



Regulação da Inteligência Artificial

benchmarking de países selecionados

Dezembro, 2022

EVE
Evidência Express

ENAP

Expediente

Presidente

Diogo Costa

Diretora-Executiva

Rebeca Loureiro de Brito

Diretora de Altos Estudos

Flávia de Holanda Schmidt

Diretor de Educação Executiva

Rodrigo Torres

Diretor de Desenvolvimento Profissional

Paulo Marques

Diretora de Inovação

Camila de Castro Barbosa Medeiros

Diretora de Gestão Interna

Alana Regina Biagi Silva Lisboa

Coordenador Geral de Ciência de Dados

Pedro Masson Sesconetto Souza

Capa e Diagramação

Samyra Lima

Imagens

Unsplash

Autoria

Ana Karolina Acris Melo

Doutoranda em Economia (UFPB).

Géssica C. Souza

Doutora em Economia (UFMG).

Amanda Corrêa Vasco

Mestre em Políticas Públicas (Insper).

Breno Salomon Reis

Mestre em Políticas Públicas (Insper).

O Evidência Express (EvEx) é uma iniciativa da Diretoria de Altos Estudos da Escola Nacional de Administração Pública (Enap) em parceria com a Universidade Federal de Brasília (UnB). A missão do EvEx é melhorar a tomada de decisão do setor público. Para isso a equipe sintetiza, produz e dissemina evidências que possam servir de base para o desenho, monitoramento e avaliação de políticas públicas.

Avaliações completas de políticas públicas são intensivas em tempo e custos. A fim de agilizar esses processos, o EvEx produz relatórios ágeis de evidências para a consolidação do conhecimento disponível e introdução de novos pontos de vista.

Os resultados dos produtos EvEx apoiam tomadores de decisão do setor público federal, subsidiando avaliações Ex Ante, Ex Post ou Análises de Impacto Regulatório. Beneficiam também os gestores públicos subnacionais, pesquisadores, docentes, servidores e demais interessados na sociedade civil.

Os produtos EvEx analisam evidências qualitativas e quantitativas, podendo ser demandados de forma avulsa ou em pacotes, sobre:

- Evolução do problema no Brasil e no mundo;
- Público-alvo de uma política;
- Causas e consequências do problema ou política;
- Soluções existentes para o problema;
- Impactos de intervenções ou políticas públicas.

Para mais informações, consulte nossa página (www.enap.gov.br/pt/servicos/avaliacao-e-organizacao-de-evidencias) ou entre em contato: evidencia.express@enap.gov.br.



Sumário Executivo

- Este *benchmarking* descreve políticas públicas internacionais, já implementadas ou em desenvolvimento, para regular o uso da Inteligência Artificial (IA). O trabalho foi elaborado para auxiliar as atividades que ocorrem no âmbito da Estratégia Brasileira de Inteligência Artificial (EBIA).
- São objetos de estudo os casos da União Europeia, Reino Unido, Estados Unidos, Austrália e Japão.
- Na União Europeia, a Lei de Inteligência Artificial se encontra em processo final de implementação e propõe a regulação de uma ampla gama de aplicações de IA por meio de uma abordagem baseada em risco para a segurança do usuário, e sanções em caso de infringimento de regras.
- No Reino Unido, o *National AI Strategy - AI Action Plan* propõe regular a IA com base em seu uso e no impacto que ela tem sobre indivíduos, grupos e empresas, com abordagem direcionada à inovação e desenvolvimento econômico.
- No Reino Unido, a IA é regulada por meio de legislação "intersectorial", com diferentes reguladores, cujo a proposta é estabelecer um conjunto de princípios intersectoriais adaptados a distintas características.
- A regulação de IA nos Estados Unidos é feita por agências reguladoras e pelos estados da federação.
- Nos Estados Unidos há ações específicas de coordenação por parte do governo para posicionar o país como liderança em pesquisa e desenvolvimento de IA, nos âmbitos público e privado. Portanto, há prioridade para o fortalecimento de pesquisa e desenvolvimento da competitividade do país em IA.
- Na Austrália, ainda não há leis específicas implementadas sobre o uso da IA. Todavia, o governo australiano está desenvolvendo o arcabouço legal e a governança necessárias para posicionar a Austrália como um líder global em IA.
- O Japão foi o segundo país a desenvolver estratégias nacionais de IA e a estabelecer metas e alocar orçamento para o tema.
- As agências governamentais japonesas adotam a abordagem de *soft-law* para lidar com possíveis vieses da tecnologia de IA. O país busca criar uma governança ágil, com o objetivo de não prejudicar investimentos e não impedir a inovação.

Sumário

1	Introdução	5
2	Metodologia	7
3	União Europeia	8
3.1	Instituições para regulação e desenvolvimento da IA	8
3.2	Mitigação de riscos e vieses da IA	8
4	Reino Unido	13
4.1	Instituições para regulação e desenvolvimento da IA	13
4.2	Mitigação de riscos e vieses da IA	14
5	Estados Unidos	16
5.1	Instituições para regulação e desenvolvimento da IA	16
5.2	Mitigação de riscos e vieses da IA	17
6	Austrália	20
6.1	Instituições para regulação e desenvolvimento da IA	20
6.2	Mitigação de riscos e vieses da IA	21
7	Japão	24
7.1	Instituições para regulação e desenvolvimento da IA	24
7.2	Mitigação de riscos e vieses da IA	26
8	Considerações Finais	29
	Referências Bibliográficas	30
	Apêndice 1	33

1. Introdução

O presente relatório apresenta os resultados de um estudo descritivo e exploratório sobre a regulação da Inteligência Artificial em múltiplos países. Ele foi desenvolvido para auxiliar as atividades que ocorrem no âmbito da Estratégia Brasileira de Inteligência Artificial (EBIA), a qual tem por objetivo nortear o Estado brasileiro em prol do desenvolvimento de ações que estimulem a pesquisa, inovação e desenvolvimento de soluções em IA, bem como seu uso consciente, responsável e ético¹.

De acordo com a Organização para a Cooperação e Desenvolvimento Econômico (OCDE), a Inteligência Artificial (IA) é uma tecnologia implantada em diversos setores, com um propósito geral e com o potencial de melhorar o bem-estar, contribuir para uma atividade econômica global sustentável positiva, aumentar a inovação e a produtividade e auxiliar nos atuais desafios globais (OCDE, 2019).

A IA é capaz de oferecer um vasto conjunto de benefícios econômicos e sociais a todo o leque de indústrias e atividades sociais ao melhorar as previsões, otimizar operações, a repartição de recursos e personalizar soluções digitais disponibilizadas às pessoas e às organizações. A utilização da IA também pode conferir importantes vantagens competitivas às empresas e contribuir para progressos sociais e ambientais. Ela pode ser utilizada, por exemplo, nos cuidados de saúde, na agricultura, na educação, no transporte e logística, na segurança, na eficiência energética, dentre muitas outras possibilidades (COMMISSION, 2021b). Diante dos impactos potenciais da IA, governos têm buscado conceber estratégias nacionais de fomento ao seu desenvolvimento. Esse movimento está sendo percebido como uma corrida internacional pela IA (IPEA, 2022).

Apesar dos benefícios da IA, sua aplicação em certos contextos pode criar riscos para usuários e desenvolvedores. Como gerenciar e preparar a estrutura regulatória para coibir tais práticas é um desafio sendo enfrentado por governos em todo o mundo. Há uma preocupação que uma eventual regulação excessiva possa inibir a inovação por parte das empresas, mas, ao mesmo tempo deixe de prevenir e mitigar danos que possam ser causados pelo uso indevido da IA. Assim, os países estão experimentando diferentes propostas que levam em conta essas duas dimensões, além de outras, como competitividade.

Este relatório apresenta o resultado de uma pesquisa que visa explorar melhor como essas dimensões são tratadas pelos países. Em particular, são foco da investigação a estrutura regulatória

¹Publicada em abril de 2021, a EBIA marca a inclusão do Brasil na lista de países com estratégias nacionais para desenvolvimento da IA. Ela elenca 73 ações estratégicas – divididas em 9 eixos temáticos – como prioritárias para realizar uma visão de futuro fundamentada nos princípios de “(i) crescimento inclusivo, desenvolvimento sustentável e bem-estar; (ii) valores centrados no ser humano e na equidade; (iii) transparência e explicabilidade; (iv) robustez, segurança e proteção e; (v) responsabilização ou *accountability*” (MCTI, 2022). Esses mesmos princípios guiam as ações dos países-membros da OCDE no tema.

existente em cada caso analisado e os mecanismos adotados por estes para assegurar uma IA responsável, que mitigue riscos e vieses na sua aplicação. São objeto de estudo os casos da União Europeia, Reino Unido, Estados Unidos, Austrália e Japão.

O relatório está dividido em oito seções, incluindo esta introdução. Na próxima seção detalha-se a metodologia adotada. Nas seções subsequentes são descritos os resultados da investigação dos casos da União Europeia, Reino Unido, Estados Unidos, Austrália e Japão, respectivamente. Por fim, são tecidas considerações finais que comparam os achados encontrados para cada caso.

Salientamos que esta pesquisa foi elaborada de maneira independente pelo Evidência Express (EvEx) nos meses de outubro e novembro de 2022, e seus resultados não representam recomendações de políticas públicas para a regulação de Inteligência Artificial. Nenhum caso apresentado de estrutura regulatória se constitui necessariamente em um modelo a ser seguido pelo Brasil. Os produtos do EvEx são elaborados como respostas ágeis, oferecendo resultados preliminares, e não devem ser tratados como evidências definitivas sobre o fenômeno estudado.

2. Metodologia

Neste trabalho foi realizado um estudo descritivo de múltiplos casos sobre a regulação da Inteligência Artificial em países selecionados. Em cada caso foram analisadas o processo de implementação das políticas públicas para a regulação da IA, além de destacar a estrutura regulatória do uso da IA e os mecanismos de mitigação de risco e vieses adotados pelos países.

Para executar a pesquisa, foi adotado o método de estudo de caso exploratório. O estudo de caso exploratório consiste em uma investigação detalhada, cujo “objetivo consiste em fazer uma análise do contexto e processos que iluminam as questões estudadas” (CASSELL; SYMON, 2004). O método é caracterizado por assumir que múltiplas perspectivas e interpretações são valiosas para caracterizar o objeto do estudo. Dessa forma, os estudos de caso são primeiro expansivos antes de serem conclusivos. Apesar disso, os bons estudos de caso têm desenho bem enquadrados (não é toda informação que importa para explicar um fenômeno ou processo). A delimitação permite a indução do raciocínio lógico e a construção de um foco progressivo que permite chegar à conclusões críveis.

