

BridgeAI: Boosting Regulatory Implementation with Data-driven insights, Global expertise, and Ethics for AI

Ana Oliveira¹, André Martins², André Valente³, Andreia Brites⁴, António Novais⁵, Bruno Ceña¹, Bruno Silva¹, Carlos Antunes³, Carme Puche Moré³, Catarina Farinha⁵, Catarina Pimentel⁴, Cristina Ferreira³, Duarte Almeida⁶, Federico Lyra⁴, Giovanni De Gregorio⁷, Helena Moniz¹, Henrique Sousa Antunes⁷, Iakovina Kindylidi⁷, Isabel Caetano⁶, Isabel Trancoso⁴, Joana Lamego⁵, Joana Portugal⁵, João Marques Martins⁴, Joe Paton⁴ (Coord.), John Krakauer², Kritika Maheshwari³, Luís Barreto Xavier⁶, Luísa Coheur⁴, Madalena Gomes Cruz⁷, Magda Cocco⁶, Maria Ana Henriques¹, Maria Roquette¹, Mariana Dalblon⁷, Mário Campolargo², Mário Figueiredo², Nicholas Kuge⁶, Nicolau Borges⁴, Nuno André¹, Paulo Dimas², Paulo Sousa Mendes⁶, Pedro Conceição², Pedro Santa Clara⁴, Pedro Vale Gonçalves^{6, 7} (Coord.), Raúl Araújo¹, Razvan Sandru³ (Coord.), Rodrigo Abreu¹, Sara Guerreiro de Sousa⁵ (Coord.), Sara Sá¹, Tito Rendas⁷, Vera Lúcia Raposo³, Virginia Dignum².

¹GT0, ²Concultores, ³GT2, ⁴GT4, ⁵GT1, ⁶GT3, ⁷GT5

Resumo executivo

Problema

O Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024, que estabelece regras harmonizadas em matéria de inteligência artificial e altera determinados atos legislativos da União (designado como “AI Act”), entrou em vigor em agosto de 2024 com uma implementação faseada, de modo a permitir aos intervenientes tempo de adaptação às diferentes fases. O Governo português, a indústria, a sociedade civil e os cidadãos em geral necessitam de compreender o Regulamento e os impactos que terá, tanto a nível nacional como internacionalmente, aquando da sua implementação total em 2027. Existem diversos desafios na preparação da implementação do AI Act:

- A Europa está a ter um papel de liderança na legislação de Inteligência Artificial (IA), promovendo em simultâneo a inovação.
- Pela primeira vez na História da Humanidade, os sistemas estão a produzir conteúdos e esses poderão causar malefícios. Estes desenvolvimentos requerem enquadramentos éticos e morais novos e novas formas transversais de pensar o trabalho, a educação ou a saúde, o que se mantém sobremaneira inexplorado.
- Serviços de IA que apoiam o processo de tomada de decisão trazem benefícios, mas também colocam riscos para as pessoas, o que enfatiza a crescente necessidade de literacia em IA para todos, incluindo administração pública, indústria e sociedade civil.

- Portugal pode liderar internacionalmente ao criar produtos responsáveis.

Principais resultados de investigação

A IA Responsável (Dignum, 2019) é um paradigma que procura solucionar os desafios descritos e que permite antecipar o AI Act:

- A IA Responsável reforça o aumento de competências e o desenvolvimento dos cidadãos em áreas emergentes (Dignum, 2019; UN Governing AI for Humanity, 2024);
- Enquadramentos éticos em conjunto com modelos de governança dinâmicos, adaptáveis, escaláveis e participativos (AI4People, Floridi et al., 2018) permitem o alinhamento dos produtos com princípios éticos e com padrões legais exigentes, promovendo inovação com confiança;
- Casos de estudo provenientes da indústria mostram que a IA Responsável promove inovação, com base em metodologias como as descritas no catálogo de risco da OCDE;
- Portugal pode aproveitar as ferramentas de avaliação de risco já existentes (por exemplo, o Guia, desenvolvido pela AMA/ARTE), que podem ser atualizadas para apoiar a gestão transversal de riscos de IA;
- A literacia e a capacitação dos cidadãos serão elementos-chave e implicam uma primeira avaliação sólida do que se sabe a nível nacional (DESI e Década Digital);
- Portugal, enquanto pólo de inovação, pode estar estritamente ligado aos ecossistemas europeu e internacional.

Recomendações principais

A implementação do AI Act requer o envolvimento de vários atores, incluindo a administração pública, a indústria e a sociedade civil. Este processo deve focar-se em ética, direito e literacia e em simultâneo na promoção da inovação. As recomendações foram desenvolvidas em grupos de trabalho (GT) com equipas multidisciplinares e correspondem a:

GT1 - Ferramentas de avaliação de risco

Recomendação #1: Identificar e priorizar os casos de uso de IA urgentes para avaliação de riscos.

