



RELATÓRIO

# AVALIAÇÃO DE IMPACTO ALGORÍTMICO PARA PROTEÇÃO DOS DIREITOS FUNDAMENTAIS

ABRIL  
2023



**LAPIN**

LABORATÓRIO DE POLÍTICAS  
PÚBLICAS E INTERNET

# Laboratório de Políticas Públicas e Internet



## REALIZAÇÃO

Laboratório de Políticas Públicas e Internet - LAPIN

## AUTORIA

Alessandra Lemos

Gabriela Buarque

Ingrid Soares

Victor Mulin

Tayrone Chiavone

## REVISÃO

Cynthia Picolo Gonzaga de Azevedo

José Renato Laranjeira de Pereira

## DIAGRAMAÇÃO

Trivela Publicidade

## COMO CITAR ESTE TRABALHO

LEMOS, Alessandra; BUARQUE, Gabriela; SOARES, Ingrid; MULIN, Victor; CHIAVONE, Tayrone. **Avaliação de Impacto Algorítmico para a proteção dos direitos fundamentais.** Relatório. Brasília: Laboratório de Políticas Públicas e Internet, 2023. Disponível em: [lapin.org.br](https://lapin.org.br). Acesso em: dd mm aaaa.



Este trabalho está licenciado sob uma licença Creative Commons Atribuição-NãoComercial-SemDerivações 4.0 Internacional (CC BY-NC-ND)



# LAPIN

LABORATÓRIO DE POLÍTICAS PÚBLICAS E INTERNET



[lapin.org.br](https://lapin.org.br)



@lapin.br



/lapinbr



/lapinbr

## Quem somos nós

O Laboratório de Políticas Públicas e Internet (LAPIN) é um centro independente de pesquisa e ação de composição multidisciplinar e com sede na capital federal brasileira. Nosso objetivo é apoiar o desenvolvimento de políticas públicas voltadas à regulação das tecnologias digitais por meio da pesquisa, articulação e da conscientização da sociedade.

Nosso trabalho consiste em (i) **investigar, analisar e compreender** os impactos sociais, econômicos, éticos e jurídicos causados pela internet e demais tecnologias digitais; (ii) **informar, incluir e ensinar** o público; e (iii) **propor soluções** inovadoras para os desafios e oportunidades trazidos pela era digital ao Brasil, e demais países latinoamericanos.

Fazemos isso por meio de pesquisas interdisciplinares, desenvolvimento de projetos, ensino, comunicação, e articulação independente voltada para as áreas de regulação, governança e políticas públicas de internet, inovação e tecnologia.

Para maiores informações sobre nossa atuação, visite nosso site: [lapin.org.br](http://lapin.org.br)