Foram realizadas buscas em páginas de sites oficiais de governos, documentos oficiais de governo, relatórios publicados por organismos multilaterais, publicações acadêmicas e relatórios técnicos de organizações privadas que atuam ou estudam o tema da Inteligência Artificial.

3. União Europeia

3.1 Instituições para regulação e desenvolvimento da IA

As soluções baseadas em IA são cada vez mais comuns e para tentar acompanhar e orientar este desenvolvimento estão surgindo iniciativas políticas em todo o mundo. Na União Europeia, a Comissão Europeia (CE) criou a Comunicação da Comissão Europeia sobre Inteligência Artificial (COMMISSION, 2018a; COMMISSION, 2018b), em 2018, iniciando o desenvolvimento de estratégias independentes pelos Estados-Membros e formalizando O Plano Coordenado para a Inteligência Artificial (COMMISSION, 2018b).

O plano descreve uma abordagem coordenada para maximizar os benefícios e enfrentar os desafios trazidos pela IA, detalhando sete objetivos, que incluem o financiamento de startups, o investimento de €1,5 bilhões em vários centros de excelência em pesquisa, o apoio a mestrados e doutorados em IA, a criação de espaços de dados europeus comuns e o desenvolvimento de diretrizes de ética com uma perspectiva global, dentre outros objetivos de desenvolvimento.

3.2 Mitigação de riscos e vieses da IA

A Comissão nomeou um grupo independente de especialistas de alto nível para desenvolver suas diretrizes de ética, publicadas em sua forma final em abril de 2019 (INTELLIGENCE, 2019a). As diretrizes listam os principais requisitos que os sistemas de IA devem atender para serem confiáveis, como a privacidade e governança de dados, a transparência e o bem-estar social e ambiental. Além disso, são listadas uma série de recomendações importantes sobre a proteção das pessoas, aumentando a aceitação de IA no setor privado, expandindo a capacidade europeia de pesquisa em IA e desenvolvendo práticas éticas de gestão de dados (INTELLIGENCE, 2019b). Essas recomendações incluem a garantia de que os dados dos usuários coletados não sejam usados para discriminá-los ilegal ou injustamente e a garantia de privacidade e proteção de dados durante todo o ciclo de vida do sistema de IA.

Em 2020, a CE publicou um *white paper* (COMMISSION, 2020) sobre IA informando os requisitos de documentação que permitiriam às autoridades garantir que os sistemas de IA de alto risco cumprissem direitos fundamentais. O documento também menciona um “rótulo de qualidade” voluntário para aplicações de IA consideradas de baixo risco e considera a exigência de uma avaliação de conformidade pré-comercialização para aplicativos de IA de “alto risco”, como o reconhecimento facial. Esse seria o elemento central de uma estrutura regulatória potencial para IA, que viria a ser formalizada em abril de 2021 com a Lei de Inteligência Artificial, mesmo ano em que o Plano Coordenado para a Inteligência Artificial foi atualizado com novas estratégias de crescimento do setor (COMMISSION, 2021a).

A proposta da Lei de Inteligência Artificial (IA) (COMMISSION, 2021b) faz parte de um pacote

abrangente de medidas que visa resolver os problemas decorrentes do desenvolvimento e da utilização da IA, examinados no *white paper*. Além disso, visa padronizar as estratégias regulatórias entre os Estados-Membros visto que a supervisão regulatória autônoma colocaria em risco o ecossistema sinérgico do mercado interno da UE. A lei propõe a regulação de uma ampla gama de aplicações de IA, alinhando-as com os valores e direitos fundamentais da UE por meio de uma abordagem baseada em risco. O âmbito, os instrumentos e o quadro de governança introduzidos pela proposta ainda estão sendo debatidos e aperfeiçoados pelos legisladores europeus. A proposta se tornará lei quando o Conselho (representando os 27 Estados-Membros da UE) e o Parlamento Europeu concordarem com uma versão comum do texto, o que está previsto para o início de 2023.

A proposta da Lei de Inteligência Artificial (IA) (COMMISSION, 2021b) descreve os vários níveis de riscos e as sanções aplicadas em caso de infringimento das regras. Os diferentes tipos de risco são: i) risco inaceitável, ii) risco elevado, iii) risco baixo ou mínimo.

Os riscos inaceitáveis representam algumas práticas de IA particularmente prejudiciais e que são proibidas, uma vez que violam os valores da UE. As proibições abrangem práticas com potencial significativo para manipular as pessoas por meio de técnicas subliminares que lhes passam despercebidas ou explorar as vulnerabilidades de grupos específicos, como as crianças ou as pessoas com deficiência. Também abrangem práticas que possam distorcer substancialmente o seu comportamento de uma forma que as torne suscetíveis a danos psicológicos ou físicos (COMMISSION, 2021b). A proposta também proíbe a classificação social assente na IA para uso geral por parte das autoridades públicas. Por último, é igualmente proibida a utilização de sistemas de identificação biométrica à distância, em tempo real, em espaços acessíveis ao público para efeitos de manutenção da ordem pública, a não ser que se apliquem determinadas exceções limitadas (COMMISSION, 2021b).

Em relação aos sistemas de IA de risco elevado, sua classificação não depende somente da função do sistema, mas também da finalidade específica e das modalidades para as quais aquele sistema é utilizado. São considerados de alto risco, por exemplo, infraestruturas críticas (como transportes), que podem colocar em risco a vida e a saúde dos cidadãos; a formação escolar que pode determinar o acesso à educação e o percurso profissional da vida de uma pessoa (ex. classificação em exames); os componentes de segurança de produtos (ex. aplicação de IA em cirurgia assistida por robô); gestão de trabalhadores (como software de triagem de currículos para processos de recrutamento); dentre outros sistemas. A lista inclui um número de sistemas de IA cujos riscos já se materializaram ou são suscetíveis de se materializar num futuro próximo. A fim de garantir que o regulamento possa ser ajustado a novas utilizações e aplicações de IA, a Comissão pode alterar a lista periodicamente (COMMISSION, 2021b).

Esses sistemas de IA de alto risco precisam cumprir um conjunto de requisitos obrigatórios horizontais e seguir procedimentos de avaliação da conformidade antes de poderem ser colocados no mercado da UE. O capítulo 3 da Lei de Inteligência Artificial (COMMISSION, 2021b) indica um conjunto evidente de obrigações horizontais impostas aos fornecedores de sistemas de IA de risco

elevado. Também são impostas obrigações proporcionadas aos utilizadores e a outros participantes de toda a cadeia de valor da IA (por exemplo, importadores, distribuidores, mandatários). Por exemplo, os fornecedores de sistemas de IA de risco elevado devem:

- Assegurar que os sistemas de IA de risco elevado cumprem os requisitos estabelecidos na lei;
- Dispor de um sistema de gestão da qualidade que cumpra o disposto proposto;
- Elaborar a documentação técnica do sistema de IA de risco elevado;
- Quando tal esteja sob o seu controle, manter os registos gerados automaticamente pelos sistemas de IA de risco elevado que fornecem;
- Assegurar que o sistema de IA de risco elevado seja sujeito ao procedimento de avaliação da conformidade aplicável, antes da colocação no mercado ou da colocação em serviço;
- Respeitar as obrigações de registo proposto;
- Adotar as medidas corretivas necessárias, se o sistema de IA de risco elevado não estiver em conformidade com os requisitos estabelecidos;
- Informar as autoridades nacionais competentes dos Estados-Membros nos quais disponibilizaram o sistema de IA ou o colocaram em serviço e, se for o caso, o organismo notificado sobre a não conformidade e quaisquer medidas corretivas tomadas;
- Apor a Marcação de conformidade CE¹ nos sistemas de IA de risco elevado para indicar a conformidade com o presente regulamento;
- Mediante pedido de uma autoridade nacional competente, demonstrar a conformidade do sistema de IA de risco elevado com os requisitos estabelecidos.

A fim de assegurar um nível elevado e coerente de proteção dos interesses públicos nos domínios da saúde, da segurança e dos direitos fundamentais, devem ser criadas normas comuns aplicáveis a todos os sistemas de IA de risco elevado. Essas normas devem ser coerentes com a Carta dos Direitos Fundamentais da União Europeia e não discriminatórias, bem como estar em consonância com os compromissos comerciais internacionais da União (COMMISSION, 2021b).

Os fornecedores de sistemas de IA de risco elevado devem criar um sistema de gestão da qualidade, o qual deve listar, de forma sistemática, procedimentos e instruções para uso da IA. A documentação do sistema deve abordar aspectos como: técnicas, procedimentos e ações sistemáticas a serem utilizadas para a concessão, controle da concessão e verificação da concessão do sistema de IA de risco elevado; exames e validações a serem realizados antes, durante e após o desenvolvimento do sistema de IA de risco elevado; a frequência com a qual os exames e validações devem ser realizados; dentre outros (COMMISSION, 2021b). O projeto da Lei de Inteligência Artificial também incentiva os países europeus a adotar *sandboxes*² regulatórios de IA para facilitar o desenvolvimento de sistemas

¹Marcação pela qual um fornecedor atesta que um sistema de IA está em conformidade com os requisitos estabelecidos no título III, capítulo 2, da Lei de Inteligência Artificial (COMMISSION, 2021b) e no restante da legislação da União aplicável que harmoniza as condições de comercialização de produtos em que seja prevista a respetiva aposição

²Um *sandbox* é um ambiente de teste isolado que permite aos usuários executar programas ou abrir arquivos sem afetar

inovadores de IA e a realização de testes sob supervisão regulatória (COMMISSION, 2021b).

No caso de sistemas de IA de risco baixo ou mínimo, apenas são propostas obrigações de transparência mínimas, permitindo que os indivíduos que interagem com o conteúdo tomem decisões informadas e tenham a opção de continuar ou parar de usar a aplicação. São exemplos desses sistemas os *chatbots*, aplicativos como videogames habilitados para IA ou filtros de *spam* (COMMISSION, 2021b).

As regras propostas para cada grau de risco seriam executadas por intermédio de um sistema de governança a nível dos Estados-Membros, aproveitando estruturas já existentes, e de um mecanismo de cooperação ao nível da UE. Cada Estado-Membro deverá designar, ou criar, uma autoridade notificadora responsável por estabelecer e executar os procedimentos necessários para a avaliação, a designação e a notificação de organismos de avaliação e fiscalização da conformidade³. Essas autoridades nomeadas representariam sua nação no Conselho Europeu de Inteligência Artificial (COMMISSION, 2021b).

