisivo no espaço da União Europeia (UE), designadamente os seguintes referentes essenciais:

- Diretiva (UE) 2016/680 do Parlamento Europeu e do Conselho, de 27 de abril de 2016 (proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais pelas autoridades competentes para efeitos de prevenção, investigação, deteção ou repressão de infrações penais ou execução de sanções penais);

- Lei n.º 59/2019, de 08 de agosto (Lei de dados pessoais para prevenção, deteção, investigação ou repressão de infrações penais);
- Regulamento (UE) 2018/1725 do Parlamento Europeu e do Conselho, de 23 de outubro de 2018 (proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais pelas instituições e pelos órgãos e organismos da União e à livre circulação desses dados);
- Regulamento (UE) 2023/2854 do Parlamento Europeu e do Conselho, de 13 de dezembro de 2023 (regras harmonizadas sobre o acesso equitativo aos dados e a sua utilização)
- Regulamento (UE) 2022/2065 do Parlamento Europeu e do Conselho de 19 de outubro de 2022 (Regulamento de Serviços Digitais – DSA);
- Convenção-Quadro do Conselho da Europa sobre a Inteligência Artificial e os Direitos Humanos, a Democracia e o Estado de Direito, adotada pelo Conselho da Europa em 17 de maio de 2024;
- Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho de 13 de junho de 2024 (Regulamento da Inteligência Artificial – AIR).

Para além destes, importa considerar igualmente os seguintes referentes:

- Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente, adotada pela CEPEJ na sua 31.ª reunião plenária (Estrasburgo, 3 e 4 de dezembro de 2018);
- Livro Branco sobre Inteligência Artificial – Uma abordagem europeia de excelência e confiança (2020);
- Carta Portuguesa de Direitos Humanos na Era Digital, aprovada pela Lei n.º 27/2021, de 17 de maio;

- Recomendação sobre a Ética da Inteligência Artificial, aprovada pela UNESCO em 23 de novembro de 2021;
- Declaração Europeia sobre Direitos e Princípios Digitais (2022);
- Princípios da OCDE sobre Inteligência Artificial, recomendação adotada em 22 de maio de 2019 e atualizada em 03 de maio de 2024;
- UNESCO, *Guidelines for the Use of AI Systems in Courts and Tribunals, 2025*;
- *Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act)*, COM (2025) 5052, 29/07/2025;
- Diretrizes sobre a utilização de inteligência artificial generativa nos tribunais, adotadas pela CEPEJ na sua 45.ª reunião plenária (Estrasburgo, 4 e 5 de dezembro de 2025)

Para a utilização de ferramentas de IA, mas principalmente para o desenvolvimento e implementação de ferramentas de IA institucionais, deverá ter-se em conta, no âmbito da regulamentação aplicável, as específicas tipologias previstas no Regulamento (UE) 2024/1689 para os sistemas de IA:

- Risco inaceitável – será proibido tudo o que for considerado uma clara ameaça para os cidadãos europeus (prática de IA proibida – artigo 5.º)
    - Sistema de IA para realizar avaliações de risco de pessoas singulares, a fim de avaliar ou prever o risco de uma pessoa singular cometer uma infração penal, com base exclusivamente no perfil de uma pessoa singular ou na avaliação dos seus traços e características de personalidade;
    - Sistemas de IA de identificação biométrica em tempo real em espaços publicamente acessíveis (...) salvo se tal utilização for estritamente necessária para atingir um dos seguintes objetivos:

- Busca direcionada a vítimas específicas de rapto, tráfico de pessoas ou exploração sexual, bem como para busca de pessoas desaparecidas;
- Prevenção de uma iminente, substancial e específica ameaça à vida ou integridade física de pessoas naturais ou de uma ameaça genuína e previsível de ataque terrorista;
- Localização ou identificação de pessoa suspeita de ter cometido um crime, para o propósito de investigação e repressão criminal ou de execução de uma pena pelos crimes constantes no anexo II punível no respetivo Estado-membro com uma pena com a moldura máxima de pelo menos 4 anos de prisão.

➢ Risco elevado – a conformidade do sistema de IA de alto risco será avaliada antes destes serem introduzidos no mercado, bem como durante todo o seu ciclo de vida (práticas de IA de alto risco – artigos 6.º e anexo III)

- Sistemas de IA destinados a serem utilizados por autoridades responsáveis pela aplicação da lei ou em seu nome, ou por instituições, agências, organismos ou organismos da União em apoio às autoridades responsáveis pela aplicação da lei ou em seu nome, para avaliar o risco de uma pessoa singular se tornar vítima de infrações penais;
- Sistemas de IA destinados a serem utilizados por ou em nome das autoridades responsáveis pela aplicação da lei ou pelas instituições, órgãos e agências da União em apoio às autoridades responsáveis pela aplicação da lei, como polígrafos e ferramentas semelhantes;
- Sistemas de IA destinados a serem utilizados por ou em nome das autoridades responsáveis pela aplicação da lei, ou pelas instituições, agências, organismos ou organismos da União em apoio às autoridades

responsáveis pela aplicação da lei, para avaliar a fiabilidade das provas no decurso da investigação ou da repressão de infrações penais;

- Sistemas de IA destinados a serem utilizados por uma autoridade judicial ou em seu nome para ajudar uma autoridade judicial na investigação e interpretação de factos e da lei e na aplicação da lei a um conjunto concreto de factos ou utilizados de forma semelhante em litígios alternativos resolução;
- Sistemas de IA destinados a serem utilizados pelas autoridades responsáveis pela aplicação da lei ou em seu nome ou pelas instituições, agências, organismos ou organismos da União em apoio às autoridades responsáveis pela aplicação da lei para avaliar o risco de uma pessoa singular cometer ou reincidir na prática de crimes, não apenas com base na definição de perfis de pessoas singulares, tal como referido no artigo 3.º, n.º 4, da Diretiva (UE) 2016/680, ou para avaliar traços e características de personalidade ou comportamentos criminosos anteriores de pessoas singulares ou grupos.