|                                                                                                                                      |           |
|--------------------------------------------------------------------------------------------------------------------------------------|-----------|
| <b>LISTA DE ABREVIATURAS E SIGLAS</b>                                                                                                | <b>5</b>  |
| <b>INTRODUÇÃO</b>                                                                                                                    | <b>6</b>  |
| <b>METODOLOGIA</b>                                                                                                                   | <b>12</b> |
| <b>PARTE I:</b> EXEMPLOS DE IMPACTOS NEGATIVOS QUE O USO DA INTELIGÊNCIA ARTIFICIAL PODE CAUSAR À PROTEÇÃO DOS DIREITOS FUNDAMENTAIS | <b>16</b> |
| <b>PARTE II:</b> ENTENDENDO AS AVALIAÇÕES DE IMPACTO ALGORÍTMICO                                                                     | <b>21</b> |
| • Avaliações de Impacto no Brasil                                                                                                    | 21        |
| • Avaliação de Impacto Algorítmico                                                                                                   | 23        |
| • Avaliação de Impacto Algorítmico vs. Relatório de Impacto da Proteção de Dados Pessoais                                            | 25        |
| • Por que realizar uma AIA?                                                                                                          | 27        |
| • Quais sistemas de IA devem ser submetidos a uma AIA?                                                                               | 30        |
| • Quais são os atores associados a sistemas de IA?                                                                                   | 34        |
| • Quem deve realizar uma AIA em cada uma das etapas do ciclo de vida de um sistema de IA?                                            | 37        |
| • Consideração de consulta a terceiros                                                                                               | 39        |
| • Publicação do resultado                                                                                                            | 40        |
| • Benefícios esperados da adoção da AIA pelo setor empresarial                                                                       | 41        |
| <b>PARTE III:</b> CONDUZINDO UMA AVALIAÇÃO DE IMPACTO ALGORÍTMICO                                                                    | <b>42</b> |
| <b>Etapa 1.</b> Descrição detalhada do sistema de IA                                                                                 | <b>43</b> |
| <b>Etapa 2.</b> Avaliação da legalidade da finalidade do sistema de IA                                                               | <b>46</b> |
| <b>Etapa 3.</b> Avaliação da necessidade de participação de terceiros                                                                | <b>47</b> |
| <b>Etapa 4.</b> Identificação e avaliação dos riscos envolvidos                                                                      | <b>48</b> |
| <b>Etapa 5.</b> Definição de salvaguardas, medidas de segurança e medidas de proteção dos direitos e liberdades dos sujeitos.        | <b>52</b> |
| <b>Etapa 6.</b> Avaliação da proporcionalidade entre finalidade e resultado                                                          | <b>53</b> |
| <b>Etapa 7.</b> Riscos residuais e a necessidade de consulta prévia a uma Autoridade de Supervisão                                   | <b>54</b> |
| <b>Etapa 8.</b> Documentação e publicação                                                                                            | <b>55</b> |
| <b>CONCLUSÃO</b>                                                                                                                     | <b>56</b> |
| <b>ANEXO I:</b> MODELO VISUAL DA AIA                                                                                                 | <b>57</b> |



## LISTA DE ABREVIATURAS E SIGLAS

- 1.** AEPD: Agência Espanhola de Proteção de Dados.
- 2.** AIA: Avaliação de Impacto Algorítmico.
- 3.** AIR: Avaliação de Impacto Regulatório.
- 4.** ANPD: Autoridade Nacional de Proteção de Dados Pessoais.
- 5.** CF/88: Constituição Federal de 1988.
- 6.** IA: Inteligência artificial.
- 7.** ICO: Information Commissioner's Office.
- 8.** LGPD: Lei Geral de Proteção de Dados Pessoais.
- 9.** OCDE: Organização para a Cooperação e Desenvolvimento Econômico.
- 10.** ONU: Organização das Nações Unidas.
- 11.** RGPD: Regulamento Geral de Proteção de Dados da União Europeia.
- 12.** RIPD: Relatório de Impacto à Proteção de Dados Pessoais.
- 13.** TRF: Tecnologias de reconhecimento facial.



# Introdução

A inteligência artificial (IA) perpassa diversos setores da economia global e esferas sociais. Além de aprimorar processos de produção e de otimizar produtos e serviços, sistemas de IA tomam cada vez mais espaço no dia a dia das empresas e dos indivíduos, gerando maiores comodidades e aumentando a qualidade de vida das pessoas. No entanto, apesar dos inúmeros e potenciais benefícios que a inteligência artificial pode apresentar, **há uma crescente preocupação com os riscos que tais sistemas podem causar a direitos fundamentais.**

Caso mal desenvolvidos ou operados, sistemas de IA têm o potencial de impactar diversos direitos e liberdades fundamentais, como o direito à igualdade e à não discriminação (art. 5º, caput, CF/88), à livre manifestação de pensamento (art. 5º, IV, CF/88), à liberdade de expressão (art. 5º, IX, CF/88) e de reunião pacífica (art. 5º, XVI, CF/88), o direito de acesso à informação (art. 5º, XXXIII, CF/88), ou mesmo afetar o bom andamento de um sistema democrático.