Se algum dos riscos e regras estipulados for infringido, seriam aplicadas sanções, divididas em três faixas:

- Até 30 milhões de euros ou 6% do volume de negócios anual global do ano financeiro anterior (o que for maior) por infrações de práticas proibidas ou não conformidade relacionada a requisitos de dados;
- Até 20 milhões de euros ou 4% do volume de negócios anual global do exercício anterior por incumprimento de qualquer um dos outros requisitos ou obrigações do Regulamento;
- Até 10 milhões de euros ou 2% do volume de negócios anual global do exercício anterior para o fornecimento de informações incorretas, incompletas ou enganosas a organismos notificados e autoridades nacionais competentes em resposta a um pedido.

Por fim, identifica-se na Lei de Inteligência Artificial (COMMISSION, 2021b) o principal instrumento de regulação da UE. A natureza horizontal da proposta procura ser coerente com as legislações em vigor dos setores em que os sistemas de IA já são utilizados ou serão provavelmente utilizados num futuro próximo. As propostas regulatórias e de conformidade devem atender as políticas de governança interna já existentes em alguns setores como o de saúde e financeiro. No caso de alguns produtos específicos, como máquinas, dispositivos médicos e brinquedos, as avaliações de conformidade deverão seguir os previstos na legislação do Novo Quadro Legislativo (NQL).

De acordo com o documento da Lei de Inteligência Artificial (COMMISSION, 2021b), a

o aplicativo, sistema ou plataforma em que são executados. Os desenvolvedores de software usam *sandboxes* para testar novos códigos de programação. Profissionais de segurança cibernética usam *sandboxes* para testar softwares potencialmente maliciosos.

³A fim de conduzir a avaliação de conformidade de sistemas de IA de acordo com a Lei de Inteligência Artificial, pesquisadores da Universidade de Oxford criaram uma ferramenta chamada capAI (FLORIDI et al., 2022). A CapAI fornece às organizações orientações práticas sobre como traduzir princípios éticos de alto nível em critérios verificáveis que ajudam a moldar o design, desenvolvimento, implantação e uso da IA ética. Essa ferramenta pode ser usada para demonstrar que o desenvolvimento e a operação de um sistema de IA são confiáveis. A ferramenta está sendo validada com as empresas no momento

escolha de um regulamento como instrumento jurídico justifica-se pela necessidade de aplicar uniformemente as novas regras, como a definição de inteligência artificial, a proibição de determinadas práticas prejudiciais possibilitadas pela IA e a classificação de determinados sistemas de IA. Para isso, o documento introduz um conjunto harmonizado de requisitos básicos no que diz respeito aos sistemas de IA classificados como de risco elevado, bem como obrigações aplicáveis aos fornecedores e aos utilizadores desses sistemas. Isso é feito na expectativa de melhorar a proteção dos direitos fundamentais e proporcionar segurança jurídica para os operadores e os consumidores.

4. Reino Unido

4.1 Instituições para regulação e desenvolvimento da IA

Os países estão em diferentes estágios de implementação de suas estratégias nacionais de IA. O Reino Unido foi um dos primeiros a desenvolver estratégias nacionais de IA, estabelecendo metas, orientações e alocando orçamento, por volta de 2017. Mas somente em julho de 2020, o *Information Commissioner's Office* (ICO) publicou Orientações sobre Inteligência Artificial e Proteção de Dados (ICO, 2020), para ajudar as organizações a mitigar os riscos de IA decorrentes de uma perspectiva de proteção de dados.

As orientações estabelecem uma estrutura e uma metodologia para auditar sistemas de IA. Elas incluem melhores práticas em conformidade com a Lei de Proteção de Dados do Reino Unido de 2018 e com as obrigações de proteção de dados sob o Regulamento Geral de Proteção de Dados da UE (*General Data Protection Regulation* - GDPR). A metodologia de auditoria é amparada na “abordagem proporcional e baseada no risco” e apresenta: i) ferramentas e procedimentos para auditorias e investigações; ii) orientação detalhada sobre IA e proteção de dados, e iii) um kit de ferramentas que fornece suporte prático para organizações que auditam a conformidade de seus próprios sistemas de IA. Além disso, as orientações abordam: i) responsabilidade e governança em IA; ii) tratamento justo, lícito e transparente; iii) minimização e segurança de dados; e iv) conformidade com os direitos individuais dos titulares de dados (ICO, 2020).

Não são considerados nas orientações os impactos da IA em áreas além da proteção de dados, que não são de competência da ICO, como saúde e financeiro. Todavia, em julho de 2022, foi lançado o *National AI Strategy - AI Action Plan* (INTELLIGENCE, 2021), cujo um dos pilares é garantir que o Reino Unido obtenha a governança nacional e internacional das tecnologias de IA para incentivar a inovação, com investimento e proteção ao público e aos valores fundamentais. Por meio de uma abordagem pró-inovação, o plano propõe a criação do Plano de Regulação Digital, cujos princípios são:

- Regular a IA com base em seu uso e no impacto que ela tem sobre indivíduos, grupos e empresas em um contexto específico e delegar aos reguladores a responsabilidade de projetar e implementar respostas regulatórias proporcionais. Isso garantirá que a abordagem seja direcionada e apoie a inovação.
- Focar na abordagem de questões em que haja evidências claras de risco real ou oportunidades perdidas. Os reguladores devem se concentrar em preocupações de alto risco, em vez de riscos hipotéticos ou baixos associados à IA. Com isso, pretende-se incentivar a inovação e evitar barreiras desnecessárias em seu caminho.
- Estabelecer um conjunto de princípios intersetoriais adaptados às características distintas da

IA, com reguladores solicitados a interpretar, priorizar e implementar esses princípios em seus respectivos setores de domínio. A fim de haver coerência e apoiar a inovação por parte dos regulados, devem ser exploradas maneiras de apoiar e incentivar a coordenação regulatória - por exemplo, trabalhando em estreita colaboração com o *Digital Regulation Cooperation Forum* (DRCF) e outros reguladores e partes interessadas. Com isso se espera que se torne a estrutura regulatória mais fácil de navegar para os regulados.

- Os reguladores devem considerar opções de intervenções mais leve, como orientação ou medidas voluntárias, em primeira instância, para que a abordagem permaneça adaptável. Na medida do possível, também será priorizado trabalhar com processos existentes, em vez da criação de novos processos.

Um dos objetivos de médio prazo do *National AI Strategy - AI Action Plan* (INTELLIGENCE, 2021) é a criação do *AI Standards Hub*¹, projeto liderado pelo *The Alan Turing Institute*. Ele almeja aumentar o envolvimento do Reino Unido no desenvolvimento de padrões técnicos globais para IA. O Hub criará ferramentas práticas e reunirá a comunidade multissetorial de IA do Reino Unido para garantir que os padrões globais da *Standards Development Organisations* (SDOs)² de IA sejam moldados por uma ampla gama de especialistas, para fornecer as ferramentas necessárias para a governança de IA.

4.2 Mitigação de riscos e vieses da IA

Não existem leis do Reino Unido explicitamente escritas para regular a IA. O setor é parcialmente regulamentado por meio de uma "colcha de retalhos" de requisitos legais e regulatórios criados para outros fins que agora também capturam usos de tecnologias de IA (INTELLIGENCE, 2022). Dessa forma, a IA é regulada por meio de uma legislação que acaba sendo intersetorial por advir de diferentes reguladores. Por exemplo, há cobertura de temas como proteção de dados (*Information Commissioner's Office*), concorrência (*Competition Markets Authority*), direitos humanos e igualdade (*Equality Human Rights Commission*). Também há legislação específicas de setor, como serviços financeiros (*Financial Conduct Authority*) e produtos médicos (*Medicines and Healthcare products Regulatory Agency*) (INTELLIGENCE, 2021).

Há legislação específica sobre dados que possui interseção com IA. A lei de proteção de dados, por exemplo, inclui requisitos específicos sobre "tomada de decisão automatizada" e o processamento mais amplo de dados pessoais, que também abrange o processamento para fins de desenvolvimento e treinamento em tecnologias de IA. O próximo Projeto de Lei de Segurança Online também contém provisões especificamente relacionadas ao projeto e uso de algoritmos.

A *Equality Human Rights Commission* identificou a IA como uma prioridade estratégica

¹Em janeiro de 2022, foi anunciado o piloto do website que promete a disponibilização de banco de dados padrões, bem como bancos de dados complementares de políticas e pesquisas, materiais de treinamento on-line e fóruns da comunidade. A plataforma pode ser acessada pelo link: <<https://aistandardshub.org/>>

²SDO é uma organização focada no desenvolvimento, publicação ou divulgação de padrões técnicos para atender às necessidades de uma indústria ou campo.

em seu Plano Estratégico 2022-2025 e se comprometeu a fornecer orientação sobre como a Lei da Igualdade se aplica ao uso de novas tecnologias, como IA, na tomada de decisões automatizada. O *Health and Safety Executive* se comprometeu a desenvolver pesquisas colaborativas com a indústria e a academia em seu Plano de Fornecimento de Ciência e Evidência 2020-2023, para determinar uma compreensão clara das implicações de saúde e segurança da IA no local de trabalho (INTELLIGENCE, 2021).

A *National AI Strategy - AI Action Plan* (INTELLIGENCE, 2021) mantém essa abordagem regulatória baseada no setor existente. De acordo com o plano, essa estratégia garante que os reguladores individuais tenham poderes para trabalhar com flexibilidade dentro de suas próprias atribuições para garantir que a IA forneça os resultados certos. O plano é introduzir princípios e regras intersetoriais adicionais, específicos para IA, para complementar o papel dos reguladores individuais, permitindo mais consistência entre os regimes existentes. A proposta de coordenação da regulação inclui os atores reguladores nacionais como a FCA (*Financial Conduct Authority*), CMA (*Competition and Markets Authority*), dentre outros. E os atores internacionais como a OECD, GPAI (*Council of Europe, Global Partnership on AI*), UNESCO, NATO (*AI Partnership for Defence*) e a *International Organisation for Standardisation and International Electrotechnical Commission* (ISO/IEC).

O Reino Unido e mais 50 países estão desenvolvendo um padrão de normalização ISO/IEC para IA que será conhecido como ISO/IEC 42001³. O desenvolvimento de padrões internacionais de IA envolve conceitos e terminologia; dados; tendência; implicações de governança; e ciclos de vida dos dados. Tal como acontece com outras normas do Sistema de Gestão, a ISO/IEC 42001 está sendo construída em torno de um processo circular de estabelecimento, implementação, manutenção e melhoria contínua da Inteligência Artificial. Este padrão ajudará as organizações a desenvolver ou usar IA de forma responsável e cumprir suas obrigações esperadas relacionadas às partes interessadas. O projeto de norma está agora em consulta pública que se encerrará em janeiro de 2023 (INTELLIGENCE, 2022).