Para os sistemas de IA de alto risco, o regime preconizado no Regulamento, designadamente nos artigos 8.º a 27.º, consagra uma série de requisitos apertados que necessariamente têm de ser cumpridos desde a sua conceção até ao final da sua utilização, de que se destacam os seguintes:

- Deve ser criado, implantado, documentado e mantido um sistema de gestão de riscos em relação aos sistemas de IA de risco elevado;
- Os conjuntos de dados para treino, validação e teste de sistemas IA de alto risco estão sujeitos a práticas de governação e gestão de dados, incluindo medidas apropriadas para deteção, prevenção e mitigação do risco de viés que possam afetar a saúde e a segurança das pessoas, ter um impacto negativo nos direitos fundamentais ou conduzir a uma discriminação proibida pelo direito da União;

- Os sistemas de IA de risco elevado devem ser concebidos e desenvolvidos de maneira que alcancem um nível apropriado de exatidão, solidez e cibersegurança e apresentem um desempenho coerente em relação a tais aspetos durante o seu ciclo de vida;
- Os sistemas de IA de alto risco devem ser concebidos e desenvolvidos de tal forma que possam ser efetivamente supervisionados por pessoas singulares durante o período em que o sistema de IA estiver em utilização (pelo menos por duas pessoas);
- Os implementadores de sistemas de IA de alto risco que tomam decisões ou auxiliam na tomada de decisões relacionadas com pessoas singulares devem informar as pessoas singulares de que estão sujeitas à utilização do sistema de IA de alto risco (aplicação do artigo 13.º da Diretiva 2016/680).

Em conclusão, a utilização, desenvolvimento e implementação de ferramentas de IA deverá apenas ocorrer após a definição dos seguintes parâmetros:

- Estabelecimento de uma carta ética, contendo, designadamente, uma política de uso responsável de ferramentas de IA aos seus utilizadores;
- Estabelecimento de padrões de explicabilidade, definindo orientações claras sobre o que constitui níveis aceitáveis de transparência e interpretabilidade em aplicações de alto risco, tendo em conta a natureza das atribuições do Ministério Público;
- Monitorização e controlo a sistemas de IA, com estruturas para monitorização e auditoria de sistemas que evoluem dinamicamente;
- Implementação de orientações que possibilitem o equilíbrio entre a privacidade e a transparência, fornecendo orientações claras sobre como implementar aplicações de IA e implementar mecanismos de supervisão (como ficheiros de registo) de forma que não entrem em conflito com as normas de privacidade.

1.2. Âmbito e Finalidade

1. Esta Carta aplica-se a todos os sistemas de IA utilizados, direta ou indiretamente, no âmbito das funções administrativas e jurisdicionais do Ministério Público, incluindo ferramentas de apoio à decisão, análise de dados jurídicos, automatização de tarefas processuais e outros sistemas de IA, particularmente aqueles considerados, na formulação do Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho de 13 de junho de 2024, como sendo de risco elevado.

2. A finalidade desta Carta é garantir que a utilização de ferramentas e aplicações de IA pelos magistrados do Ministério Público é efetuada com respeito pelos direitos fundamentais, pela dignidade da pessoa humana, pela proteção de dados pessoais e pelos princípios constitucionais de justiça, imparcialidade, transparência, responsabilidade e segurança jurídica.

1.3. Princípios Éticos Fundamentais

1. Princípio do respeito pelos Direitos Fundamentais, visando garantir que a conceção e implementação de instrumentos e serviços de IA respeita e é compatível com os Direitos Fundamentais consagrados designadamente na Constituição da República Portuguesa e na Convenção Europeia dos Direitos Humanos, incluindo o direito a um processo justo, à privacidade e à proteção de dados pessoais.

2. Princípio da não discriminação e igualdade de tratamento, em que a utilização ou desenvolvimento de sistemas de IA não pode produzir, amplificar ou reforçar qualquer forma de discriminação direta ou indireta com base em características pessoais ou sociais, devendo ser adotadas medidas para identificar e corrigir viés algorítmico, visando prevenir especificamente o desenvolvimento ou a intensificação de qualquer discriminação entre indivíduos ou grupos de indivíduos.

3. Princípio da transparência e da explicabilidade, no sentido em que as recomendações baseadas em IA devem ser transparentes e, tanto quanto possível, explicáveis em termos

compreensíveis para utilizadores e demais intervenientes, designadamente processuais, devendo os sistemas de IA ser auditados e dispor de mecanismos de contestação, garantindo a responsabilização humana pelas decisões finais.

4. Princípio da Proteção de Dados e da Privacidade, no sentido em que a utilização de IA deve obedecer rigorosamente às regras de proteção de dados, particularmente no que respeita à legalidade do tratamento, limitação da finalidade do tratamento, minimização de dados, conservação limitada de dados, integridade dos dados, confidencialidade e proteção dos direitos dos respetivos titulares.

5. Princípio da Supervisão Humana ("sob controlo do utilizador"), no sentido de que as ferramentas de IA devem ser utilizadas como suporte à decisão humana, e não como substitutas de juízos de valor ou de decisões do Ministério Público, devendo a autoridade humana manter sempre o controlo e poder de revisão das decisões suportadas por IA, assim impedindo uma abordagem meramente prescritiva e acrítica.

6. Princípio da qualidade e segurança, no sentido em que os sistemas de IA devem ser concebidos e testados de forma a assegurar a sua robustez técnica, fiabilidade, segurança e respeito pelos direitos fundamentais, incluindo avaliações de impacto prévias e monitorização contínua dos resultados e, no que concerne ao tratamento de decisões e dados judiciais, a utilização de fontes certificadas e dados intangíveis com modelos concebidos de forma multidisciplinar, em ambiente tecnológico seguro.

1.4. Utilização de Ferramentas de IA em Geral

A. Confidencialidade e dever funcional

1. Os magistrados do Ministério Público mantêm integralmente o dever de confidencialidade, de proteção de dados pessoais e de respeito pelo segredo de justiça, pelo segredo profissional e pelo segredo de Estado quando utilizam ferramentas de IA.

2. A utilização de sistemas de IA não exonera o utilizador da responsabilidade pessoal pelo conteúdo produzido ou revisto com apoio tecnológico.

3. É proibida a delegação, explícita ou implícita, de juízo jurídico ou apreciação probatória em sistemas automatizados sem intervenção, revisão e decisão do magistrado.

B. Proibição de utilização para avaliação preditiva individual

1. É proibida a utilização de ferramentas de IA para:

a) Prever a probabilidade de condenação;
b) Estimar risco de reincidência;
c) Sugerir medida de coação com base em perfilização automatizada.