Diante da grande influência dessa tecnologia em modelos de negócio, a tendência é que empresas e organizações que desenvolvem ou operem sistemas de IA considerados de risco para os direitos de pessoas físicas sejam cada vez mais chamadas para apresentar informações sobre seu funcionamento, de modo a cumprir com os princípios da transparência, da responsabilização e da prestação de contas.

Nesse sentido, a **Avaliação de Impacto Algorítmico (AIA) é vista como um instrumento de governança que possibilita ao desenvolvedor ou aplicador da tecnologia identificar e reduzir possíveis riscos que determinado sistema de IA possa causar aos direitos fundamentais.**



Assim, a proposta deste relatório é, inicialmente, apresentar uma série de importantes conceitos associados à AIA, visando conscientizar e incentivar a sua adoção por empresas e organizações que buscam atuar de forma responsável com a inteligência artificial.

Por não existir ainda um padrão universalmente estabelecido sobre como conduzir uma Avaliação de Impacto Algorítmico, este relatório apresenta uma metodologia que pretende orientar a condução de uma AIA pelos responsáveis por um sistema cujo risco pretende ser mensurado.

Dada a relevância do tema no debate regulatório no Brasil, compete sublinhar que a **Comissão de Juristas** instituída pelo Ato do Presidente do Senado nº 4, de 2022, destinada a subsidiar a elaboração de minuta de substitutivo para instruir a apreciação dos Projetos de Lei nº 5.051/2019, 21/2020, e 872/2021, que têm como objetivo estabelecer princípios, regras, diretrizes e fundamentos para regular o desenvolvimento e a aplicação da inteligência artificial no país, apresentou relatório final com previsões sobre a avaliação de impacto, dedicando uma seção inteira para tal instrumento.<sup>1</sup>

O anteprojeto de lei, em síntese, descreve etapas metodológicas e firma o caráter público da **avaliação de impacto**, que poderá contar com a criação de uma base de dados com os respectivos instrumentos. Ainda, foram estipuladas duas ferramentas de avaliação de impacto, nomeadas como (i) **Avaliação preliminar** e (ii) **Avaliação de impacto algorítmico**.

Em ordem cronológica, o **art. 13** da minuta dispõe que todo sistema de inteligência artificial passará por **avaliação preliminar** realizada pelo fornecedor para classificação de seu grau de risco, previamente à sua colocação no mercado ou utilização em serviço, indicando as finalidades ou aplicações indicadas. Ademais, prevê, no parágrafo terceiro, que a autoridade competente poderá determinar a reclassificação do sistema de IA, mediante notificação prévia, bem como determinar a realização de avaliação de impacto algorítmico para instrução da investigação em curso.

<sup>1</sup> Relatório Legislativo, Relatório Final aprovado em 06 de dezembro de 2022. Disponível em: <https://legis.senado.leg.br/comissoes/comissao?codcol=2504>. Acesso em: 23 jan 2023.



O parágrafo quarto prossegue aduzindo que se o resultado da reclassificação identificar o sistema de inteligência artificial como de alto risco, a realização de avaliação de impacto algorítmico e a adoção de medidas de governança serão obrigatórias, sem prejuízo de eventuais penalidades em caso de avaliação preliminar fraudulenta, incompleta ou inverídica.

A necessidade da realização de **avaliação de impacto algorítmico** vem estampada no **art. 22**, que a caracteriza como obrigação dos agentes de inteligência artificial sempre que o sistema for considerado como de alto risco pela avaliação preliminar. Na sequência, o **art. 23** estipula que a avaliação será realizada por profissional ou equipe de profissionais com conhecimentos técnicos, científicos e jurídicos necessários para realização do relatório com independência funcional, cabendo à autoridade competente regulamentar os casos em que a realização será necessariamente conduzida por profissional ou equipe externas ao fornecedor.