³O *Standardization Roadmap Artificial Intelligence*, desenvolvido pelo *German Commission for Electrical, Electronic Information Technologies* (MAACK et al., 2020) apresenta o estado atual dos padrões e especificações internacionais para a IA e formulam algumas recomendações para as avaliações de conformidade e certificação. O relatório lista nas Tabelas 10, 11 e 12 os padrões e especificações existentes que tratam explicitamente de aplicações de IA. Além disso, fornece uma visão geral dos padrões e especificações que ainda não fazem declarações detalhadas sobre a aplicação de componentes de IA, mas são particularmente relevantes para a padronização ou aplicação de IA e fornece informações sobre projetos de padronização atuais.

5. Estados Unidos

5.1 Instituições para regulação e desenvolvimento da IA

Diante dos impactos potenciais de longo alcance da Inteligência Artificial (IA) e da importância atribuída ao tema ao longo dos últimos anos, governos têm buscado conceber estratégias nacionais de fomento à IA. Esse movimento está sendo percebido como uma corrida internacional pela IA (IPEA, 2022).

Nos Estados Unidos, essa corrida culminou no *National Initiative Act* (Lei da Iniciativa Nacional de Inteligência Artificial) de 2020 (GPO, 2020), o qual estabeleceu um comitê multidisciplinar americano, o *National Artificial Intelligence Initiative* (NAII), para tratar de temas relacionados à IA. Efetivamente, a NAII coordena as atividades ligadas à IA em andamento nas agências do governo federal e garante que cada agência se mantenha em contato e se informe sobre o que a outra está fazendo. A NAII tem como responsabilidade realizar essa gestão e implementar um programa coordenado entre agências do governo federal que visa acelerar o desenvolvimento da IA no território norte-americano. Com isso se espera posicionar os Estados Unidos como liderança em pesquisa e desenvolvimento de IA, nos âmbitos público e privado, e em todos os setores da economia e da sociedade.

A fim de assessorar o presidente e as entidades governamentais em tópicos relacionados ao *National Initiative Act*, foi criado em abril de 2022 o *National AI Advisory Committee* - NAIAC (NAIAC, 2022). Este colegiado é coordenado pelo Departamento de Comércio do Governo americano, e a sua composição inclui 26 representantes do setor privado, sociedade civil, entidades não governamentais e da academia, com liderança e expertise multidisciplinar em IA. Mediante diretriz do Congresso, o NAIAC pode ser consultado e prover recomendações sobre diversos tópicos relacionados à NAII, como: o estado da competitividade americana em IA; o estado da ciência em torno da IA; questões relacionadas à força de trabalho de IA; como a IA pode potencializar oportunidades em diversas regiões; como fazer o melhor uso dos recursos disponíveis na iniciativa; oportunidades de cooperação internacional; a necessidade de atualizar a NAII; questões relacionadas à prestação de contas; gestão da NAII; dentre outros.

A estrutura regulatória atual do uso de Inteligência Artificial no país possui caráter setorial e ocorre por meio da atuação de agências reguladoras e dos estados da federação (ROBERTS HUW; COWLS, 2021). Algumas cidades do país introduziram suas próprias normas na tentativa de mitigar possíveis impactos negativos do uso de Inteligência Artificial. É o caso da cidade de São Francisco, por exemplo, em 2019, ao tornar-se a primeira cidade do país a proibir o uso de reconhecimento facial por autoridades governamentais. A cidade de Portland é um segundo exemplo na construção de leis mais restritivas sobre o uso de reconhecimento facial e na proibição do uso por empresas privadas e uso governamental deste tipo de tecnologia (SIMONITE, 2021). Essas medidas demonstram as soluções

desenvolvidas para a obtenção de resultados éticos nessas cidades em consequência da ausência de regulamentações nacionais claras. No entanto, as medidas locais não afetam o âmbito federal (PACK, 2022).

5.2 Mitigação de riscos e vieses da IA

Ao longo dos anos e com o avanço da IA em diversos setores, métricas, ferramentas de avaliação e padrões tecnológicos foram desenvolvidos. O *National Institute of Standards and Technology* (NIST) define por meio da Ordem Executiva - EO 13859, de 2019, quais são as melhores práticas para o uso da IA.¹ Com isso, em resposta a Ordem Executiva, foram estabelecidas as orientações para uma regulamentação de aplicações de IA no setor privado através do Memorando M-21-06 (EOP, 2020). O documento expressa que ações regulatórias ou não regulatórias pelas agências federais devem priorizar a inovação e o crescimento da IA no país: “quando permitido por lei, ao decidir se e como regulamentar uma área que pode afetar as aplicações de IA, as agências devem avaliar o efeito da possível regulamentação sobre a inovação e o crescimento da IA” (NAII, 2022).

As recomendações do governo apontam que as agências reguladoras devem evitar uma abordagem que prejudique a posição dos Estados Unidos como líder global em inovação de IA. Como estabelecido pelo *National Artificial Intelligence Initiative* (NAII, 2022), onde a IA envolve riscos, as agências devem considerar os benefícios e custos potenciais de empregar a IA. Em outros cenários, a orientação para as agências é que usem sua autoridade para lidar com leis estaduais inconsistentes, onerosas e duplicadas que impeçam o surgimento de um mercado nacional. No entanto, quando um padrão nacional uniforme para um aspecto específico da IA não for essencial, as agências devem considerar renunciar à ação regulatória.

Alguns princípios são norteadores para que agências formulem abordagens regulatórias e não regulatórias para o projeto, desenvolvimento, implantação e operação de aplicativos de IA, tanto gerais quanto específicos do setor:

- Confiança pública na IA;
- Participação Pública;
- Integridade Científica e Qualidade da Informação;
- Avaliação e Gestão de Riscos;
- Benefícios e Custos;
- Flexibilidade;
- Justiça e Não Discriminação;
- Divulgação e Transparência;

¹Em 2019, o documento presidencial do Gabinete Executivo do Presidente publicou a Ordem Executiva 13859 com políticas e princípios para a liderança americana em inteligência artificial: os Estados Unidos devem promover avanços tecnológicos em IA; impulsionar o desenvolvimento de padrões técnicos apropriados; treinar as gerações atuais e futuras de trabalhadores americanos com as habilidades para desenvolver e aplicar tecnologias de IA; promover a confiança do público nas tecnologias de IA e proteger as liberdades civis; e promover um ambiente internacional que apoie a pesquisa e inovação americana em IA e abra mercados para as indústrias americanas de IA, ao mesmo tempo em que proteja a vantagem tecnológica em IA (E.O.13859, 2019).

- Segurança e Proteção; e,
- Coordenação Interinstitucional.

Em resposta ao Memorando M-21-06 e em uma tentativa de garantir uma IA responsável, agências como o *Department of Health and Human Services*, o *Department of Veterans Affairs*, a *United States Agency for International Development* e o *Department of Energy* prontamente publicaram seus princípios e orientações para o uso de IA. Da mesma forma, agências independentes, como o *Federal Trade Commission* (FTC), emitiram orientações sobre como utilizar a inteligência artificial de maneira transparente, explicável, justa, empiricamente sólida e responsável.

Já as diretrizes para o uso de Inteligência Artificial pelo Governo Federal foram estabelecidas a partir da Ordem Executiva 13960, com o objetivo de promover o uso confiável e garantir a transparência na divulgação de informações relevantes sobre o uso de IA para as partes interessadas apropriadas, incluindo o Congresso e o público, na medida do possível e de acordo com leis e políticas aplicáveis. Essas diretrizes orientam as agências a catalogar seus casos de uso de IA e convocam a *General Services Administration* (GSA) e o *Office of Personnel Management* para aprimorar a experiência na implementação da IA. De acordo com o princípio da transparência, portanto, a partir de 2022 as agências passaram a realizar um inventário anual de seus casos de uso de IA, divulgando-os, conforme orientação da Ordem Executiva 13960. No nível de agência e departamento, vários novos centros e escritórios de IA foram estabelecidos para promover a adoção de IA confiável nesses departamentos.

O país recentemente adotou também um guia para proteção contra as ameaças da IA. O *Blueprint for an AI Bill of Rights* não constitui uma nova lei ou uma política regulatória do atual governo dos Estados Unidos. O documento funciona como uma declaração de princípios, seguindo cinco premissas básicas para proteger o público na era da Inteligência Artificial:

1. Os sistemas automatizados devem ser eficazes e seguros;
2. Os usuários de tais sistemas devem ser protegidos contra a discriminação algorítmica, e o sistema deve ser projetado e usado de maneira equitativa;
3. As pessoas devem poder controlar como seus dados são usados e não devem estar sujeitas a práticas abusivas de dados;
4. Os usuários devem saber por que e como um sistema de IA fez sua determinação;
5. As pessoas devem ter a opção de optar por não tomar decisões de IA e recorrer a um ser humano se o sistema apresentar um erro, falhar ou se elas quiserem contestar a decisão.

O foco em limitar o excesso regulatório, portanto, traduz-se no que pode ser considerado como uma “Boa Sociedade de Inteligência Artificial” nos Estados Unidos. Em grande parte, a governança da IA encontra-se nas mãos do setor privado. Embora o *National AI Initiative Act* sinalize a atual necessidade de coordenação e transparência sobre o desenvolvimento de IA e pareça um passo promissor para medidas éticas e regulatórias, o fortalecimento de pesquisa e desenvolvimento focado em melhorar a competitividade do país segue sendo a pauta prioritária (ROBERTS HUW; COWLS, 2021).

Portanto, de acordo com Pack (2022), os Estados Unidos ainda dependem amplamente de agências preexistentes para administrar e fazer cumprir essas regulamentações; daqui para frente, pode-se esperar que outras entidades governamentais emitam orientações sobre como abordarão a regulamentação de tecnologias e aplicativos de Inteligência Artificial que se enquadram no âmbito de sua autoridade.

6. Austrália

6.1 Instituições para regulação e desenvolvimento da IA

O governo australiano, junto a empresas privadas, conceitua a inteligência artificial como uma solução tecnológica empolgante para fortalecer a economia e melhorar a qualidade de vida de todos os australianos, remodelando todos os setores, profissões e vidas. Não há leis específicas que regulem a IA na Austrália até o momento. Portanto, pela Austrália ainda estar em um estágio de desenvolvimento da regulação, não é possível discorrer sobre responsabilidades e divisão de competências entre os níveis de governo federal e subnacional ou o quão (des)centralizado setorialmente é a regulação. Todavia, o governo australiano está construindo as estruturas políticas regulatórias necessárias para perseguir o objetivo de posicionar a Austrália como um líder global em IA (DISER, 2021). As iniciativas incluem a implementação de três estratégias:

- *Australia's AI Action Plan* (DISER, 2021);
- *AI Technology Roadmap* (DISER;CSIRO, 2019); e
- *Australia's AI Ethics Framework*, que abarca oito princípios concebidos para garantir que a IA seja segura, protegida e confiável (DISER, 2019).