2. Qualquer sistema de análise estatística de natureza preditiva para além do disposto no número anterior deve ter natureza meramente auxiliar e nunca vinculativa.

C. Regras de conduta fundamentais

1. Na utilização de ferramentas de IA, os magistrados do Ministério Público devem assumir que:

a) A ferramenta de IA é um instrumento auxiliar, nunca um decisor;
b) O magistrado mantém sempre controlo e responsabilidade sobre o conteúdo do seu trabalho;
c) A proteção dos dados pessoais, processuais e o segredo de justiça prevalecem sobre qualquer conveniência tecnológica.

2. Em caso de dúvida sobre os limites da utilização de ferramentas de IA, não deve a mesma ser utilizada.

1.5. Utilização de Ferramentas Não Institucionais de IA

A. Âmbito das ferramentas abrangidas

1. Consideram-se ferramentas não institucionais de inteligência artificial os sistemas de IA disponibilizados no mercado para utilização transversal, que não tenham sido especificamente certificadas para uso institucional, incluindo sistemas de IA generativa, assistentes virtuais, ferramentas de análise de texto, sumarização, tradução automática, pesquisa jurídica automatizada e apoio à redação.

2. Incluem-se igualmente sistemas disponibilizados em ambiente *cloud*, plataformas externas e quaisquer ferramentas cujo processamento ocorra fora da infraestrutura tecnológica controlada pelo Ministério Público.

B. Princípio da proibição de inserção de dados processuais identificáveis

1. É proibida a introdução, em ferramentas não institucionais de IA, de:

a) Dados pessoais identificados ou identificáveis relativos a intervenientes processuais, designadamente arguidos, vítimas, testemunhas ou terceiros;

b) Elementos constantes de processos sujeitos a segredo de justiça;

c) Peças processuais integrais ou excertos que permitam a identificação do processo;

d) Informação classificada, sensível ou protegida por dever de confidencialidade funcional.

2. Para efeitos do número anterior, considera-se dado identificável qualquer informação que, direta ou indiretamente, permita identificar uma pessoa singular, em especial por referência a um identificador, como por exemplo um nome, um número de identificação, dados de localização, identificadores por via eletrónica ou a um ou mais elementos específicos da identidade física, fisiológica, genética, mental, económica, cultural ou social dessa pessoa singular.

C. Tratamento de dados pessoais e anonimização

1. Sempre que, excecionalmente, se utilize uma ferramenta não institucional de IA para apoio técnico (por exemplo, estruturação de texto ou revisão linguística), os dados devem ser previamente:

a) Anonimizados de forma irreversível; ou

b) Pseudonimizados, desde que o processo decorra exclusivamente em ambiente institucional seguro.

2. A anonimização deve eliminar:

a) Nomes próprios;
b) Números de identificação civil, fiscal ou processual;
c) Moradas e contactos;
d) Datas, dados de localização, identificadores por via eletrónica, um ou mais elementos específicos da identidade física, fisiológica, genética, mental, económica, cultural ou social que permitam identificação contextual;
e) Qualquer elemento factual singular que permita a reidentificação.

3. A mera substituição de nomes por iniciais não constitui anonimização suficiente quando o contexto permita a reidentificação.

D. Segredo de Justiça

1. É expressamente proibida a introdução, em ferramentas não institucionais de IA, de qualquer informação abrangida pelo segredo de justiça.

2. Mesmo após cessação formal do segredo de justiça, mantém-se o dever de reserva relativamente a elementos não públicos do processo.

3. A utilização de IA não pode, direta ou indiretamente, permitir:

a) A identificação do objeto do processo;
b) A reconstrução de diligências processuais;
c) A inferência sobre estratégias investigatórias;
d) A divulgação de prova ainda não produzida em audiência pública.

2.6. Desenvolvimento e Utilização de Ferramentas Institucionais de IA

A. Ferramentas institucionais de IA

1. Consideram-se ferramentas institucionais de inteligência artificial os sistemas de IA desenvolvidos, disponibilizados ou aprovados pela PGR para utilização pelos magistrados do Ministério Público.

2. As ferramentas institucionais de IA devem ser sujeitas a auditorias periódicas, internas e externas, para garantir conformidade com as regras legais aplicáveis e com os princípios estabelecidos nesta Carta.

3. A utilização de ferramentas institucionais de IA deve ser registada em sistema interno próprio, podendo recorrer a registos de utilização, indicando:

a) A ferramenta utilizada;
b) A finalidade da sua utilização;
c) O tipo de dados tratados; e
d) O responsável pela utilização da ferramenta de IA .

B. Avaliação de Impacto antes da Implementação

Antes da implementação de qualquer ferramenta institucional de IA, deverá ser realizada uma Avaliação de Impacto sobre os Direitos Fundamentais que identifique riscos, medidas de mitigação e mecanismos de supervisão humana do sistema a implementar.

C. Supervisão, Auditoria e Monitorização Contínua

1. Será estabelecido um comité de supervisão de inteligência artificial da PGR, responsável pela supervisão ética e legal das ferramentas institucionais de IA.

2. As regras específicas de utilização das ferramentas institucionais de IA são definidas e aprovadas pelo Procurador-Geral da República, sob proposta do comité de supervisão de inteligência artificial da PGR.

D. Tratamento de dados processuais

1. O tratamento de dados processuais pode ocorrer em ferramentas institucionais de IA, com respeito pelos regimes de segredo, designadamente de justiça, e de confidencialidade, e desde que:

a) Tal tratamento seja expressamente autorizado pela PGR, nas regras de utilização ou para finalidade específica;
b) As ferramentas de Inteligência Artificial estejam sujeitas a Avaliação de Impacto sobre Direitos Fundamentais;
c) As ferramentas de IA estejam conformes com o Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho de 13 de junho de 2024 e com as demais regras legais aplicáveis, designadamente em matéria de proteção de dados pessoais; e
d) As ferramentas de IA estejam instaladas em infraestruturas próprias ou sob controlo institucional da PGR.

1. Adicionalmente, as ferramentas de IA classificadas como de "*alto risco*" ao abrigo do Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho de 13 de

junho de 2024 exigem cumprimento integral dos requisitos ali previstos, designadamente quanto a:

a) Gestão de risco;
b) Governança de dados;
c) Documentação técnica;
d) Registo de *logs* e auditoria;
e) Supervisão humana obrigatória.