O parágrafo primeiro do **art. 24** segue aduzindo que a avaliação de impacto considerará e registrará, ao menos: **(a)** os riscos conhecidos e previsíveis associados ao sistema de inteligência artificial à época em que foi desenvolvido, bem como os riscos que podem razoavelmente dele se esperar; **(b)** benefícios associados ao sistema de inteligência artificial; **(c)** probabilidade de consequências adversas, incluindo o número de pessoas potencialmente impactadas; **(d)** gravidade das consequências adversas, incluindo o esforço necessário para



mitigá-las; **(e)** lógica de funcionamento do sistema de inteligência artificial; **(f)** processo e resultado de testes e avaliações e medidas de mitigação realizadas para verificação de possíveis impactos a direitos, com especial destaque para potenciais impactos discriminatórios; **(g)** treinamento e ações de conscientização dos riscos associados ao sistema de inteligência artificial; **(h)** medidas de mitigação e indicação e justificação do risco residual do sistema de inteligência artificial, acompanhado de testes de controle de qualidade frequentes; **(i)** medidas de transparência ao público, especialmente aos potenciais usuários do sistema, a respeito dos riscos residuais, principalmente quando envolver alto grau de nocividade ou periculosidade à saúde ou segurança dos usuários, nos termos dos artigos 9º e 10 da Lei nº 8.078, de 11 de setembro de 1990 (Código de Defesa do Consumidor).

O parágrafo segundo determina, ainda, que em atenção ao princípio da precaução, quando da utilização de sistemas de inteligência artificial que possam gerar impactos irreversíveis ou de difícil reversão, a avaliação de impacto algorítmico levará em consideração também as evidências incipientes, incompletas ou especulativas.



Em continuidade, os parágrafos terceiros e quarto estabelecem que a autoridade competente poderá estabelecer outros critérios e elementos para a elaboração da avaliação de impacto, incluindo a participação dos diferentes segmentos sociais afetados, conforme risco e porte econômico da organização, bem como a regulamentação da periodicidade de atualização, considerando o ciclo de vida dos sistemas de alto risco e os campos de aplicação, podendo incorporar melhores práticas setoriais.

Desta forma, foram criadas duas “camadas” avaliativas. A primeira delas (avaliação preliminar) a ser elaborada pelo próprio fornecedor da tecnologia, visando classificar o seu grau de risco, podendo esta ser reclassificada pelo regulador. A segunda camada trata-se da avaliação de impacto algorítmico, mais robusta e obrigatória para sistemas considerados de alto risco.

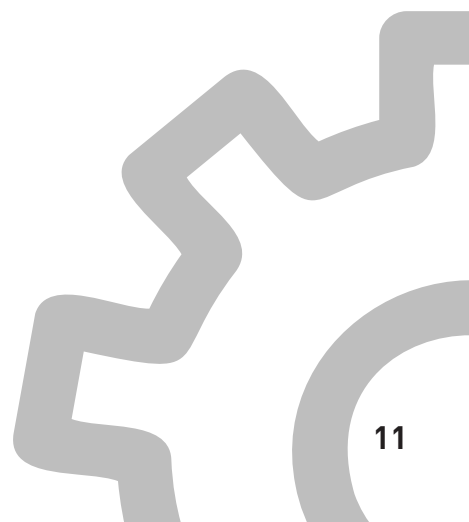
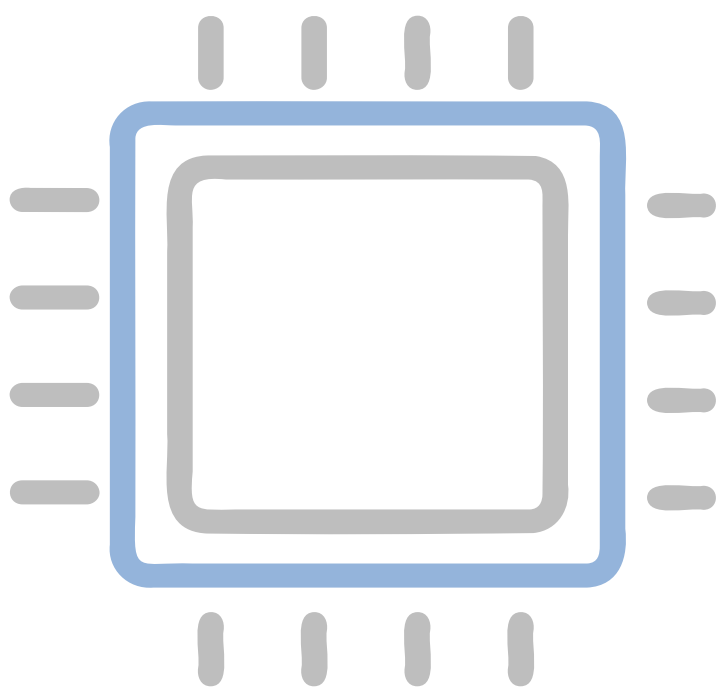
A minuta do substitutivo segue para análise do Senado Federal, podendo ainda sofrer acréscimos, modificações ou supressões. Desse modo, sem pretensão de esgotar o tema, **o presente relatório visa discutir a proposta da avaliação de impacto algorítmico a partir do atual estado da arte, pontuando que a dinamicidade do tema impõe sua revisitação e reflexão constantes.**