Todos os documentos-chave destacados acima mencionam a importância da ética e da segurança, especialmente no equilíbrio de interesses e incentivos comerciais, mas os métodos e recursos de desenvolvimento, implantação e monitoramento de IA ainda não são completamente compreendidos nas agências governamentais ou nos setores clientes (os que consomem a IA). Dessa forma, ainda não é possível tornar a funcionalidade de IA suficientemente transparente e compreensível para ter confiança em todas as suas aplicações.

Seguir uma abordagem tradicional para regulamentar novas tecnologias com base em como elas funcionam tornou-se insuficiente devido ao rápido avanço da tecnologia, criando uma disparidade entre o ritmo da mudança e a demora na resposta regulatória. Dessa forma, tentar vincular a regulamentação da IA à própria tecnologia subjacente (como os algoritmos de computador) criam limitações que costumam dificultar o controle dos inúmeros riscos associados às tecnologias de IA.

Após o lançamento do documento *Ethics Guidelines on Artificial Intelligence* pela União Europeia, em 2019, a Comissão Australiana de Direitos Humanos (ou *Australian Human Rights Commission (AHRC)*) publicou o *Human Rights and Technology Discussion Paper*, que apresenta a visão da comissão sobre regulação e propostas abrangentes para garantir que o governo e empresas australianas protejam os direitos humanos ao projetar, desenvolver e usar tecnologias com inteligência artificial (AHRC, 2019).

Uma forma prática de regulamentação que suporte as mudanças tecnológicas futuras da IA

parece ser o ponto de partida tanto para a UE quanto para a AHRC. Porém, cada uma possui as suas especificidades. A estrutura regulatória da União Europeia é frequentemente baseada em diferentes e específicos casos de uso da IA. Por exemplo, a proposta da UE pretende proibir o uso de IA para rastrear pessoas usando informações biométricas (como varreduras de reconhecimento facial) e para classificar seu comportamento, além de proibir seu uso enquanto ferramenta de pontuação social em casos de entrevistas de emprego e pedidos de empréstimo (COMMISSION, 2021b). A abordagem da UE tem a característica de fornecer orientação explícita para os casos específicos aos quais ela se dirige. Por outro lado, a Comissão Australiana de Direitos Humanos adotou uma abordagem diferente, cujo objetivo é desenvolver uma estrutura legal para IA que se concentre em regular o resultado da interação da IA com humanos, em vez de seus usos pretendidos (AHRC, 2019; AHRC, 2021). Por exemplo, em vez de proibir casos de uso específicos e correr o risco de erros de semântica, ele recomenda que qualquer indivíduo afetado seja notificado quando a inteligência artificial for materialmente usada na tomada de uma decisão administrativa. Essa notificação deve incluir informações sobre como um indivíduo afetado pode contestar a decisão. Dessa forma, independentemente do caso de uso, se as organizações tiverem que informar às pessoas sempre que a IA for usada para tomar uma decisão que as afete, ainda haverá um forte incentivo para garantir que sua IA seja desenvolvida usando as melhores práticas aceitas e seja adequada para processar dados privados.

6.2 Mitigação de riscos e vieses da IA

A confiabilidade no governo usando algoritmos automatizados para a tomada de decisão encontra-se baixa entre os cidadãos australianos devido aos esquemas *Robodebit* revelados em 2019¹, método ilegal de avaliação e recuperação automatizada de dívidas que ocorria desde 2016 (BRAITHWAITE, 2019). Como resultado da baixa confiança nos sistemas de IA por parte da população, em 2019 o governo australiano passou a financiar o desenvolvimento de uma estrutura de ética nacional para orientar empresas e governos que desenvolvem e implementam IA na Austrália. Chamada *Australia's Ethics Framework*, a estrutura inclui oito princípios de ética da IA para ajudar a reduzir o risco de impactos negativos da IA e garantir que o uso da IA seja sustentado por padrões de boa governança. Os princípios básicos da IA consistem em: 1- Não prejudicar; 2- Conformidade Regulatória; 3- Proteção e privacidade; 4- Equidade; 5- Transparência e Explicabilidade; 6- Contestação; e 7- Responsabilidade (DISER, 2019). O *Australia's Ethics Framework* examina questões-chave por meio da exploração de uma série de estudos de caso e tendências do debate ético de IA na Austrália e no mundo.

Além de conter os princípios de ética para uso da IA, o *Australia's Ethics Framework* possui um kit de ferramentas que fornece abordagens práticas e acessíveis para aproveitar o melhor do que a IA

¹O esquema *Robodebit* visava substituir o antigo sistema manual de cálculo de pagamentos indevidos e emissão de avisos de dívida para beneficiários de bem-estar por um sistema automatizado de correspondência de dados que comparava os registros do governo australiano com os dados de renda média do Escritório de Impostos Australiano. O esquema foi denunciado devido a alegações de avisos de dívidas falsas ou calculadas incorretamente (cerca de 470.000 dívidas emitidas indevidamente a serem pagas), preocupações com impactos na saúde física e mental dos destinatários de avisos de dívidas, e questões sobre a legalidade do esquema (BRAITHWAITE, 2019).

tem a oferecer. O kit de ferramentas abrange: 1- Avaliação de Impacto; 2- Revisão Interna ou Externa; 3- Avaliações de risco; 4- Diretrizes de Melhores Práticas; 5- Padrões da Indústria; 6- Colaboração; 7- Mecanismo de monitoramento e melhoria; 8- Mecanismo de recurso; 9- Consulta (DISER, 2019). O terceiro item desse kit, a avaliação de risco, aparece como uma etapa fundamental para o tratamento dos riscos apresentados pelo uso do sistema de IA. Ela é útil para fornecer um limite e gatilhos para ações adicionais e processos de mitigação de riscos. Como pode ser observado na Figura 6.1, a estrutura para avaliar fatores de risco, proposta pelo DISER (2019), examina probabilidade de risco (raro, improvável, possível, provável, quase certo), tipo de risco (risco insignificante, risco menor, risco moderado, grande risco, risco crítico) e a consequência (baixa, moderada, alta e extrema). Quando um risco tem alta probabilidade de ocorrência e resultados mais negativos, as consequências tornam-se mais graves. A partir disso, a Figura A1, no Apêndice, examina os fatores que podem fazer com que um aplicativo de IA contenha mais riscos. Conforme observado na figura, diferentes cenários podem conter mais ou menos riscos dependendo das circunstâncias, mas o guia fornece uma visão geral das áreas que provavelmente contêm riscos que devem ser considerados antes da implementação da IA. Além disso, conforme o tipo de risco, o relatório ainda sugere uma variedade de ações que podem ser tomadas para mitigar o risco.

Figura 6.1: Estrutura de avaliação de risco para sistemas de IA

Probabilidade de risco	Consequência				
	Risco insignificante	Risco menor	Risco moderado	Grande risco	Risco crítico
Raro	Baixa	Baixa	Moderada	Alta	Alta
Improvável	Baixa	Moderada	Moderada	Alta	Extrema
Possível	Baixa	Moderada	Alta	Alta	Extrema
Provável	Moderada	Alta	Alta	Extrema	Extrema
Quase certo	Moderada	Alta	Alta	Extrema	Extrema

Fonte: Adaptado de DISER (2019).

De acordo com DISER (2019), para implementar ética na IA não é necessário reinventar leis e padrões éticos, mas sim atualizar os existentes e torná-los adequados ao contexto da IA. Esses esforços são essenciais para que o cidadão ganhe confiança na aplicação de tecnologias e produtos de IA. Entretanto, por enquanto, não há leis específicas que regulem a IA na Austrália. Todavia, conforme descrito pelo *Australia's Ethics Framework*, já há uma série de leis que poderiam moldar indiretamente a adoção e implementação dessas tecnologias, incluindo aquelas relacionadas à privacidade e segurança de dados, regulamentação de serviços financeiros e as leis de propriedade intelectual.

Não há direito de propriedade para uma IA ou algoritmo. No entanto, se o algoritmo de IA for implementado via software, ele será protegido pela *Copyright* de 1968 (Lei de Direitos Autorais) ligada à regulação de softwares. Além disso, também não há leis específicas de propriedade de dados na Austrália, mas algumas leis se aplicam à segurança de dados e à privacidade das informações, como a *Privacy Act* de 1988 (Lei de Privacidade), que estabelece procedimentos de coleta, uso e divulgação

de informações pessoais². Uma vez que não há direitos autorais gerais sobre dados na Austrália, o *Australia's Ethics Framework* recomenda que as empresas usem acordos comerciais para esclarecer seus direitos e as posições acordadas sobre como cada parte terá acesso aos dados, como poderá usá-los e em que medida. Além disso, o regime de responsabilidade civil da Austrália não contempla ou aborda danos ou prejuízos resultantes do uso da IA. Porém, o *Australia's Ethics Framework* sugere que a reparação das vítimas passe a ser abordada contratualmente de modo que o contrato indique claramente quem é o responsável por qualquer dano resultante do uso da IA.

²Em 2018, a *Privacy Act* foi alterada para incluir um esquema de violação de dados notificável, como parte da tomada de decisões computadorizada (PARLIAMENT OF AUSTRALIA, 2018).

7. Japão

7.1 Instituições para regulação e desenvolvimento da IA

Nas últimas décadas, o governo japonês e o setor privado, fizeram grandes investimentos em tecnologias de inteligência artificial (IA), no intuito de aumentar a eficiência econômica e a qualidade de vida da população, atender à escassez de mão de obra e impulsionar a competitividade futura do país em uma realidade em que a sociedade se encontrará envelhecida. A partir da iniciativa governamental, a IA está sendo incorporada rapidamente em produtos e processos, como automação de processos robóticos (RPA), aprendizado de máquina, processamento de linguagem natural, auxiliando a encontrar novos medicamentos e terapias (TAEIHAGH, 2021).