Recomendação #2: Atualizar o Guia, para que se torne uma ferramenta relevante, prática e flexível e para mapear e mensurar os riscos da IA, tornando-a obrigatória no processo de licitação pública (Guia 2.0).

Recomendação #3: Centralizar a avaliação de riscos da IA numa única autoridade pública.

Recomendação #4: Estabelecer parcerias com organizações terceiras.

GT2 - Ética e processos regulatórios

Recomendação #1: Modelo dinâmico de governação baseado no AI4People.

Recomendação #2: Zonas Livres Tecnológicas (ZLT, do inglês “sandboxes”) com supervisão ética.

GT3 - AI Act interface e implementação

Recomendação #1: Atualizar as políticas e os programas de formação do setor público

Recomendação #2: Estabelecer uma rede de apoio de peritos

Recomendação #3: Fortalecer a monitorização e planeamento da defesa cibernética

Recomendação #4: Atualizar e harmonizar a documentação de políticas nacionais sobre IA

GT4 - Literacia em IA e capacitação avançada

Recomendação #1: Lançar uma campanha de literacia em IA, adaptada a vários públicos

Recomendação #2: Promover parcerias público-privadas e apoio governamental

Recomendação #3: Avaliar o nível de literacia em IA em Portugal através de um inquérito nacional

GT5 - Esforços regulatórios para além da Europa

Recomendação #1: Estabelecer Portugal como um pólo de coordenação de ZLT.

Recomendação #2: Coordenação internacional através de parcerias europeias de confiança

Recomendação #3: Desenvolver plataformas dedicadas que permitam e facilitem o licenciamento de obras protegidas para efeitos de treino de IA.

Recipientes

Ministro Adjunto do Gabinete do Primeiro Ministro e da Reforma do Estado, Agência para a Reforma Tecnológica do Estado, IP (ARTE, formerly AMA), Ministra da Juventude e Modernização, Ministro da Educação e Ciência, Ministro da Economia, Centro de Cibersegurança Nacional, Autoridade Nacional de Comunicações (ANACOM).

GT1 – Ferramentas de avaliação de risco

O Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho Europeu introduz uma estrutura baseada em risco, que define diferentes níveis de risco de acordo com os usos específicos dos sistemas de IA, proibindo aplicações de IA com risco inaceitável e introduzindo obrigações rigorosas para sistemas de IA de alto risco. Apesar de ser uma abordagem flexível, que permite que o regulamento se adapte à evolução tecnológica, o Regulamento não é facilmente aplicável, pois deixa muitos conceitos abertos a interpretações e muitos outros a serem operacionalizados por normas harmonizadas (que ainda estão a ser elaboradas).

Com base na análise de casos de estudo concretos de diversos parceiros da indústria (Priberam, Sword Health, Unbabel e Youverse) de áreas diferenciadas e diferentes níveis de avaliação de risco, que vão desde aplicações na área da saúde até à biometria, o GT1 estabeleceu uma metodologia para avaliar riscos sem perder de vista o desenvolvimento de produtos inovadores.

Em colaboração com a Agência de Modernização Administrativa (AMA), agora denominada ARTE, o GT1 testou a viabilidade de utilização da GuIA, uma ferramenta para avaliação de riscos éticos, para uma utilização transversal da ferramenta em serviços públicos e privados.

Com base na avaliação multidisciplinar realizada (envolvendo todos os GT), determinou-se também que, para avaliar adequadamente os riscos e equilibrá-los com a inovação, é necessário um ecossistema de especialistas de excelência, que já se encontra disponível em Portugal.

A partir da análise de estudos de caso, ficou claro que uma abordagem única para a avaliação de riscos da IA é inadequada, uma vez que diferentes aplicações exigem diferentes níveis de escrutínio, avaliação e mecanismos de mitigação de riscos também eles diferentes.

Para lidar com as complexidades da avaliação de riscos de IA, Portugal deve priorizar estratégias práticas, escaláveis e inovadoras. São passos essenciais: identificar os casos de uso de IA mais urgentes, atualizar a [GuIA](#) para uma ferramenta de avaliação de riscos dinâmica e acionável, centralizar a governação de riscos de IA e fomentar colaborações público-privadas. Estas medidas não só irão melhorar a adoção responsável de IA na administração pública, como também posicionar Portugal como líder em inovação em IA.