2.7. Formação

1. Os magistrados do Ministério Público que utilizem sistemas de IA devem receber formação adequada sobre ética, direitos fundamentais e limitações tecnológicas da Inteligência Artificial.

2. A formação pode incidir, designadamente, sobre:

a) A utilização de ferramentas de IA, institucional ou geral;
b) Proteção de dados;
c) Segurança da informação;
d) Riscos de enviesamento algorítmico; e
e) Limites éticos e legais da IA na Justiça.

2.8. Disposição final

Esta Carta deve ser revista periodicamente, tendo em conta o progresso tecnológico, as melhores práticas internacionais e novas orientações europeias ou nacionais.

2. Política de Auditoria e Monitorização Técnica

Política de auditoria e monitorização técnica de sistemas institucionais de Inteligência Artificial

Capítulo I

Objeto e âmbito de Aplicação

Artigo 1.º

(*Objeto*)

1. A presente política de auditoria e monitorização técnica de sistemas institucionais de Inteligência Artificial estabelece os princípios, procedimentos e mecanismos de auditoria e monitorização aplicáveis aos sistemas de Inteligência Artificial (IA) autorizados para utilização no Ministério Público.

2. Consideram-se sistemas institucionais de inteligência artificial os sistemas de IA desenvolvidos, disponibilizados ou aprovados pela Procuradoria-Geral da República (PGR) para utilização pelos magistrados do Ministério Público.

2. Esta política visa assegurar a conformidade legal dos sistemas institucionais de IA, mitigar os riscos técnicos e jurídicos do seu desenvolvimento e disponibilização para utilização e garantir que, em todas as fases, os sistemas de IA dispõem de supervisão humana efetiva.

Artigo 2.º

(*Âmbito de aplicação*)

1. A política de auditoria e monitorização técnica de sistemas institucionais de IA aplica-se a:

a) Sistemas de IA desenvolvidos pela PGR;
b) Sistemas de IA contratados a fornecedores externos;
c) Sistemas de IA integrados em plataformas da infraestrutura tecnológica da PGR, através nomeadamente de soluções de interoperabilidade;
d) Sistemas de IA disponibilizados ao Ministério Público pelo Ministério da Justiça.

2. A presente política aplica-se sempre a sistemas de IA classificados como de risco elevado nos termos do artigo 6.º e anexo III do Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho de 13 de junho de 2024.

Capítulo II

Princípios Orientadores

Artigo 3.º

(*Princípios aplicáveis*)

A política de auditoria e monitorização técnica de sistemas institucionais de IA rege-se pelos seguintes Princípios:

a) Legalidade e Conformidade Permanente;
b) Soberania Digital e dos Dados;
c) Rastreabilidade e Auditabilidade;
d) Transparência Técnica;
e) Transparência e Supervisão Humana Obrigatória;
f) Segurança e Minimização de Dados; e
g) Prevenção de Enviesamento Algorítmico.

Artigo 4.º

(*Princípio da legalidade e conformidade permanente*)

1. Os sistemas institucionais de IA utilizados no Ministério Público devem observar, em todo o seu ciclo de vida, o cumprimento integral do direito da União Europeia (UE) e do direito interno aplicável, em especial do Regulamento (UE) 2024/1689 (Regulamento da Inteligência Artificial) e do Regulamento (UE) 2016/679 (Regulamento Geral sobre a Proteção de Dados).

2. A conformidade legal referida no número anterior é objeto de verificação contínua, mediante auditorias periódicas e sempre que ocorra qualquer modificação relevante do sistema de IA, nomeadamente alterações ao modelo, aos dados de treino, às regras de negócio, às funcionalidades aplicacionais ou à finalidade da sua utilização.

3. A deteção de desconformidades graves determina a suspensão imediata da utilização do sistema até à reposição da conformidade legal e técnica.

Artigo 5.º

(*Princípio da soberania digital e dos dados*)

1. A conceção, aquisição, alojamento, operação e manutenção de sistemas institucionais de IA utilizados no Ministério Público devem assegurar a soberania digital e dos dados da instituição, garantindo o controlo efetivo sobre as infraestruturas tecnológicas, os modelos algorítmicos de IA, os dados de treino, os dados tratados pelo sistema de IA e os dados de utilização.

2. A utilização de infraestruturas, serviços em nuvem, modelos de IA ou componentes tecnológicos fornecidos por entidades terceiras está sujeita a avaliação prévia do risco para a soberania digital, a confidencialidade dos processos, o segredo de justiça e a proteção de dados pessoais.

3. É proibida a transferência, o acesso remoto ou o tratamento de dados processuais, dados pessoais ou metadados relevantes por entidades localizadas fora do Espaço da União Europeia.

4. O tratamento de dados e disponibilização de serviços em sistemas de IA contratados ou protocolados com entidades terceiras só pode ocorrer se estiverem asseguradas garantias adequadas e mecanismos juridicamente válidos de transferência de dados e se tal utilização tiver sido expressamente autorizada pelo Procurador-Geral da República, com base na apreciação do Comité de Supervisão de Inteligência Artificial.

5. Os sistemas institucionais de IA devem, sempre que tal seja tecnicamente viável e economicamente proporcionado, ser alojados em infraestruturas sob controlo direto da PGR ou em ambientes de nuvem soberana localizados no território nacional ou da União Europeia, com garantias de jurisdição e de não sujeição a regimes jurídicos extracomunitários suscetíveis de comprometer o segredo de justiça ou a proteção de dados.

6. É proibida a utilização de sistemas de IA que impliquem:

a) Uma dependência tecnológica estrutural (*lock-in*), incompatível com a autonomia institucional do Ministério Público;
b) A impossibilidade de controlo efetivo em caso de migração de dados e modelos; e
c) A sujeição a obrigações de divulgação de dados a autoridades terceiras em termos incompatíveis com o direito nacional e da União Europeia.

7. Os cadernos de encargos e contratos públicos de desenvolvimento, aquisição ou licenciamento para sistemas institucionais de IA devem conter cláusulas específicas relativas:

a) À localização e jurisdição dos modelos de computação e dos dados treino e de utilização;

b) Ao controlo institucional dos modelos de IA e dos registos (*logs*) e metadados associados à utilização e auditoria do sistema;
c) Ao direito e regime de auditoria técnica;
d) À reversibilidade tecnológica e portabilidade dos dados; e
e) À suspensão ou cessação segura do serviço e eliminação certificada dos dados de utilização.