No Japão, o campo da IA não é dominado por apenas um tipo de ator, mas existe de uma colaboração de atores nos domínios público, privado e de pesquisa. Como já foi dito anteriormente, não existe regulamentação abrangente da IA no Japão, no entanto, o governo japonês oferece algumas diretrizes que auxiliam e incentivam a utilização da IA de formar responsável. O Gabinete Japonês tem um papel de coordenação em relação a questões estratégicas. A coordenação de esforços e troca de informações é de responsabilidade do Conselho de Ciência, Tecnologia e Inovação (CSTI). Sob a liderança do Primeiro-Ministro e do Ministro de Estado para a Política de Ciência e Tecnologia, o CSTI supervisiona toda a ciência e tecnologia do país, aconselha e formula políticas. A execução da estratégia de IA está dividida em três ministérios: Ministério da Administração Interna e Comunicação (MIC), Ministério da Economia, Comércio e Indústria (METI) e Ministério da Educação, Cultura, Desporto, Ciência e Tecnologia (MEXT). O MIC é responsável pelo impacto social e pelas necessidades governamentais em relação à IA. Finalmente, o MEXT é responsável pelos aspectos educacionais e de pesquisa, promovendo a pesquisa, bem como novos programas de estudo e outras ações para enfrentar a escassez de recursos humanos (JAPAN, 2022a). Não foi possível observar como estão definidas as responsabilidades da regulação aos níveis federal e subnacional.

Os países estão em diferentes estágios de implementação de suas estratégias nacionais de IA. Depois do Canadá, o Japão foi o segundo país a desenvolver estratégias nacionais de IA estabelecendo metas e alocando orçamento (OECD, 2021). Desde 2019, o Gabinete Japonês vem formulando uma estratégia de IA com o objetivo de fornecer um pacote abrangente de políticas sobre IA para enfrentar os desafios sociais do Japão e melhorar a competitividade de suas indústrias (JAPAN, 2019a). Em sua versão mais recente, lançada em 22 de abril de 2022, a Estratégia de IA de 2022, descreve os cinco objetivos estratégicos a seguir e estabelece um plano de ação alinhado com esses objetivos estratégicos (JAPAN, 2022a):

- Estabelecer um sistema e infraestrutura técnica capaz de proteger, na medida do possível, a vida e a propriedade dos residentes contra crises iminentes, como pandemias e desastres de

grande escala, e estabelecer uma estrutura para a operação adequada e sustentável de tal sistema e infraestrutura técnica.

- Desenvolver os recursos humanos mais capazes para a era da IA e tornar-se um país que pode atrair talentos em todo o mundo e estabelecer uma estrutura para atingir esses objetivos de forma sustentável.
- Tornar-se o precursor em aplicativos de IA em indústrias do mundo real para aumentar a competitividade industrial.
- Estabelecer um conjunto de sistemas técnicos para alcançar uma “sociedade sustentável que englobe a diversidade” e estabelecer uma estrutura para operar esses sistemas.
- Sob a liderança do Japão, estabelecer uma rede internacional de pesquisa, treinamento e infraestrutura social em IA para acelerar a pesquisa e o desenvolvimento de IA e o desenvolvimento de recursos humanos.

Para atingir os objetivos supracitados, a estratégia de IA descreve um roteiro e um plano de ação que consiste em seis componentes:

1. Desenvolvimento de recursos humanos para IA;
2. Implementação de aplicações de IA em cinco áreas prioritárias (saúde, agricultura, infraestrutura, transporte e desenvolvimento regional);
3. Promoção de pesquisa e desenvolvimento;
4. Apoio a entidades que utilizam IA (pequenas empresas e municípios);
5. Construção de infraestrutura de dados; e
6. Princípios éticos na utilização de IA.

Até o momento, não há uma regulamentação abrangente de IA no Japão, no entanto, o Gabinete Japonês e vários ministérios têm buscado influenciar a IA a partir de várias estratégias e diretrizes. Dessa forma, o Japão adota o chamado *soft law* para preencher a lacuna de governança sobre a regulamentação da IA. Esse inclui vários tipos de instrumentos que não são diretamente executáveis pelo governo. Há ainda abordagens desprovidas de caráter jurídico, como diretrizes profissionais, padrões privados, códigos de conduta e melhores práticas. O relatório intitulado *Governance Innovation* afirma que é importante que o governo e o setor privado trabalhem juntos para estabelecer padrões, diretrizes e outras “soft laws” para reforçar os esforços das empresas para implementar a governança ágil¹ (METI, 2021).

Em outubro de 2016, o Ministério de Assuntos Internos e Comunicações do Japão organizou um grupo de estudo composto por participantes multissetoriais da indústria, da academia e do setor privado chamado de "Conselho de Promoção para a Sociedade de Rede IA", para aprofundar a discussão sobre a rede AI e formular diretrizes. Este conselho publicou dois documentos: o primeiro, em 2017, chamado

¹De acordo com o METI (2021), a governança das sociedades futuras exigirá métodos de governança que sejam flexíveis e adaptáveis a circunstâncias, em oposição a abordagens em que regras e procedimentos detalhados são previamente definidos ou fixados. Como alternativa, o governo japonês busca implementar a governança ágil, que é um modelo de governança projetado para executar ciclos contínuos e rápidos de “análise de condições e riscos”, “estabelecimento de metas”, “desenho de sistema”, “implementações”, “avaliações”, e “melhorias”.

de *Draft AI R&D guidelines for international discussion*², no qual o Japão reconhecia os grandes benefícios que a IA traria para a sociedade e a economia e propunham princípios de desenvolvimento de IA direcionados aos desenvolvedores de IA (SUDOH; MITOMO et al., 2017); e o segundo, *AI Utilization Guidelines*, publicado em agosto de 2019, que fornecia orientações para usuários de IA incluindo provedores de serviços de IA e usuários comerciais de sistemas de IA (SUDOH; MITOMO et al., 2019).

7.2 Mitigação de riscos e vieses da IA

Riscos foram identificados na utilização da IA, como falta de transparência e perda de controle. Tendo isso em vista, em março de 2019, o Gabinete Japonês divulgou um documento intitulado *Social Principles of Human Centric AI*, com o objetivo de promover a implementação social apropriada e proativa da IA, proteger os usuários e evitar a propagação de riscos, além de promover uma sociedade centrada no ser humano e pronta para a IA. Os princípios atrelados a este documento não são princípios legais em si, mas podem ser usados como diretrizes para interpretar as leis existentes (JAPAN, 2019b).

Acerca das diretrizes governamentais, no dia 28 de janeiro de 2022, o Ministério da Economia, Comércio e Indústria (METI) emitiu suas *Governance Guidelines for Implementation of AI Principles* (a versão 1.0 foi lançada em 15 de janeiro de 2021). Essas diretrizes, integram os "Princípios sociais da IA centrada no ser humano" e descrevem o que deve ser colocado em prática ao aderir esses princípios. Dessa forma, as Diretrizes de Governança para Implementação dos Princípios de IA destinam-se a apoiar a implementação dos princípios sociais da IA centrada no ser humano pelas entidades envolvidas no desenvolvimento e operação da tecnologia habilitada para IA. Embora não sejam juridicamente vinculados, fornecem objetivos de ação a implementar, exemplos práticos de cada objetivo de ação e exemplos de como os desvios dos objetivos de ação podem ser avaliados, servindo assim como uma referência útil (JAPAN, 2022b).

De acordo com a *Governance Guidelines for Implementation of AI Principles*, o Ministério da Economia, Comércio e Indústria (METI) optou por diretrizes voluntárias para o uso de inteligência artificial, uma abordagem diferente do que está acontecendo na União Europeia, onde os reguladores usam regras mais rígidas para regular o viés implícito na tomada automática de decisões do aprendizado de máquina. Ainda no documento, o METI ressalta que as empresas devem ter transparência e responsabilidade no uso de aprendizado de máquina, algoritmos e processos automáticos de tomada de decisão, compreendendo os impactos positivos e negativos que os sistemas de IA podem ter³. Além

²Estas Diretrizes consistem nos Princípios de PD de IA, entre eles estão: 1) Princípio da colaboração - Os desenvolvedores devem prestar atenção à interconectividade e interoperabilidade dos sistemas de IA; 2) Princípio da transparência - Os desenvolvedores devem prestar atenção à verificabilidade das entradas/ saídas de sistemas de IA e a explicabilidade de seus julgamentos; 3) Princípio da controlabilidade - Os desenvolvedores devem prestar atenção à controlabilidade dos sistemas de IA; 4) Princípio da segurança- Os desenvolvedores devem levar em consideração que os sistemas de IA não prejudicarão a vida, o corpo, ou propriedade de usuários ou terceiros por meio de atuadores ou outros dispositivos.

³Todas as diretrizes apontadas pelo *Governance Guidelines for Implementation of AI Principles*, incluindo análise de riscos, compreensão da aceitação social e a definição de metas, entre outras, são baseadas em exemplos práticos e hipotéticos para a implementação dos princípios sociais (JAPAN, 2022b). Por exemplo, para compreender a aceitação social do desenvolvimento e operação do sistema de IA, o relatório recomenda a realização de pesquisas com questionários

disso, sob uma abordagem baseada em risco, as empresas também são aconselhadas a definir metas de governança de IA e examinar como resolver quaisquer lacunas, fazendo com que as equipes de gerenciamento preparem um plano de resposta a incidentes e avaliem a situação de aceitação social.

Embora os sistemas de IA possam ter um impacto positivo nas empresas, ajudando-as a superar a escassez de RH, melhorar a produtividade e desenvolver negócios de alto valor agregado, o desenvolvimento e a operação de sistemas de IA são acompanhados de riscos exclusivos da IA. Diferente da proposta da União Europeia, que classifica aplicações de IA de riscos mínimos a inaceitáveis e visa regular sistemas de IA de alto risco, o governo japonês espera que as empresas compreendam a essência de cada uma das diretrizes do documento *Governance Guidelines for Implementation of AI Principles* como companheiras para seus esforços, em vez de uma adesão rígida a elas. O relatório *Governance Innovation* afirma que modelos convencionais de governança que colocam leis e regulamentos no centro enfrentam dificuldades para acompanhar a velocidade da inovação. Assim, tais modelos de governança têm causado problemas onde as leis não conseguem controlar novos riscos que podem ser trazidos pelas inovações, além de dificultar o desenvolvimento de inovações (METI, 2021).

Para enfrentar esses desafios, o Japão dispensa os modelos de governança convencionais baseados em regras e adota modelos baseados em metas que possam orientar entidades, como empresas. O *Social Principles of Human Centric AI* estabelece as metas para o uso da IA, no qual os desenvolvedores e provedores de serviços de IA devem definir e cumprir seus próprios objetivos (princípios de desenvolvimento e utilização de IA), que devem ser implementados de acordo com seus próprios propósitos e meios de desenvolvimento e operação de IA, dentre outros⁴. Por outro lado, um dos problemas das regulamentações baseadas em metas é que elas podem criar uma grande lacuna entre as metas compartilhadas pela sociedade e as operações práticas para atingir as metas no nível corporativo, já que essas metas dependem de iniciativas voluntárias de cada empresa.