Ao adotar uma abordagem estruturada e colaborativa, Portugal pode garantir que a IA serve o interesse público, fortalece a capacidade institucional e contribui para o crescimento económico – transformando, em

última análise, os desafios regulamentares em oportunidades para uma inovação responsável.

GT2 – Ética e processos regulatórios

O lento desenvolvimento de normas éticas e a sua eventual formalização em estruturas vinculativas contrastam com a rápida evolução das tecnologias. Isso é particularmente relevante no caso de IA, caracterizada por um crescimento exponencial e adoção pelo público em geral, incluindo órgãos públicos e indústria. Devido a essa rápida evolução e à capacidade dos sistemas de IA desenvolverem novas funcionalidades (Bommasani et al., 2021), ou seja, funcionalidades não previstas com base no treino inicial, os riscos ou desafios que podem representar são frequentemente imprevisíveis. Em contrapartida, as estruturas e diretrizes éticas clássicas propõem um modelo de governação estático e hierárquico que promove a aplicação de um conjunto fixo de regras a um amplo espectro de fenómenos. Os modelos “top-down” (diretrizes estabelecidas de cima para baixo) dependem da imposição de um conjunto preestabelecido de normas – frequentemente como uma solução “única para todos” – por uma autoridade centralizada. No entanto, argumenta-se há muito tempo que essas soluções “únicas para todos” são ineficientes para lidar com fenómenos complexos ou com requisitos e condições locais (Homsy et al., 2019). As abordagens “bottom-up” (participação de baixo para cima) têm sido comumente propostas como alternativas aos modelos de “top-down”, em que atores descentralizados – indivíduos, associações ou instituições comerciais – podem optar pela autorregulação, desenvolvendo o seu próprio conjunto de diretrizes éticas. Contudo, as abordagens “bottom-up” estão sujeitas a dificuldades, como a incompatibilidade com regulamentações nacionais ou internacionais, ou a falta de reconhecimento ou adoção. Os modelos estáticos de governança “top-down” não conseguem acompanhar o rápido desenvolvimento da IA, nem os riscos emergentes que esta possa representar,

enquanto as abordagens “bottom-up” correm o risco de não terem apoio institucional ou impacto a longo prazo (Floridi, 2019).

O GT2 propõe um modelo de governação adaptável a questões emergentes, que incentiva a participação “bottom-up”, para facilitar a identificação de problemas emergentes, e escalável, para atender à ampla implementação de sistemas de IA. Seguindo trabalhos já existentes na área – principalmente no âmbito do quadro *AI4People* (Floridi et al., 2018; Pagallo et al., 2019) – propomos um sistema de governação multinível que: a) aborde, primeiramente, os desafios e riscos imediatos relativos a direitos e valores essenciais de máxima prioridade (por exemplo, dignidade humana, integridade e privacidade), que devem ser respeitados em qualquer desenvolvimento de IA; b) incentive contribuições “bottom-up” de entidades civis, comerciais e científicas com base em dados pós-comercialização, algoritmos e avaliações de impacto; c) implemente mecanismos de coordenação que permitam o diálogo mutuamente enriquecedor entre entidades governamentais e civis, comerciais ou científicas, por meio dos quais a legislação existente possa ser atualizada, para atender às necessidades emergentes do desenvolvimento e da implementação de IA. O modelo de governação delineado foi elaborado com base em desafios existentes e possivelmente emergentes que os sistemas de IA podem apresentar, identificados numa análise cuidadosa de estudos de caso dos parceiros da indústria, realizada pelo GT2.

O apoio do IA Act às ZLT, ambientes de testagem (do inglês “sandboxes”) é fundamental para equilibrar a inovação com a supervisão ética. Os ambientes de testagem permitem melhorias iterativas, conformidade legal e salvaguardas sociais, fomentando a inovação e minimizando danos não

intencionais. Recomendamos que as “sandboxes” sejam implementadas com serviços de consultoria ética e jurídica integrados, que auxiliem os desenvolvedores a sanar eventuais problemas que possam ter passado despercebidos. Tais serviços podem incentivar os desenvolvedores, especialmente os menos favorecidos, a optar por projetos

eticamente conscientes e, conseqüentemente, promover uma inovação justa e responsável.

Uma abordagem holística e adaptativa que integre previsão ética, sensibilidade contextual e colaboração das partes interessadas é essencial para concretizar o potencial da IA, mitigando os seus riscos.