8. O cumprimento do presente artigo é objeto de verificação periódica no âmbito da política de auditoria e monitorização técnica, devendo qualquer risco relevante para a soberania digital determinar a suspensão da utilização do sistema até à sua mitigação.

Artigo 6.º

(*Rastreabilidade e auditabilidade*)

1. Os sistemas institucionais de IA devem assegurar, por si ou através de serviço aplicacional, a existência de mecanismos de registo automático de todas as operações (*logs*) que permitam a rastreabilidade *ex post* das suas funcionalidades relevantes.

2. Os registos referidos no número anterior devem permitir a reconstrução das interações relevantes entre o sistema institucional de IA e os respetivos utilizadores, incluindo, designadamente:

a) Os dados de entrada (*inputs*);
b) Os resultados produzidos (*outputs*);
c) A versão do modelo utilizada;
d) A identificação do utilizador interveniente; e
e) A data e hora da operação.

3. Os registos devem ser conservados em ambiente seguro, com garantias de integridade, confidencialidade e controlo de acessos, por prazo não inferior a 5 anos.

4. É proibida a utilização de sistemas de IA que, pela sua arquitetura técnica, impeçam ou dificultem significativamente a auditoria institucional ou jurisdicional.

Artigo 7.º

(*Transparência técnica*)

1. Os sistemas institucionais de IA devem ser acompanhados de documentação técnica suficiente para permitir a compreensão funcional do seu modo de funcionamento pela PGR, pelo Comité de Supervisão de Inteligência Artificial, pelas equipas técnicas e pelos utilizadores, designadamente os magistrados do Ministério Público.

2. A documentação referida no número anterior deve incluir, nomeadamente:

a) A descrição da finalidade do sistema de IA;
b) As funcionalidades e os limites de utilização do sistema de IA;
c) As fontes e a natureza dos dados de treino;
d) As métricas de desempenho e fiabilidade; e
e) O compêndio dos riscos identificados e respetivas medidas de mitigação.

3. Não é permitida a utilização de sistemas de IA cuja explicabilidade funcional mínima seja recusada pelo Comité de Supervisão de Inteligência Artificial ou seja tecnicamente inviável, em termos que comprometam o exercício da supervisão humana efetiva do sistema de IA.

Artigo 8.º

(*Transparência e supervisão humana obrigatória*)

1. A utilização de sistemas institucionais de IA está sujeita a supervisão humana efetiva, não podendo tais sistemas substituir o exercício funcional dos magistrados do Ministério

Público, ao nível da tramitação, da decisão, da valoração jurídica ou da apreciação da prova.

2. É proibida qualquer forma de automatização decisória materialmente vinculativa em matéria que afete direitos, liberdades e garantias, ou que interfira com o exercício da ação penal e demais competências próprias do Ministério Público.

3. A utilização por magistrados de sistemas de IA no âmbito das funções próprias do Ministério Público, designadamente em sede processual, deve ser objeto de informação no processo, indicando pelo menos a ferramenta de IA e a finalidade para a qual aquela foi utilizada, em termos a definir por Diretiva do Procurador-Geral da República.

4. A utilização processual de sistema de IA por magistrados do Ministério Público é passível de verificação, assegurando-se a possibilidade de revisão, correção ou rejeição dos *outputs* do sistema, garantindo o direito de revisão por parte dos demais intervenientes processuais.

5. Os utilizadores institucionais devem receber formação adequada em literacia digital e algorítmica, nos termos a definir pelo Conselho Superior do Ministério Público.

Artigo 9.º

(*Segurança e minimização de dados*)

1. Os sistemas institucionais de IA devem respeitar os princípios da minimização de dados, da integridade e da confidencialidade, nos termos do artigo 5.º do RGPD.

2. O tratamento de dados pessoais no âmbito de sistemas de IA está sujeito a medidas técnicas e organizativas adequadas de segurança, nos termos do artigo 32.º do RGPD e do artigo 15.º do Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho de 13 de junho de 2024.

3. É proibida a introdução de dados reais de processos judiciais ou de inquérito penal em ferramentas de IA não institucionais de uso geral.

4. A utilização de dados pessoais em sistemas de IA depende da realização prévia de avaliação de impacto sobre a proteção de dados, nos termos do artigo 35.º do RGPD, sempre que o tratamento de dados seja suscetível de implicar elevado risco para os direitos, liberdades e garantias dos respetivos titulares.

5. Devem ser adotadas medidas de pseudonimização ou anonimização sempre que tal seja tecnicamente possível e compatível com a finalidade do sistema de IA.

Artigo 10.º

(*Prevenção de enviesamento algorítmico*)

1. Os sistemas institucionais de IA devem ser concebidos, avaliados e monitorizados de forma a prevenir, detetar e mitigar enviesamentos suscetíveis de produzir efeitos discriminatórios ou de comprometer a imparcialidade nas decisões do Ministério Público.

2. Para efeitos do disposto no número anterior, devem ser realizados testes periódicos de enviesamento, incluindo:

a) Análise da representatividade e neutralidade dos dados de treino;
b) Avaliação de variáveis-proxy potencialmente discriminatórias; e
c) Verificação de impactos diferenciados dos *outputs* do sistema de IA sobre grupos vulneráveis.

3. A deteção de enviesamento relevante determina a suspensão do sistema até à implementação de medidas corretivas adequadas.

4. Os resultados dos testes periódicos de enviesamento devem ser objeto de relatório técnico submetido ao Comité de Supervisão de Inteligência Artificial, para efeito de avaliação e mitigação dos efeitos de enviesamento que venham a ser detetados.

Capítulo III

Estrutura de Governação de Sistemas Institucionais de IA

Artigo 11.º

(*Comité de supervisão de inteligência artificial*)

1. A governação dos sistemas institucionais de IA é efetuada através do Comité de Supervisão de Inteligência Artificial (CSIA).