Sobre diretrizes relacionadas a riscos e alguns desenvolvimentos recentes, em 2018, o Ministério da Economia, Comércio e Indústria (METI) publicou o documento “Diretriz para Contratos de Utilização de IA e Dados”, que diz respeito às novas questões legais incluem como lidar com os direitos de propriedade e identificar quem é legalmente responsável caso ocorra algum incidente no desenvolvimento de software e utilização da tecnologia de IA (METI, 2018). Em julho de 2020, a Agência de Assuntos do Consumidor concluiu que é desejável que a IA seja usada de maneira inteligente, melhorando a alfabetização do cliente. Para atingir esse objetivo, a agência lançou o *AI Utilization Handbook - Keys to Effectively Using AI*. A seção básica do manual descreve a estrutura, características e limitações da IA, bem como os tipos de serviços que utilizam IA. A seção de ponto de verificação fornece informações sobre os mecanismos e considerações para cada tipo de produto e

de consumidores publicadas por governos, instituições públicas e privadas, além de citar o exemplo da Agência de Assuntos do Consumidor, no Grupo de Trabalho de IA sob o Comitê de Desafios à Digitalização do Consumidor, que realizou pesquisas por questionário e publicou seus resultados sobre o status da compreensão dos consumidores sobre IA.

⁴De acordo com Japan (2019b), os princípios sociais da IA compreendem o seguinte: 1- princípio centrado no ser humano; 2- princípio da educação e alfabetização; 3- princípio da proteção da privacidade; 4- princípio de garantia da segurança; 5- princípio da concorrência leal; 6- princípios de justiça, responsabilidade e transparência; e 7- princípio da inovação.

serviço, bem como considerações de gerenciamento de dados, levando em consideração o uso real do consumidor de serviços baseados em IA.

No entanto, as questões de insegurança dos consumidores em relação à utilização da IA não podem ser resolvidas simplesmente melhorando a alfabetização do consumidor. Essa direção de melhorar a alfabetização do consumidor é baseada na suposição de que as empresas que usam IA fornecem informações apropriadas aos consumidores. Portanto, uma diretriz intermediária para preencher a lacuna entre os Princípios de IA e a prática em nível corporativo é considerada necessária no sentido de facilitar esse fornecimento de informações.

Mais recentemente, o relatório *AI Governance in Japan Ver. 1.1* publicado em 2021 e elaborado pelo *Ministry of Economy, Trade and Industry* (METI), concluiu que as regulamentações juridicamente vinculativas para sistemas de IA são consideradas desnecessárias no momento. Mesmo que as discussões sobre os requisitos juridicamente vinculativos sejam realizadas no futuro, a avaliação de riscos deve ser implementada levando em consideração não apenas os riscos, mas também os benefícios potenciais. O relatório afirma que deve-se considerar a possibilidade de que certos riscos possam ser eliminados devido ao desenvolvimento de tecnologias. Por exemplo, embora a tecnologia de reconhecimento facial baseada em IA seja usada para *smartphones* e outros dispositivos, a tecnologia é eficaz na prevenção de login não autorizado por terceiros, e muitos usuários estão cientes de que a IA é usada e podem optar por não usá-la.

O relatório *AI Governance in Japan*, ainda recomenda que em certas áreas, como o setor automotivo e de saúde, as regulamentações devem estar sob responsabilidade das organizações responsáveis pelas leis do setor. Por fim, como diretrizes intermediárias não vinculativas não são legalmente obrigatórias, pode não haver incentivo suficiente para usá-las. Como recomendação futura, o relatório sugere mecanismos que associem o uso da diretriz a benefícios. Por exemplo, a empresa que respeita o princípio da IA receberá pontos adicionais no processo de compras governamentais do sistema de IA.

Resumindo, as agências governamentais japonesas adotam a abordagem de *soft-law*, exceto para algumas áreas específicas⁵, para lidar com possíveis abusos e vieses da tecnologia de IA. O Japão busca uma governança ágil, com o objetivo de não prejudicar investimentos e não impedir a inovação, tornando a abordagem mais flexível do METI adequada para o país. O Japão aparece como um estado proativo, que estimula ativamente o desenvolvimento e a implantação de tecnologias de IA.

⁵Em Janeiro de 2022, Comissão de Proteção de Informações Pessoais, ou PPC, passou a estudar o uso de imagens faciais e dados de câmeras de segurança, depois que a principal operadora ferroviária do Japão implantar câmeras equipadas com recursos de reconhecimento facial em estações para detectar ex-presidiários e suspeitos de terrorismo. Além disso, em março de 2022, o METI e o Ministério da Administração Interna e Comunicação (MIC) revisaram o *Guidebook for Utilization of Camera Images*, que especificam as considerações para a revisão e implementação de projetos que utilizam imagens de câmeras e para estabelecer um entendimento mútuo entre consumidores e empresas que podem ser capturados em imagens de câmeras (METI; MIC, 2022).

8. Considerações Finais

Este trabalho analisou características sobre o desenvolvimento e regulamentação de Inteligência Artificial na União Europeia, Reino Unido, Estados Unidos, Austrália e Japão.

Na União Europeia, há um longo projeto de lei sobre regulação da IA em fase final de aprovação. Ele é centrado no consumidor e possui uma abordagem de risco com níveis que variam de sistemas que são totalmente banidos (como por exemplo, sistemas de pontuação social) para sistemas que exigem disposições mínimas de transparência (ex. filtros de spam). A lei procura garantir que os sistemas de IA de alto risco cumpram os direitos fundamentais e a segurança do usuário, por meio de uma supervisão regulatória exercida pelo estado e por entidades privadas. São propostas obrigações proporcionais aos produtores, fornecedores, utilizadores e outros participantes de toda a cadeia de valor da IA. A regulação é feita por organismos de avaliação de conformidade fiscalizados por cada Estado Membro.

No Reino Unido, observa-se a construção de uma Estratégia Nacional de IA com uma visão de inovação e oportunidade em primeiro plano, tentando se posicionar com menos rigor em relação à proteção de dados e regulamentações. Com uma política regulatória intersetorial, a estratégia não deixa claro as prováveis posições a serem alcançadas sobre a responsabilidade legal e não divulga a metodologia de desenvolvimento do regulamento/normas ou como os *stakeholders* serão engajados no processo, além do fato de não explicitar as responsabilidades e as competências aos níveis dos países membros do Reino Unido.

Já nos Estados Unidos, leva-se em consideração a "Boa Sociedade de Inteligência Artificial", priorizando o desenvolvimento e a corrida pela liderança do país em IA. Apesar dos esforços recentes em produzir princípios norteadores para proteção contra os riscos existentes no uso da IA, o foco em melhorar a competitividade do país segue sendo pauta prioritária. A estrutura regulatória segue tendo caráter setorial e ocorre por meio da atuação de agências e dos estados da federação.

A Austrália, por sua vez, encontra-se em um estágio de desenvolvimento de estruturas políticas e regulatórias da IA. Não há, até o momento, leis específicas sobre o uso de Inteligência Artificial. No entanto, o governo australiano desenvolveu recentemente princípios básicos de ética como uma tentativa de reduzir os impactos negativos e aumentar a confiabilidade do uso desse tipo de tecnologia. Assim como nos Estados Unidos, o objetivo da construção de uma estrutura de IA pelo governo prioriza posicionar a Austrália como líder global.

No Japão, a estrutura é colaborativa. Embora não haja uma regulamentação de IA abrangente no país, o governo oferece diretrizes para o desenvolvimento de Inteligência Artificial responsável e em colaboração com atores nos domínios públicos, privados e de pesquisa. O chamado "*soft law*" impera quando trata-se de governança sobre IA. O Japão busca uma governança ágil, com o objetivo de não prejudicar investimentos e não impedir a inovação,

Dentre os casos analisados, a partir da Tabela 8.1, é possível observar que todos possuem uma estratégia nacional para fomentar o desenvolvimento da IA. Essa estratégia se traduz na criação de instituições, como comissões, que funcionam como órgãos coordenadores das iniciativas governamentais e como ponte entre os setores privado e público. Apesar dos países possuírem estratégias, a maioria ainda está amadurecendo a sua estrutura regulatória e nos próximos meses novas diretrizes e documentos sobre o tema devem ser publicados. Apesar dessa ser a situação em termos absolutos na “corrida pela IA”, alguns países estão mais avançados do que outros. Em termos relativos, apenas a Austrália aparece como um país atrás dos demais em discussões sobre como se será feita a regulação pelas agências de governo (se de forma mais setorial, focando na especificidade da área de aplicação da IA, ou de forma geral).

Tabela 8.1: Características da regulação e do desenvolvimento da IA

	Há estratégia nacional para IA?	Estágio de maturidade da regulação	Tipo de abordagem regulatória	Há mecanismos para mitigar riscos?	Há normas padrões para IA?
UE	Sim	Avançado	Geral e Setorial	Sim	Não
Reino Unido	Sim	Avançado	Setorial	Sim	Sim
EUA	Sim	Avançado	Setorial	Sim	Não
Austrália	Sim	Inicial	Não Definido	Sim	Não
Japão	Sim	Avançado	Setorial	Sim	Não

Fonte: Elaboração própria com base nos dados da pesquisa.

Explorar a estrutura regulatória dos países se constitui como um importante exercício para o caso brasileiro, pois o país se encontra em um estágio de desenho das políticas públicas sobre esse tema. O que se percebe analisando os casos explorados é que as diretrizes e ações para fomento e regulação da IA são muito recentes e, portanto, ainda é cedo para mensurar as consequências das estratégias escolhidas por cada país. Não é possível apontar qual modelo adotado por é o melhor e os resultados deste *benchmarking* não devem ser interpretados dessa forma. Todavia, é possível observar suas características e refletir sobre como o Brasil pode avançar na sua estratégia para IA. Apesar das políticas públicas para regular e fomentar a IA ainda sejam recentes, é importante que o Brasil se engaje nessas discussões, pois há uma corrida internacional entre os países para tornar suas organizações privadas e públicas mais competitivas no uso da IA (IPEA, 2022).

Referências Bibliográficas

AHRC. Human rights and technology discussion paper. Australian Human Rights Commission, 2019. Citado 2 vezes nas páginas 20 e 21.

AHRC. Human rights and technology final report. Australian Human Rights Commission, 2021. Citado na página 21.

BRAITHWAITE, V. Beyond the bubble that is robodebt: How governments that lose integrity threaten democracy. *Australian journal of social issues*, Wiley Online Library, v. 55, n. 3, p. 242–259, 2019. Citado na página 21.