O presente relatório foi dividido em três partes. A **primeira** é reservada a **contextualizar e apresentar exemplos de riscos** que o uso da inteligência artificial pode causar à proteção dos direitos fundamentais. Destaca-se que sistemas de IA tratam dados e informações de carácter pessoal e não pessoal, e que, consequentemente, os impactos aos direitos fundamentais não se limitam apenas a questões de privacidade e proteção de dados pessoais.

A **segunda parte** discorre sobre um conjunto de **conceitos-chave associados à AIA**. Os conceitos escolhidos objetivam criar um conhecimento sólido a respeito de questões procedimentais, como **(a)** o que é uma AIA; **(b)** quais tipos de aplicações devem ser analisadas; **(c)** quem e em quais etapas deve-se conduzir uma AIA; **(d)** se é necessária a participação de terceiros externos; e, **(e)** como pode ser feita a publicação do resultado da avaliação. Além de questões procedimentais, também apresentamos justificativas e benefícios para convencer e conscientizar sobre a importância da AIA.

A **terceira e última parte** apresenta e explica um conjunto de **oito etapas** concretas que o responsável pelo sistema de IA deve seguir para realizar uma AIA. Estas etapas foram traduzidas em um modelo visual da AIA, apresentado no **Anexo I**.



Para a primeira e segunda parte deste relatório - riscos da IA aos direitos fundamentais e conceitos-chaves associados à AIA -, a metodologia usada foi a de **revisão de literatura** publicada no domínio da ética e regulação de inteligência artificial. As fontes de pesquisa foram extraídas de trabalhos de:

- Atores governamentais;
- Autoridades de controle e supervisão da área de proteção de dados pessoais;
- Organizações internacionais, como Organização das Nações Unidas (ONU) e Organização para a Cooperação e Desenvolvimento Económico (OCDE);
- *Think Tanks* e organizações da sociedade civil;
- Academia.

Por outro lado, a terceira parte, que explica como conduzir uma AIA, foi desenvolvida a partir da análise de ao menos **três propostas de avaliações de impacto para sistemas de IA**.

A primeira proposta que serviu de inspiração adveio das discussões da Autoridade de Proteção de Dados do Reino Unido, **Information Commissioner's Office (ICO)**, que parte da realização de um Relatório de Impacto à Proteção de Dados Pessoais (RIPD),<sup>2</sup> instrumento previsto nas leis de proteção de dados brasileira e europeia para a mensuração de risco para atividades de tratamento de dados pessoais, para avaliar os riscos provenientes de sistemas de IA.<sup>3</sup> Essa é uma linha de raciocínio que a Agência Espanhola de Proteção de Dados (AEPD) também defende.<sup>4</sup>

<sup>2</sup> A nomenclatura europeia para este instrumento é "Avaliação de Impacto da Proteção de Dados Pessoais", definida no artigo 35 do Regulamento Geral sobre a Proteção de Dados da União Europeia (RGPD). Contudo, por uma questão de facilitar a leitura e o entendimento, optamos por utilizar a nomenclatura encontrada na Lei Geral de Proteção de Dados Pessoais (LGPD).