GT3 – AI Act interface e implementação

O Regulamento (UE) 2024/1689 da União Europeia sobre Inteligência Artificial é a referência mundial para a governação responsável de IA (“IA Act”), exigindo uma reforma institucional imediata e profunda a nível nacional. Ao entrarmos no último trimestre de 2025, a Administração Pública Portuguesa (AP) já opera ao abrigo de disposições-chave deste regulamento histórico. O objetivo central deste relatório é traçar o caminho para a plena conformidade a nível nacional, assegurando que o nosso compromisso com o fomento da inovação seja rigorosamente equilibrado com o imperativo constitucional de salvaguardar os direitos fundamentais e manter a integridade administrativa. A governação da IA não é meramente uma questão técnica; é uma prioridade política que exige uma ação legislativa estratégica, recursos dedicados e uma liderança decisiva.

Conforme descrito no documento [UN Governing AI for Humanity \(2024\)](#), a legislação sobre IA é um esforço global e coordenado. A UE parece atuar isoladamente nesse processo, mas o AI Act é uma declaração forte e necessária da Europa, que serviu de base para iniciativas globais. Apesar das diferenças globais (abordadas no GT5), o Regulamento estabelece o mote para diversos esforços e estratégias coordenadas a nível global. Portanto, as nossas recomendações estão alinhadas com as melhores práticas globais, permitindo que a administração pública (e indústrias) avance rumo à convergência em políticas de IA responsável.

Portugal também tem sido pioneiro em diversos aspetos da governação de IA e das políticas públicas digitais. Em 2019, tornou-se um dos primeiros Estados-Membros da UE a publicar

um quadro ético oficial para a IA no setor público — o Guia de IA para uma Inteligência Artificial Ética, Transparente e Responsável na Administração Pública, desenvolvido pela Agência para a Modernização Administrativa (AMA). Este documento antecipou muitos dos princípios fundamentais do Regulamento, como a responsabilidade, a transparência, a explicabilidade, a equidade e a ética, tendo posicionado Portugal entre os primeiros países europeus a traduzir compromissos éticos abstratos em prática administrativa. Paralelamente, as Zonas Livres Tecnológicas (ZLTs), criadas pelo Decreto-Lei n.º 62/2021, estabeleceram um dos primeiros mecanismos legais da Europa para a realização de testes em contexto real de tecnologias emergentes, sob condições supervisionadas e juridicamente seguras. No âmbito da investigação e da coordenação, Portugal tem desempenhado um papel de liderança em projetos financiados pela UE, focados na gestão de riscos, no desenvolvimento de capacidades e na implementação responsável de sistemas de IA no setor público. O país também abriga uma das estruturas de cibersegurança mais avançadas do sul da Europa, através do Centro Nacional de Cibersegurança (CNCS), que já integrou as considerações sobre os riscos da IA nos seus programas de monitorização, treino e resiliência nacional. Essas iniciativas pioneiras demonstram que Portugal não apenas seguiu as tendências europeias, mas também ajudou a moldá-las, fornecendo uma base sólida para passar da liderança normativa à excelência operacional no âmbito da Lei de IA da UE.

Em todo o mundo, os governos estão a migrar de princípios abstratos para a governação

operacional de IA. Os Estados Unidos, através do *NIST AI Risk Management Framework* e do recém-criado Instituto de Segurança de IA, começaram a institucionalizar testes de ataque a sistemas (do inglês “red teaming”), exigindo que tanto desenvolvedores privados quanto agências públicas realizem avaliações estruturadas de robustez, antes da implementação de um sistema. O Reino Unido seguiu um caminho semelhante: o Instituto de Segurança de IA e a *Red Team Network* interdepartamental realizam avaliações contínuas de modelos em larga escala, para identificar vieses, segurança e uso indevido, com os resultados obtidos a serem incorporados diretamente nos módulos de aquisição e certificação. França, Alemanha e Holanda estão a desenvolver estruturas comparáveis dentro das suas agências digitais nacionais, integrando células de teste de IA em estruturas de segurança cibernética existentes. Singapura e Canadá foram mais além, incorporando mesmo ferramentas de garantia de IA em contratos de compras públicas, tornando os fornecedores legalmente responsáveis por demonstrar conformidade através de documentação de rastreabilidade, robustez e impacto ético.

Essas experiências demonstram uma convergência em direção a três mecanismos complementares que podem ser desenvolvidos em Portugal, tendo em vista a aplicação plena e eficaz do IA Act .

Em primeiro lugar, os testes adversariais, ou “red teaming”, estão a tornar-se um padrão de salvaguarda internacional. Estes testes combinam *hacking* ético e ataques aos sistemas, para identificar vulnerabilidades ocultas, vieses e riscos de manipulação. Nos EUA e no Reino Unido, esses exercícios são conduzidos em conjunto por equipas de IA e cibersegurança, sob supervisão coordenada de agências nacionais. Portugal pode replicar esse modelo, estabelecendo um programa conjunto entre a ANACOM e o Centro Nacional de Cibersegurança (CNCS), integrando processos de “red teaming” na governação de sistemas de IA de risco elevado e de uso geral na Administração Pública. Os resultados desses exercícios devem alimentar diretamente as obrigações de gestão de riscos

e monitorização pós-mercado, previstas na Lei de IA e na Diretiva NIS2, transformando a conformidade em resiliência proativa.