2. O CSIA é criado por despacho do Procurador-Geral da República, tendo composição multidisciplinar e contendo na sua composição:

a) Dois magistrados do Ministério Público;
b) O responsável de segurança da informação da Procurador-Geral da República;
c) O encarregado de proteção de dados da PGR;
d) Dois especialistas de sistemas e tecnologias de informação da PGR, sendo um da área do desenvolvimento aplicacional e outro da área da administração de sistemas.

3. O CSIA poderá ainda integrar especialistas técnicos em IA e especialistas com competências em proteção de dados designados para o efeito.

4. O funcionamento do Comité de Supervisão de Inteligência Artificial será objeto de regulamento próprio, aprovado pelo Procurador-Geral da República.

Artigo 12.º

(*Competência geral do Comité de Supervisão de Inteligência Artificial*)

1. Compete ao CSIA assegurar que a conceção, desenvolvimento, aquisição, implementação e utilização dos sistemas institucionais de IA ocorrem em conformidade

com os princípios e regras legais e regulamentares aplicáveis, a carta ética de utilização de IA pelo Ministério Público e a presente política de auditoria e monitorização técnica de sistemas institucionais de IA.

2. Compete ao CSIA, designadamente:

a) Aprovar previamente os sistemas institucionais de IA;
b) Monitorizar os sistemas institucionais de IA;
c) Ordenar e validar a realização de auditorias aos sistemas institucionais de IA;
d) Proceder à avaliação de impacto de sistemas institucionais de IA; e
e) Recomendar a suspensão e inativação de sistemas institucionais de IA.

3. A competência do CSIA abrange todo o ciclo de vida dos sistemas institucionais de IA do Ministério Público.

Artigo 13.º

(*Proteção de dados pessoais, segredo de justiça e outros segredos legalmente protegidos*)

1. O CSIA articula-se com o Encarregado de Proteção de Dados (EPD) da PGR para validação da avaliação de Impacto sobre a Proteção de Dados (AIPD) e para o acompanhamento da conformidade dos sistemas institucionais de IA com o RGPD.

2. Compete ao CSIA, em relação aos sistemas institucionais de IA e respetivos modelos de dados e de dados de treino:

a) Verificar a aplicação do princípio da minimização de dados;
b) Assegurar a compatibilidade dos fluxos de dados com o regime do segredo de justiça;
c) Avaliar riscos de transferência internacional de dados;
d) Determinar medidas adicionais de pseudonimização, anonimização ou segregação de ambientes.

3. O CSIA pode determinar a suspensão imediata do sistema institucional de IA em caso de risco grave para os direitos dos titulares dos dados pessoais ou de risco grave de violação do segredo de justiça.

Artigo 14.º

(*Supervisão humana, formação e cultura institucional*)

1. Compete ao CSIA assegurar que os sistemas institucionais de IA são configurados para permitir uma supervisão humana efetiva.

2. Em qualquer sistema institucional de IA, o CSIA deve validar:

a) Os requisitos mínimos de validação humana dos *outputs* do sistema institucional de IA, e
b) Os perfis de acesso e níveis de autonomia do sistema institucional de IA.

3. A PGR promove uma cultura institucional de utilização crítica e responsável da IA, cabendo ao CSIA emitir propostas de guias práticos, orientações internas e planos de formação em literacia algorítmica ao Procurador-Geral da República.

Artigo 15.º

(*Transparência técnica e explicabilidade institucional*)

1. Compete ao CSIA exigir aos fornecedores externos e/ou equipas internas de desenvolvimento aplicacional a disponibilização de documentação adequada e atualizada para garantir a explicabilidade técnica e funcional dos sistemas institucionais de IA.

2. Para tanto, o CSIA deve:

a) Recusar a autorização prévia de sistemas sem descrição adequada do processo *input-output* ("*caixa-negra*" não auditável); e

b) Impor requisitos mínimos de explicabilidade na documentação técnica e funcional do sistema de IA.

3. O CSIA pode implementar modelos de fichas técnicas institucionais descritivas do funcionamento dos sistemas institucionais de IA, cujo preenchimento deve ser assegurado pelos fornecedores dos sistemas, internos ou externos.

4. A insuficiência de transparência técnica e/ou funcional constitui fundamento de recusa de autorização prévia do sistema de IA e a falta da sua atualização pode implicar suspensão ou inativação do sistema.

Artigo 16.º

(*Documentação do ciclo de vida do sistema institucional de IA*)

1. O CSIA organiza e mantém atualizado um *dossier* contendo toda a documentação técnica e funcional referente ao sistema institucional de IA.

2. O *dossier* é composto por, pelo menos, os seguintes elementos:

a) Cadernos de encargos e relatórios de avaliação procedimental;
b) Manuais técnicos e funcionais do sistema, entregues pelos fornecedores internos ou externos;
c) Histórico de atualizações do sistema de IA;
d) Memória descritiva do sistema de IA, sua finalidade e classificação de risco;
e) Relatórios de autorizações e auditorias ao sistema;
f) Relatórios de incidentes; e
g) Relatórios de medidas de mitigação de risco implementadas.

Artigo 17.º

(*Avaliação e autorização prévia de sistemas institucionais de IA*)

1. Compete ao CSIA proceder à avaliação prévia da conformidade legal, ética e técnica de qualquer sistema institucional de IA, desenvolvido, protocolado ou adquirido pela PGR, antes da sua utilização pelo Ministério Público.

2. A autorização prévia a sistema institucional de IA depende da verificação cumulativa dos seguintes requisitos:

a) Conformidade com os requisitos legais e regulamentares aplicáveis, designadamente quando estejam em causa sistemas de IA de alto risco;
b) Avaliação de impacto sobre a proteção de dados;
c) Compatibilidade com o cumprimento das regras sobre o segredo de justiça e a confidencialidade processual;
d) Respeito pelos princípios da supervisão humana, transparência técnica e não discriminação e soberania digital; e
e) Conformidade com a política de auditoria e monitorização técnica de sistemas institucionais de IA e a política de segurança da informação.

3. A autorização prévia pode ser:

a) Plena;
b) Condicionada a medidas técnicas e organizativas adicionais; ou
c) Temporária, sob monitorização permanente.

4. Pode ser proposta pelo CSIA ao Procurador-Geral da República a recusa de implementação de sistema de IA avaliado negativamente, mediante decisão fundamentada.