<sup>3</sup> Information Commissioner's Office (ICO). Guidance on AI and Data Protection. jul 2020, versão 0.0.22. Disponível em: <https://ico.org.uk/for-organisations/guide-to-data-protection/key-dp-themes/guidance-on-artificial-intelligence-and-data-protection/>; Information Commissioner's Office (ICO). **What are the accountability and governance implications of AI?** Disponível em: <https://ico.org.uk/for-organisations/guide-to-data-protection/key-dp-themes/guidance-on-ai-and-data-protection/what-are-the-accountability-and-governance-implications-of-ai/>. Acesso em: 31 jul 2021.

<sup>4</sup> Agencia Española de Protección de Datos (AEPD). **Adecuación al RGPD de tratamientos que incorporan inteligencia artificial**. 2020. Disponível em: <https://www.aepd.es/conocimiento/guia/adequacion-al-rgpd-de-tratamientos-que-incorporan-inteligencia-artificial-una-introduccion/>. Idem. **Audit requirements for personal data processing activities involving AI**. 2021. Disponível em: <https://www.aepd.es/sites/default/files/2021-01/requisitos-auditorias-tratamientos-incluyan-ia-en.pdf>. Acesso em: 31 jul 2021.



Além disso, é possível encontrar inúmeros artigos acadêmicos que igualmente apoiam a adoção de RIPD para promover sistemas de IA que respeitem a equidade e a não discriminação.<sup>5</sup>

A segunda proposta é de autoria da **ECP | Platform for the Information Society**,<sup>6</sup> que publicou em 2018 as orientações denominadas “*Artificial Intelligence Impact Assessment - Roadmap for conducting the AIIA*”.<sup>7</sup> Essa proposta de avaliação de impacto foi desenvolvida a partir das Diretrizes para regras de conduta de sistemas autônomos (“*Handreiking Voor GedragsRegels AutoNome Systemen*”), publicadas pela própria ECP ainda em 2006.<sup>8</sup>

<sup>5</sup>JANSSEN, Heleen. Detecting new approaches for a fundamental rights impact assessment to automated decision-making. **International Data Privacy Law**, Oxford, v. 10, n. 1, p. 76-106, fevereiro 2020. Disponível em: <https://doi.org/10.1093/idpl/ipz028>. Acesso em: 31 jul 2021; KAMINSKI, Margot E.; MALGIERI, Gianclaudio Algorithmic impact assessments under the GDPR: producing multi-layered explanations. **International Data Privacy Law**, Oxford, v. 11, n. 2, abril 2021, p. 125-144. Disponível em: <https://academic.oup.com/idpl/article/11/2/125/6024963>. Acesso em: 31 jul 2021; MANTELERO, Alessandro; ESPOSITO, Maria Samantha. An evidence-based methodology for human rights impact assessment (HRIA) in the Development of AI Data-Intensive Systems. **The Computer Law and Security Report**, Reino Unido, v. 41, jul 2021. Disponível em: [https://www.sciencedirect.com/science/article/pii/S0267364921000340#cit\\_12](https://www.sciencedirect.com/science/article/pii/S0267364921000340#cit_12). Acesso em: 31 jul 2021; IVANOVA, Yordanka. The data protection impact assessment as a tool to enforce non-discriminatory AI. In: SCHWEIGHOFER, Erich et al (EE.). **Privacy technologies and policy**. Charm, Springer International, 2020; METCALF, Jacob et al. Algorithmic impact assessments and accountability: the co-construction of impacts. In: ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY (FAccT '21), 2021, evento online. Proceedings.

<sup>6</sup>ECP, ou Electronic Commerce Platform, é uma plataforma independente situada na Holanda em que o governo, a comunidade empresarial e as organizações da sociedade civil se reúnem para cooperar e trocar conhecimento para explorar as oportunidades em um contexto de sociedade da informação, bem como para mitigar as ameaças colocadas pelo mundo digital. Disponível em: <https://ecp.nl/>. Acesso em 31 jul 2021.