Em segundo lugar, a maioria das jurisdições de países líderes está a criar redes permanentes de especialistas para manter a capacidade de dar resposta ágil e especializada. O *Comité de l’Intelligence Artificielle de Confiance* da França, o *AI Advisory Council* do Canadá e o *Model AI Governance Framework Working Group* de Singapura reúnem especialistas da administração pública, da academia e do setor privado, de modo a fornecerem orientação técnica a reguladores e implementadores. Portugal pode adotar um modelo semelhante, formando uma rede transversal de especialistas nacionais sob a coordenação da ANACOM e do CNCS, fornecendo suporte ágil às autoridades competentes e entidades públicas. Essa rede deve emitir orientações práticas, organizar workshops técnicos e manter um repositório dinâmico de boas práticas acessível a todos os níveis de governo.

Em terceiro lugar, as melhores práticas internacionais apontam para a necessidade de documentação política harmonizada e operacional. O *Algorithmic Transparency Recording Standard* do Reino Unido, a *Directive on Automated Decision-Making* do Canadá e o conjunto de ferramentas *AI Verify* de Singapura traduzem princípios éticos amplos em procedimentos administrativos diários, modelos e listas de verificação. Portugal pode seguir este caminho, revendo documentos existentes, como o Guia de Ação para uma Inteligência Artificial Ética, Transparente e Responsável da AMA, e integrando-os num único manual de conformidade. Isto deverá clarificar as responsabilidades entre a ANACOM, a CNCS e os vários reguladores setoriais identificados no artigo 77.º do IA Act, ao mesmo tempo que fornece cláusulas modelo para aquisições, conjuntos de dados e procedimentos de supervisão humana.

Ao integrar essas práticas internacionais comprovadas — como os testes adversariais (“red teaming”), redes de especialistas e conjuntos de ferramentas políticas harmonizadas como instrumentos de clareza

jurídica —, Portugal pode acelerar a implementação da Lei de Inteligência Artificial e posicionar-se entre as principais jurisdições europeias em governação responsável de IA. Essa abordagem transforma a conformidade

em confiança: garantindo que todos os sistemas de IA implementados ou supervisionados pelo Estado sejam, não apenas legais, mas também confiáveis, explicáveis e seguros.

GT4 – Literacia em IA e capacitação avançada

O rápido desenvolvimento e a adoção de tecnologias de IA representam tanto oportunidades quanto desafios. Por um lado, a IA pode impulsionar o crescimento económico, fomentar a inovação e otimizar serviços em diversas áreas. Por outro lado, sem acesso equitativo às competências em IA, essa mudança tecnológica pode exacerbar as desigualdades existentes e expor as sociedades a uma gama mais ampla de riscos. Uma forma de abordar essas questões é promover a literacia em IA e o treino avançado.

A literacia em IA e o treino avançado têm significados diferentes para cada público-alvo. Por exemplo, para estudantes (do ensino básico, secundário, universidades, politécnicos e escolas profissionais), significa compreender os princípios fundamentais, os aspectos técnicos e as aplicações. Para políticos e tomadores de decisão, além dos princípios fundamentais, deve abranger as aplicações em diversos setores (por exemplo, administração pública, media, saúde, segurança, defesa, etc.), as regulamentações existentes, outros modelos legais em vigor e os riscos para indivíduos, populações e sistemas.

Uma vez que já foi realizado um trabalho significativo, tanto a nível internacional como nacional, por instituições públicas e privadas, seria sensato aproveitar esses esforços em vez de começar do zero, garantindo que todas as iniciativas caminham na mesma direção e contribuem coletivamente para a literacia em IA e a formação avançada no país.

Por fim, neste momento, não existe uma percepção qualitativa do estado da literacia em IA em Portugal, o que seria fundamental para compreender o sucesso (ou o fracasso) destes esforços de formação.