CASSELL, C.; SYMON, G. *Essential guide to qualitative methods in organizational research*. [S.l.]: sage, 2004. Citado na página 7.

COMMISSION, E. Communication from the commission to the european parliament, the european council, the council, the european economic and social committee and the committee of the regions on artificial intelligence for europe. European Commission, Brussels, Belgium, 2018a. Citado na página 8.

COMMISSION, E. Communication from the commission to the european parliament, the european council, the council, the european economic and social committee and the committee of the regions on artificial intelligence for europe - coordinated plan on artificial intelligence (com(2018) 795 final). European Commission, Brussels, Belgium, 2018b. Citado na página 8.

COMMISSION, E. White paper on artificial intelligence – a european approach to excellence and trust. European Commission, Brussels, Belgium, 2020. Citado na página 8.

COMMISSION, E. Communication from the commission to the european parliament, the european council, the council, the european economic and social committee and the committee of the regions on artificial intelligence for europe - coordinated plan on artificial intelligence - review. European Commission, Brussels, Belgium, 2021. Citado na página 8.

COMMISSION, E. Proposal for a regulation of the european parliament and of the council laying down harmonised rules on artificial intelligence and amending certain union legislative acts (artificial intelligence act). European Commission, Brussels, Belgium, 2021. Citado 6 vezes nas páginas 5, 8, 9, 10, 11 e 21.

DISER. Australia's ai ethics framework. Department of Industry, Science, Energy and Resources, 2019. Citado 4 vezes nas páginas 20, 21, 22 e 34.

DISER. Australia's ai action plan. Department of Industry, Science, Energy and Resources, 2021. Citado na página 20.

DISER;CSIRO. Ai technology roadmap. Department of Industry, Science, Energy and Resources; Commonwealth Scientific and Industrial Research Organisation, 2019. Citado na página 20.

- E.O.13859. *Maintaining American Leadership in Artificial Intelligence*. Washington, DC, 2019. Citado na página 17.
- EOP. *Memorandum for the heads of executive departments and agencies*. Washington, DC, 2020. Citado na página 17.
- FLORIDI, L. et al. capai-a procedure for conducting conformity assessment of ai systems in line with the eu artificial intelligence act. *Available at SSRN 4064091*, 2022. Citado na página 11.
- GPO. *National Artificial Intelligence Initiative Act of 2020*. Washington, DC, 2020. Citado na página 16.
- ICO, I. C. O. *Guidance on ai and data protection*. Information Commissioner's Office, UK, 2020. Citado na página 13.
- INTELLIGENCE, E. C. European Commission High-Level Expert Group on A. *Ethics guidelines for trustworthy ai*. European Commission, Brussels, Belgium, 2019. Citado na página 8.
- INTELLIGENCE, E. C. European Commission High-Level Expert Group on A. *Policy and investment recommendations for trustworthy ai*. European Commission, Brussels, Belgium, 2019. Citado na página 8.
- INTELLIGENCE, O. for A. *National ai strategy*. Office for Artificial Intelligence, UK, 2021. Citado 3 vezes nas páginas 13, 14 e 15.
- INTELLIGENCE, O. for A. *Establishing a pro-innovation approach to regulating ai*. Office for Artificial Intelligence, UK, 2022. Citado 2 vezes nas páginas 14 e 15.
- IPEA. *Exame comparativo das estratégias nacionais de inteligência artificial: Argentina, brasil, chile, colômbia e coreia do sul*. IPEA, 2022. Citado 3 vezes nas páginas 5, 16 e 30.
- JAPAN, C. O. of. *Ai strategy 2019*. 2019. Citado na página 24.
- JAPAN, C. O. of. *Social principles of human centric ai (2019)*. 2019. Citado 2 vezes nas páginas 26 e 27.
- JAPAN, C. O. of. *Ai strategy 2022*. 2022. Citado na página 24.
- JAPAN, C. O. of. *Governance guidelines for implementation of ai principles*. 2022. Citado na página 26.
- MAACK, S. et al. *Deutsche normungsroadmap künstliche intelligenz*. 2020. Citado na página 15.
- MCTI. *Estratégia de IA*. 2022. <<https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/inteligencia-artificial>>. Acesso em: 07 de dez. de 2022. Citado na página 5.
- METI. *Contract guidelines on utilization of ai and data, version 1.1*. Ministry of Economy, Trade and Industry, 2018. Citado na página 27.
- METI. *Governance innovation ver. 2 - a guide to designing and implementing agile governance*. Ministry of Economy, Trade and Industry, 2021. Citado 2 vezes nas páginas 25 e 27.
- METI; MIC. *Guidebook for utilization of camera images ver. 3.0*. Ministry of Economy, Trade and Industry and Ministry of Internal Affairs and Communications, 2022. Citado na página 28.

- NAIAC. *The National AI Advisory Committee*. 2022. <<https://www.ai.gov/naiac/>>. Acesso em: 07 de dez. de 2022. Citado na página 16.
- NAII. *National AI Initiative*. Washington, DC, 2022. Citado na página 17.
- OCDE. *Artificial intelligence in society*. 2019. Citado na página 5.
- OECD. *State of implementation of the OECD AI principles: insights from national AI policies*. [S.l.]: OECD Paris, 2021. Citado na página 24.
- PACK, H. *Regulation of artificial intelligence in the united states*. 2022. Citado 2 vezes nas páginas 17 e 19.
- ROBERTS HUW; COWLS, J. H. E. M. F. T. A. T. M. F. L. *Achieving a good ai society: Comparing the aims and progress of the eu and the us*. *Science and Engineering Ethics*, v. 27, n. 68, 2021. Citado 2 vezes nas páginas 16 e 18.
- SIMONITE, T. *Portland’s face-recognition ban is a new twist on ‘smart cities’*. *Wired*, 2021. Citado na página 16.
- SUDOH, O.; MITOMO, H. et al. *Draft ai r&d guidelines for international discussions*. In: *Tokyo: The Conference toward AI Network Society*. [S.l.: s.n.], 2017. Citado na página 26.
- SUDOH, O.; MITOMO, H. et al. *Ai utilization guidelines.practical reference for ai utilization*. In: *Tokyo: The Conference toward AI Network Society*. [S.l.: s.n.], 2019. Citado na página 26.
- TAEIHAGH, A. *Governance of artificial intelligence*. *Policy and Society*, Oxford University Press, v. 40, n. 2, p. 137–157, 2021. Citado na página 24.

Apêndice 1

Figura A1: Fatores que determinam o risco da IA - Baseado no Australia's Ethics Framework

Fator de risco	Avaliação do risco de acordo com a natureza da atividade da IA				
	Insignificante	Menor	Moderado	Alto	Crítico
Proteção de privacidade	IA não usa dados sensíveis ou privados	Aplicação utiliza uma pequena quantidade de dados privados	IA usa um grande número de dados privados das pessoas	A IA foi projetado de forma a possibilitar a coleta de informações sobre indivíduos sem seu consentimento expresso	IA usa um grande banco de dados de dados privados ou confidenciais.
Imparcialidade	IA não afeta os direitos de uma pessoa.	Há opt-in e opt-out na aplicação.	Há opt-in, mas o opt-out não é claro.	A pessoa não tem a opção de cancelar sua inscrição no aplicativo e a IA se afeta os direitos, é de uma única pessoa.	A aplicação afeta os direitos de um grande grupo de pessoas.
Dano físico	Aplicativo não pode controlar ou influenciar outros sistemas.	O aplicativo controla equipamentos incapazes de causar danos significativos, incômodo público ou vigilância.	A aplicação pode controlar equipamentos pesados ou materiais perigosos, mas o risco de prejudicar pessoas ou causar incômodo ao público é limitado.	A IA controla equipamentos pesados/materiais perigosos e espera-se que opere em um espaço público.	A IA pode controlar equipamentos que podem causar perda de vida ou o equipamento foi projetado para coletar informações pessoais secretamente.
Contestabilidade	O aplicativo opera com base no opt-in e nenhuma intervenção é necessária para reverter o resultado caso alguém decida não participar.	O aplicativo é ativado automaticamente, mas há uma notificação clara para o usuário e é fácil cancelar sua inscrição. Também é fácil obter ajuda humana para fazer isso.	O aplicativo será amplamente utilizado entre o público, mas há poucos ou nenhum recurso humano disponível para prover auxílio.	Não há uma maneira fácil de optar por não participar.	O aplicativo não permite descadastramento e a pessoa afetada não dispõe de meios para alterar esse resultado.
Responsabilidade	Há clara responsabilização pelos resultados gerados.	Há precedente legal relacionado aos resultados entregues e há uma clara cadeia de responsabilidade.	Há delimitação razoavelmente clara de responsabilidade entre usuários e provedores.	Há pouco ou nenhum precedente legal para este aplicativo e não há separação de responsabilidade entre usuários e provedores.	A responsabilidade legal pelo resultado do uso é pouco clara.
Regulatório e Observância Legal	Consentimento obtido para uso de dados. Aplicação opera em território legal familiar.	Uma aplicação idêntica ou semelhante de IA tem precedente legal demonstrando conformidade. O consentimento é facilmente obtível.	Há pouco precedente legal para o pedido, mas houve procura por extenso aconselhamento jurídico e o pedido foi revisado por terceiros.	Não há precedente legal para o pedido e nenhum terceiro ou especialista jurídico foi consultado. O consentimento não é claro.	Nenhum consentimento obtido para uso de grandes quantidades de dados privados. Não há precedente legal claro.
Transparência e explicabilidade	Entradas, algoritmo e resultados gerados da IA são bem compreendidas.	O aplicativo usa uma IA difícil de explicar, como redes neurais, mas as entradas são claras e não há casos de saídas totalmente inesperadas e inexplicáveis.	O aplicativo tem saídas inesperadas, mas elas são revisadas periodicamente até serem compreendidas. A revisão externa e a colaboração são promovidas.	As saídas do aplicativo são inexplicáveis. A revisão é de eficácia limitada para entendê-los.	As entradas não são controladas, o algoritmo não é bem compreendido e as saídas não são compreendidas.
Número de pessoas afetadas	A aplicação não afetará os indivíduos.	O aplicativo será executado apenas dentro de uma organização e afetará um pequeno número de pessoas.	O algoritmo afetará uma pequena comunidade de pessoas bem informadas sobre seu uso.	O aplicativo afetará um grande número de pessoas em todo o país.	O aplicativo afeta um público nacional ou global de pessoas.

Fonte: Adaptado de DISER (2019).

Evidência Express

EVEX **ENAP**