<sup>7</sup>ECP | Platform for the Information Society. **Artificial Intelligence Impact Assessment - Roadmap for conducting the AIIA**. 2019. Disponível em: <https://futurium.ec.europa.eu/en/european-ai-alliance/best-practices/ai-impact-assessment-code-conduct>. Acesso em: 31 jul 2021.

<sup>8</sup>SCHERMER, Bart W. **Handreiking voor gedragsregels autonome systemen**. ECP.NL, 2006. Relatório técnico. Disponível em: <https://ecp.nl/wp-content/uploads/2017/04/Handreiking-voor-gedragsregels-autonome-systemen-2006.pdf>. Acesso em: 31 jul 2021.

Para trazer uma perspectiva do olhar do setor empresarial sobre AIA, a terceira proposta analisada foi a do **Open Loop**, um programa apoiado pelo Facebook AI, denominada “*AI Impact Assessment: A Policy Prototyping Experiment*”.<sup>9</sup> Este relatório apresenta os resultados do programa de prototipagem de leis realizado com dez organizações europeias a respeito das avaliações de impacto de decisão automatizada.

Todas as propostas influenciaram na elaboração da metodologia de execução da AIA apresentada pelo LAPIN, constituída por oito etapas.

A **etapa 1**, referente à **descrição detalhada do sistema de IA**, é algo presente em todas as propostas analisadas, notadamente sobre a indicação do contexto, da finalidade, da lógica de aplicação do sistema e do tipo de dado utilizado para basear seu funcionamento. Além disso, incluiu-se na etapa 1 a descrição dos benefícios da aplicação de IA, que é apresentada como uma etapa isolada na proposta da ECP. Já a **etapa 2**, em que se **avalia a legalidade da finalidade e da lógica de operação do sistema de IA**, e que foi uma inspiração retirada nos estudos da ECP.

A **etapa 3**, sobre a avaliação da **necessidade de participação de terceiros**, foi inspirada na própria metodologia do RIPD.<sup>10</sup> Considerando que alguns usos de sistemas de IA podem ter uma repercussão a nível social, como, por exemplo, um sistema de moderação de conteúdo online ou um sistema de reconhecimento facial, a ideia de permitir maior abertura da AIA para a sociedade revelou-se um importante aspecto na construção da metodologia deste instrumento.

Por sua vez, a **etapa 4**, que se refere à **identificação e avaliação dos riscos**, foi baseada no relatório do Open Loop, que pretende que a avaliação de impacto seja conduzida de modo a identificar e reduzir riscos que possam afetar negativamente pessoas físicas. Essa é uma ampliação do objeto de proteção se comparada às outras propostas, que se limitam a proteger tão somente os direitos e interesses das pessoas físicas.

<sup>9</sup> ANDRADE, Norberto Nuno Gomes de; KONTSCIEDER, Verena. **AI impact assessment: a policy prototyping experimente**. Open Loop, 2021, p. 30, 35-36 e 75-76. Relatório técnico. Disponível em: [https://openloop.org/wp-content/uploads/2021/01/AI\\_Impact\\_Assessment\\_A\\_Policy\\_Prototyping\\_Experiment.pdf](https://openloop.org/wp-content/uploads/2021/01/AI_Impact_Assessment_A_Policy_Prototyping_Experiment.pdf). Acesso em: 31 jul 2021.

<sup>10</sup> O relatório de impacto no Regulamento Geral Europeu de Proteção de Dados Pessoais é previsto no art. 35. Já na LGPD, tal previsão é encontrada em seu art. 38, além das menções nos art. 4º, parágrafo terceiro, art. 5º, XVII e art. 10, parágrafo terceiro.

A **etapa 5**, que define **salvaguardas, medidas de segurança e de proteção dos direitos e liberdades** dos sujeitos afetados, também é algo comum em todas as propostas. Já a **etapa 6**, em que se avalia a **proporcionalidade entre finalidade e resultado**, é uma etapa característica do RIPD definida pelo art. 35 do Regulamento Geral de Proteção de Dados europeu.