Ao lançar uma campanha nacional de literacia em IA, adaptada a diferentes grupos, é

essencial abordar tanto as vias educacionais formais quanto as informais, garantindo uma colaboração abrangente com públicos diversos. Algumas estratégias incluem:

i) Educação Formal: integrar a literacia em IA no currículo nacional em todos os níveis de ensino, aproveitando iniciativas já existentes, como as da Academia [ubbu](#) e da [DGE](#). Promover metodologias de aprendizagem ativa, incluindo a aprendizagem baseada em problemas, como as implementadas com sucesso, por exemplo, pelo [IEFP](#) e pela [ANQEP](#).

ii) Educação Informal: ampliar o conhecimento sobre IA para o público em geral, por meio de campanhas públicas direcionadas, utilizando meios de comunicação social, televisão e rádio, para destacar os impactos e as oportunidades da IA; colaborar com museus e bibliotecas, para criar materiais educativos acessíveis e compreensíveis sobre IA. Capacitar organizações não governamentais (ONGs) e organizações comunitárias para liderar iniciativas que visem aumentar o conhecimento público sobre IA, com foco em áreas rurais e sub-representadas, a fim de garantir o acesso equitativo aos recursos.

Para promover a iniciativa de literacia em IA, Portugal pode desenvolver parcerias entre agências governamentais, a indústria e ONGs, cujos esforços devem: incentivar o investimento do setor privado em programas de literacia em IA, especialmente por parte de empresas com forte presença em Portugal (por exemplo, nos setores de tecnologia, finanças e saúde); aumentar o financiamento da UE para a transformação digital a fim de apoiar projetos de literacia em IA; criar uma plataforma centralizada, na qual os indivíduos e as organizações possam aceder a recursos de

literacia em IA (como o [MLU-explAIIn](#)), incluindo ferramentas digitais, programas educativos (como o [pix](#)) e sessões de formação em português (como [MILObs](#), [LEME](#)); programas para promover o intercâmbio entre profissionais de diversas áreas da sociedade e especialistas em IA. Desta forma, a identificação de ferramentas tecnológicas para resolver desafios profissionais específicos seria mais facilmente promovida e facilitada (por exemplo, [Pólos de Inovação Digital](#)).

A partir dos indicadores do Índice de Economia e Sociedade Digitais (DESI) de 2024/2025 e do Relatório da Década Digital de 2024, observa-se um foco claro na utilização da IA no contexto empresarial. No entanto, existe uma lacuna significativa em termos de métricas de literacia

em IA. Assim, a realização de um inquérito nacional para avaliar o nível de literacia em IA em Portugal seria um passo fundamental. Tal esforço permitiria obter métricas de literacia em IA que abrangem os setores educativo, social e regional.

Embora um inquérito inicial pudesse ajudar a orientar os esforços para as recomendações aqui propostas, este deverá ser um esforço contínuo, de modo a avaliar a eficácia da implementação das políticas e da campanha de literacia em IA anteriormente recomendada. O trabalho desenvolvido pelo [INCoDe](#) ou pela AMA (agora denominada Agência para a Reforma Tecnológica do Estado) poderá servir de base para este esforço.

GT5 – Esforços regulatórios para além da Europa

A regulamentação da IA tornou-se uma prioridade global. Embora as abordagens variem significativamente entre jurisdições, a análise comparativa revela preocupações recorrentes: a necessidade de conciliar a inovação com a segurança, a responsabilização e o respeito pelos direitos fundamentais. Os marcos regulatórios são indispensáveis para a confiança pública, mas também podem aumentar os custos de conformidade e retardar a experimentação. A forma como cada sistema jurídico lida com essa tensão depende da sua cultura constitucional — da sua tolerância ao risco, do seu desenho institucional e da sua concepção de interesse público.

O Reino Unido, os Estados Unidos e o Japão seguem modelos descentralizados, setoriais e predominantemente baseados em normas não vinculativas, enquanto a União Europeia aplica uma abordagem horizontal e juridicamente vinculativa, fundamentada na avaliação da conformidade e na garantia dos direitos fundamentais. Para os decisores políticos europeus, o desafio reside em manter elevados padrões de proteção, criando simultaneamente condições que incentivem a liderança tecnológica e a inovação responsável.

De acordo com o [Relatório do Índice de IA de 2025 da Universidade de Stanford](#), os governos de todo o mundo estão a aumentar drasticamente tanto a regulamentação da IA, quanto o investimento público, mas a coordenação global permanece fragmentada. Portugal pode posicionar-se como uma referência no sul da Europa, para a implementação responsável e ética da IA, priorizando a transparência, a segurança e a colaboração transfronteiriça. Para alcançar esse objetivo, Portugal deve estabelecer estruturas de supervisão claras para as “sandboxes” ou ZLT, para apoiar ambientes de testagem de IA e incorporar salvaguardas éticas e de direitos humanos como fundamento da sua estrutura experimental.