Artigo 18.º

(*Determinação e matriz de risco de sistemas institucionais de IA*)

1. Compete ao CSIA aprovar a matriz de risco institucional aplicável a sistemas de IA utilizados no Ministério Público, assegurando a sua coerência com o sistema de gestão de riscos previsto no artigo 9.º do Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho de 13 de junho de 2024.

2. Na determinação do risco de sistemas institucionais de IA, compete ao CSIA:

a) Classificar os sistemas institucionais de IA quanto ao nível de risco jurídico, ético, tecnológico e reputacional;
b) Definir medidas proporcionais de mitigação;
c) Impor limites funcionais ou de contexto de utilização e funcionalidades dos sistemas institucionais de IA;
d) Determinar a necessidade de auditorias extraordinárias para sistemas institucionais de IA de maior risco.

Capítulo IV

Auditorias

Artigo 19.º

(*Regra Geral*)

1. Os sistemas institucionais de IA são obrigatoriamente sujeitos a auditoria previamente à sua implementação e durante todo o seu ciclo de vida.

2. As auditorias são determinadas e coordenadas pelo CSIA e podem ser realizadas por equipa interna da área informática da PGR, com assessoria técnica e/ou jurídica, designadamente de auditor de cibersegurança, ou por equipa externa contratada para o efeito.

3. Todas as auditorias a sistemas de IA comportam as seguintes vertentes:

a) Jurídica, incluindo designadamente a conformidade com as normas legais e regulamentares aplicáveis;
b) Técnica, incluindo designadamente os vetores de robustez aplicacional, tecnologia empregue, interoperabilidade interna e externa, integridade dos modelos e dados de treino, cibersegurança e registos de utilização;
c) Organizativa, incluindo designadamente os aspetos de gestão, perfis e regimes de acesso, incluindo o controlo de acesso aplicacional, edição e monitorização de funcionalidades e conteúdos e propostas de formação; e
d) Funcional, incluindo designadamente a análise das principais funcionalidades, modos de utilização, adequação das funcionalidades às finalidades próprias da ação do Ministério Público, podendo ainda incluir a realização de testes com utilizadores da PGR.

Artigo 20.º

(*Auditoria prévia*)

1. Previamente à implementação em produção de qualquer sistema institucional de IA, deve ser realizada uma auditoria prévia, que deve incluir, para além dos requisitos gerais, os seguintes itens de avaliação:

a) Impacto sobre direitos fundamentais;
b) Impacto sobre proteção de dados pessoais;
c) Análise de classificação de risco;
d) Testes de robustez e cibersegurança;
e) Testes de enviesamento algorítmico; e
f) Avaliação de documentação e explicabilidade do sistema de IA.

2. Nenhum sistema institucional de IA pode ser implementado sem relatório favorável de auditoria prévia.

Artigo 21.º

(*Auditoria periódica*)

1. Todos os sistemas institucionais de IA em produção são objeto de auditoria periódica de monitorização, a qual deve ocorrer com a seguinte periodicidade:

a) Bienal, para sistemas institucionais de IA classificados como sistemas de risco moderado; e
b) Anual, para sistemas institucionais de IA classificados como sistemas de risco elevado.

2. Para além dos requisitos gerais, as auditorias periódicas devem incluir os seguintes itens de avaliação:

a) Verificação de registos (*logs*) de utilização e resultado;
b) Testes de precisão e fiabilidade;
c) Avaliação de eventuais desvios de desempenho;
d) Verificação de conformidade com os limites estipulados de utilização; e
e) Análise de incidentes registados.

3. O CSIA valida os relatórios de auditoria periódica positiva e acompanha a execução das medidas corretivas assinaladas.

Artigo 22.º

(*Auditoria extraordinária*)

1. O CSIA pode determinar a realização de auditorias extraordinárias sempre que:

a) Ocorra incidente de segurança;
b) Ocorra situação de violação de dados pessoais;

c) Seja detetado um enviesamento relevante;
d) Se verifiquem indícios de desconformidade legal;
e) Exista alteração substancial do modelo de IA ou das condições de funcionamento do sistema de IA.

2. O CSIA valida os relatórios de auditoria extraordinária e, não sendo caso de recomendação da suspensão ou inativação do sistema institucional de IA, acompanha a execução das medidas de mitigação até à eliminação das desconformidades assinaladas.

Capítulo IV

Monitorização

Artigo 23.º

(*Monitorização contínua*)

1. O CSIA é responsável e supervisiona a monitorização contínua dos sistemas institucionais de IA em produção, competindo-lhe designadamente:

a) Definir indicadores de risco e desempenho;
b) Analisar relatórios periódicos de funcionamento;
c) Acompanhar incidentes, designadamente incidentes graves;
d) Determinar a comunicação de incidentes às entidades competentes, quando legalmente exigido.

2. Todos os sistemas institucionais de IA devem ser concebidos para produzir e manter, na própria aplicação ou através de serviço aplicacional, os seguintes registos automáticos de utilização (*logs*):

a) Utilizador;
b) Data e hora de utilização;
c) Finalidade declarada;

d) Registo de *input*;
e) Registo de *output*;
f) Tipo de operação realizada; e
g) Versão do sistema e modelo de IA utilizado.

3. Para além dos registos de utilização, devem ser gerados os seguintes indicadores de desempenho do sistema institucional de IA:

a) Percentagem de erro factual;
b) Taxa de incidência de respostas não conformes;
c) Frequência de *outputs* rejeitados por revisão humana, e
d) Identificação de padrões discriminatórios.

4. Os registos indicados nos números anteriores devem ser:

a) Imutáveis;
b) Mantidos em base de dados com encriptação dos registos nativa, salvo quando acedidos legitimamente;
c) Conservados pelo período de 5 anos; e
d) Acessíveis apenas ao CSIA e a entidades autorizadas.

5. Todo o acesso aos registos previstos neste artigo é registado automaticamente em base de dados autónoma, devendo conter os seguintes elementos:

a) Utilizador;
b) Data e hora de consulta;
c) Finalidade da consulta;
d) Registo da consulta.