As “sandboxes” ou ZLT não devem ser um espaço para desregulamentação, mas sim um ambiente estruturado em que a inovação é testada sob supervisão, com mecanismos integrados para mitigação de riscos, transparência e proteção do cidadão. Cada projeto participante deve demonstrar conformidade com os princípios de equidade, responsabilidade e transparência, suportados em planos documentados de gestão de riscos e de simulação de ataques cibernéticos. Um

comité de ética dedicado, integrando representantes do setor público, da academia e da sociedade civil, poderá supervisionar todas as atividades das ZLT, de forma a garantir que a experimentação esteja alinhada com as obrigações constitucionais e os direitos fundamentais.

Em paralelo, Portugal deve reforçar a cooperação internacional. A colaboração com Espanha continua a ser uma opção importante, mas parcerias equivalentes ou mais abrangentes poderiam também ser desenvolvidas com França, Canadá, Singapura e Japão, países que estabeleceram ambientes de teste de IA reconhecidos internacionalmente, integrando princípios éticos desde a fase de conceção. Portugal poderia, portanto, posicionar-se não só dentro de uma rede europeia de ambientes de teste interoperáveis, coordenada pelo European AI Office, mas também como uma porta de entrada do sul da Europa para estruturas globais de experimentação. Ao alinhar-se com iniciativas como o AI Verify de Singapura, o Directive on Automated Decision-Making sandbox do Canadá ou o Regulatory Sandbox Scheme do Japão, Portugal pode garantir que as lições aprendidas circulem também para a Europa e que as normas éticas e de direitos humanos sejam harmonizadas

internacionalmente. Este modelo aberto permitiria a Portugal implementar ambientes de teste conjuntos em setores prioritários como a saúde, a energia, a administração pública e a defesa e atrair investigação e investimento globais focados em IA responsável e centrada no ser humano.

Por fim, Portugal deve promover o desenvolvimento responsável de IA, facilitando o acesso a dados de treino lícitos e de alta qualidade, ao mesmo tempo que protege a propriedade intelectual e os direitos morais. Plataformas de licenciamento dedicadas ao treino de IA ajudariam a garantir a conformidade e a transparência, apoiando a investigação e o desenvolvimento sem prejudicar os direitos dos autores.

Em resumo, a liderança portuguesa deve concentrar-se em três pilares: (i) ambientes de teste éticos e baseados em direitos que tornem a inovação responsável a norma; (ii) coordenação internacional através de parcerias europeias de confiança; e (iii) acesso justo e transparente aos dados que respeite a criatividade e a legalidade. Esta abordagem alinharia o país com as melhores práticas globais, reforçaria a confiança pública e posicionaria Portugal como um centro credível e inovador para IA responsável na Europa.

Conclusões

*A equipa do BridgeAI considera que existem três recomendações transversais principais de todos os GTs, consultores e discussões públicas havidas: **continuidade** do trabalho em IA, **agilidade** para implementar as recomendações e adaptá-las aos contextos atuais e atrair e reter **talento** em IA.*

O BridgeAI testou a hipótese de que, ao criar espaços de aprendizagem multissetoriais adequados, apoiados por ferramentas tecnológicas que permitam agilizar a comunicação e extrair informações, é possível criar conhecimento acionável (Cross et al., 2004; Stern et al., 2021; Jagannathan et al., 2023) a partir de casos de estudo reais é

possível impulsionar a implementação eficaz dos processos de tomada de decisão. A equipa:

i) demonstrou que é possível alcançar um consenso alargado entre os atores fundamentais, permitindo a implementação mais eficiente das regulamentações;

ii) propôs resultados concretos, fundamentados em casos de estudo reais da indústria;

iii) realizou a prova de conceito da metodologia de conhecimento acionável, que pode ser aplicada a outras áreas temáticas relevantes para as políticas públicas;

iv) contribuiu para a implementação do EU AI Act em Portugal.

A implementação sustentável do EU AI Act exige um diálogo contínuo entre todas as partes relevantes, para garantir que a formulação de políticas e a sua implementação eficaz consigam acompanhar a evolução dinâmica neste campo tecnológico e regulamentar.

Como tal, e em consonância com a visão de impulsionar a economia nacional com base

numa profunda reforma e digitalização das estruturas e fluxos de trabalho relacionados com o Estado - considerada de grande relevância e alinhado com a criação do Ministério da Reforma do Estado, liderado pelo Ministro Matias - gostaríamos de deixar uma recomendação para a continuação deste tipo de iniciativas, incluindo aquelas com apoio financeiro no âmbito dos mecanismos de ciência para políticas públicas, que permitam o fluxo sustentável de conhecimento aplicável para a formulação de políticas públicas.