Capítulo V

Gestão de Incidentes

Artigo 24.º

(*Incidente em sistema institucional de IA*)

1. Considera-se incidente em sistema institucional de IA qualquer evento, falha, anomalia, uso indevido, comportamento inesperado ou disfunção técnica ou organizativa, ocorrida durante a conceção, implementação, operação, monitorização ou desativação de um sistema de IA, que:

a) Comprometa ou possa comprometer a conformidade legal do sistema institucional de IA;
b) Afete ou possa afetar a segurança, integridade, disponibilidade, rastreabilidade ou fiabilidade do sistema;
c) Coloque em risco, ainda que potencialmente, a proteção de dados pessoais;
d) Afete a qualidade, correção ou previsibilidade dos *outputs* do sistema; ou
e) Comprometa o cumprimento das demais regras e princípios aplicáveis, designadamente quanto à segurança da informação, soberania digital ou supervisão humana.

2. Constituem, nomeadamente, incidentes, os seguintes eventos:

a) Falhas técnicas relevantes do modelo;
b) Erros sistemáticos de *output*;
c) Acessos não autorizados aos sistemas institucionais de IA ou aos registos de utilização (*logs*);
d) Perda de integridade de dados;
e) Desvios ao funcionamento autorizado;
f) Utilização fora do âmbito funcional autorizado; e
g) Falhas nos mecanismos de registo de utilização (*logs*) ou de supervisão humana.

3. A ocorrência de incidente em sistema institucional de IA é objeto de registo, análise e avaliação pelo CSIA, designadamente nos termos dos procedimentos de auditoria.

Artigo 25.º

(*Incidente grave em sistema institucional de IA*)

1. Considera-se incidente grave em sistema institucional de IA todo o incidente que, pela sua natureza, impacto ou potencial de impacto, envolva:

a) Risco elevado ou efetivo para direitos, liberdades e garantias constitucionalmente protegidos;
b) Violação do segredo de justiça, de outros segredos legalmente protegidos ou da confidencialidade processual;
c) Violação de dados pessoais suscetível de implicar elevado risco para os titulares dos dados;
d) Enviesamento algorítmico relevante com impacto discriminatório ou potencialmente discriminatório;
e) Falha de supervisão humana suscetível de conduzir a decisões automatizadas materialmente vinculativas;
f) Comprometimento grave da segurança, integridade ou disponibilidade do sistema institucional de IA; ou
g) Desconformidade grave com os requisitos aplicáveis a sistemas de IA de alto risco previstos nos artigos 8.º a 15.º Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho de 13 de junho de 2024.

2. São sempre classificados como incidentes graves os seguintes eventos:

a) Qualquer acesso, transferência ou tratamento de dados processuais fora da UE;
b) Utilização de dados reais de processos judiciais em ferramentas institucionais de IA não autorizadas para esse efeito;
c) Impossibilidade de auditoria ou rastreabilidade de *outputs* relevantes, ou

d) Ocorrência de falhas técnicas que alterem significativamente o funcionamento do sistema institucional de IA.

3. A verificação de incidente grave determina:

a) A suspensão imediata da utilização do sistema institucional de IA;
b) A ativação pelo CSIA do plano de resposta a incidentes;
c) A comunicação às entidades competentes quando legalmente exigido;
d) A abertura de processo interno de averiguação e reporte ao Procurador-Geral da República.

4. Em caso de incidente grave, para além do mais, o CSIA pode impor medidas imediatas de mitigação e proceder à suspensão cautelar da utilização do sistema institucional de IA.

Artigo 26.º

(*Deveres de comunicação e resposta a incidente em sistema institucional de IA*)

1. Qualquer utilizador institucional que tenha conhecimento de incidente ou incidente grave deve comunicá-lo de imediato ao CSIA.

2. Os incidentes graves são comunicados pelo CSIA ao Encarregado de Proteção de Dados para efeitos de avaliação das obrigações de notificação nos termos do RGPD, e ao responsável pela segurança da informação e cibersegurança da PGR.

3. O CSIA aprova e mantém atualizado um Plano de Resposta a Incidentes em sistema institucional de IA, contendo designadamente:

a) Os procedimentos de contenção de danos;
b) As medidas de mitigação aplicáveis;
c) Os fluxos de comunicação interna;

d) Os critérios e modelos de reporte a entidades externas, designadamente à Comissão Nacional de Proteção de Dados (CNPD) e ao Centro Nacional de Cibersegurança (CNCS), e
e) Os procedimentos de reposição de conformidade do sistema institucional de IA.

Capítulo VI

Contratação Pública

Artigo 27.º

(*Soberania digital e contratação pública de sistemas IA*)

1. O CSIA deve pronunciar-se, previamente à decisão de adjudicação em procedimento pré-contratual de aquisição de sistema de IA ou de desenvolvimento de sistema de IA, sobre:

a) Os riscos para a soberania digital;
b) A dependência tecnológica estrutural (*lock-in*);
c) A jurisdição e localização dos dados do sistema, incluindo os dados de treino; e
d) As cláusulas contratuais referentes a auditoria, reversibilidade e eliminação de dados.

2. É obrigatório um parecer do CSIA previamente à decisão de contratar quando estejam em causa sistemas de IA de alto risco.

Artigo 28.º

(*Cláusulas contratuais e controlo de fornecedores externos de sistemas IA*)

1. Os cadernos de encargos e contratos públicos para a aquisição ou desenvolvimento de sistemas institucionais de IA devem prever sempre:

a) A garantia de tratamento de dados exclusivamente na UE;
b) A proibição absoluta de reutilização de dados para treino;
c) Obrigatoriedade de sujeição do sistema a auditorias internas e externas;
d) Obrigação de confidencialidade e respeito pelo segredo de justiça,
e) Direito de inspeção técnica.

2. O CSIA pode propor à Secretaria-Geral da PGR cláusulas específicas para garantia da conformidade legal e sustentabilidade técnica dos sistemas institucionais de IA que venham a ser adquiridos ou desenvolvidos por via da contratação pública.

3. O sistema institucional de IA objeto de contratação pública pode, sob proposta do CSIA, ser suspenso ou mesmo descontinuado, para além das demais causas que venham a ser detetadas em auditoria, monitorização contínua ou por ocorrência de incidente grave, quando deixe de cumprir os requisitos legais e/ou exista incumprimento contratual do fornecedor.

Capítulo VII

Disposições Finais

Artigo 29.º

(*Revisão da Política*)

A presente política de auditoria e monitorização técnica de sistemas institucionais de IA deve ser revista bienalmente ou sempre que exista alteração legislativa relevante da UE ou nacional.