Bibliografia seletiva

- [1.] Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., Arx, S. von, Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N. S., Chen, A. S., Creel, K. A., Davis, J., Demszky, D., ... Liang, P. (2021). On the Opportunities and Risks of Foundation Models. ArXiv, abs/2108.07258.
- [2.] Cross, R., & Sproull, L. (2004). More Than an Answer: Information Relationships for Actionable Knowledge. *Organization Science*, 15(4), 446–462. <https://www.jstor.org/stable/30034748>
- [3.] Dignum, V. (2019). *Responsible artificial intelligence: how to develop and use AI in a responsible way*. Springer, Springer International Publishing: Cham, Switzerland, 2019; ISBN 978-3-030-30370-9.
- [4.] Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482>
- [5.] Floridi, L. (2019). Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical. *Philosophy & Technology*, 32(2), 185–193. <https://doi.org/10.1007/s13347-019-00354-x>
- [6.] *GulA: Guia para a Inteligência Artificial, versão 1.2 (2025). Agência para a Modernização Administrativa, IP.* <https://bo.mosaico.gov.pt/api/assets/mosaico-site/ff6944f0-8046-46ef-8ae0-36126b5f3b1f>
- [7.] Homsy, G. C., Liu, Z., & Warner, M. E. (2019). Multilevel governance: Framing the integration of top-down and bottom-up policymaking. *International Journal of Public Administration*, 42(7), 572–582.
- [8.] Jagannathan, K., et al. (2023, June). A research agenda for the science of actionable knowledge: Drawing from a review of the most misguided to the most enlightened claims in the science-policy interface literature. *Environmental Science & Policy*, 144, 174–186. <https://www.sciencedirect.com/science/article/pii/S1462901123000680>
- [9.] National Artificial Intelligence Initiative Office. The National AI Advisory Committee. Recommendations: Enhancing AI Literacy for the United States of America, report, November (2023). (<https://digital.library.unt.edu/ark:/67531/metadc2349497/m1/1/>: accessed October 30, 2025). University of North Texas Libraries, UNT Digital Library, <https://digital.library.unt.edu>; crediting UNT Libraries Government Documents Department.
- [10.] OECD. (2024). *Governing with Artificial Intelligence: Are Governments Ready?* OECD Artificial Intelligence Papers No. 20. <https://doi.org/10.1787/26324bc2-en>
- [11.] Pagallo, U., Aurucci, P., Casanovas, P., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Schafer, B., & Valcke, P. (2019). AI4People - On Good AI Governance: 14 Priority Actions, a S.M.A.R.T. Model of Governance, and a Regulatory Toolbox (SSRN Scholarly Paper 3486508). <https://papers.ssrn.com/abstract=3486508>
- [12.] Portugal 2025 Relatório por país sobre a Década Digital (2025). <https://digital-strategy.ec.europa.eu/pt/factpages/portugal-2025-digital-decade-country-report>
- [13.] Regulation (EU) 2024/1689 on Artificial Intelligence <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>
- [14.] Stern, M. J., Briske, D. D., & Meadow, A. M. (2021, February 19). Opening learning spaces to create actionable knowledge for conservation. *Conservation Science and Practice*, 3(5). <https://conbio.onlinelibrary.wiley.com/doi/10.1111/csp2.378>
- [15.] United Nations. (2024). *Governing AI for Humanity*. <https://doi.org/10.18356/9789211067873>



COMO CITAR ESTE DOCUMENTO

Oliveira, A., Martins, A., Valente, A., Brites, A., Novais, A., Ceña, B. et al. (2026). *BridgeAI: Boosting Regulatory Implementation with Data-driven insights, Global expertise, and Ethics for AI*, S4P-23 Policy Brief 5644/2023. PLANAPP – Centro de Planeamento e de Avaliação de Políticas Públicas.




CONTACTO

science4policy@planapp.gov.pt

COPYRIGHT

© PLANAPP, 2026

-  [Ciência para as políticas públicas](#)
-  [PLANAPP](#)
-  [Newsletter](#)

-  [PLANAPP](#)
-  [@planapp](#)
-  [PLANAPP podcasts](#)



Este *policy brief* foi desenvolvido no âmbito do Science4Policy 2024 (S4P-24): Concurso de Estudos de Ciência para as Políticas Públicas, uma iniciativa do Centro de Planeamento e de Avaliação de Políticas Públicas (PLANAPP), em parceria com a Fundação para a Ciência e a Tecnologia (FCT), financiada pelo Plano de Recuperação e Resiliência de Portugal. Linha temática S4P-24/09: Digitalização, inovação e qualificação / Antecipar a regulação da Inteligência Artificial.

O conteúdo é da exclusiva responsabilidade dos seus autores e não vincula nem compromete o PLANAPP nem a FCT.