Na **etapa 7**, a análise de **riscos residuais** está presente tanto nas discussões sobre o RIPD para sistemas de IA quanto na proposta da ECP. Contudo, a consulta prévia a uma autoridade supervisora setorial quando há ainda um alto risco residual é premissa que foi inspirada diretamente no RGPD.

Por fim, a **etapa 8** aborda tanto a questão da **documentação da AIA**, que é algo inerente ao próprio processo de avaliação, quanto sobre sua **publicação**. A sugestão sobre a deliberação da publicação de um resumo da AIA advém das discussões acadêmicas sobre a RIPD para sistemas de IA.

Ressalte-se que o texto apresentado pela **Comissão de Juristas** dispõe, no art. 24, que a metodologia da avaliação de impacto conterà, ao menos, as etapas de: I - preparação; II - cognição do risco; III - mitigação dos riscos encontrados; IV - monitoramento. Verifica-se, assim, que as etapas ora sugeridas estão em conformidade com a previsão do anteprojeto de lei, em que pese utilizem nomenclaturas diversas.

Isso porque as etapas 1, 2 e 3 convergem com a etapa de preparação; a etapa 4 equivale à cognição do risco; as etapas 5, 6 e 7 convergem com a de mitigação dos riscos encontrados; e a etapa 8 alia-se à constante necessidade de controle social e etapa de monitoramento. Destaca-se, ainda, que as etapas previstas na minuta substitutivo são mínimas, sendo possível acrescentar outras medidas de segurança que possam avaliar os riscos com mais robustez. Com efeito, **no presente relatório optou-se por detalhar as fases de modo mais pormenorizado**, com vistas a facilitar a compreensão pelos atores de IA e permitir maior segurança na condução do procedimento.

Importa ressaltar que a literatura científica está em fase de desenvolvimento. Nesse contexto, o presente relatório visa fomentar a discussão sobre o tema no Brasil, baseando-se em documentos que tratam a questão por meio de revisão bibliográfica e buscando contextualizar criticamente a proposição da avaliação de impacto



## PARTE I: EXEMPLOS DE IMPACTOS NEGATIVOS QUE O USO DA INTELIGÊNCIA ARTIFICIAL PODE CAUSAR À PROTEÇÃO DOS DIREITOS FUNDAMENTAIS

Se, por um lado, é inegável o aspecto benéfico do desenvolvimento e utilização de tecnologias de inteligência artificial, por outro, é necessário estar atento aos possíveis impactos negativos que elas podem ter na proteção e fruição de direitos fundamentais.

Considerando a característica prototipal dos interesses apontados - que remontam aos direitos de intimidade, liberdade e igualdade - optamos por destacar seis direitos fundamentais que podem ser colocados em risco. Assim, pode-se apontar os **direitos à privacidade** (art. 5º, X, CF/88), **à proteção de dados pessoais** (art. 5º, LXXIX, CF/88), à **liberdade de expressão** (art. 5º, IX, CF/88), à **liberdade de reunião pacífica** (art. 5º, XVI, CF/88), à igualdade e à não discriminação (art. 5º, caput, CF/88), ao **direito de petição** (art. 5º, XXXIV, CF/88) e ao **devido processo legal** (art. 5º, inciso LIV, CF/88).

O **risco à privacidade e à proteção de dados pessoais** reside no potencial que as tecnologias de inteligência artificial têm de coletar, processar e armazenar dados em uma escala gigantesca tal como ocorre, por exemplo, nas tecnologias de reconhecimento facial (TRF). Um dos perigos apontados é a possibilidade de países e empresas utilizarem estas tecnologias para monitorar, analisar, prever e até mesmo manipular o comportamento de seus cidadãos ou clientes.<sup>11</sup>

<sup>11</sup> ORGANIZAÇÃO DAS NAÇÕES UNIDAS. Alto Comissariado de Direitos Humanos. **O direito à privacidade na era digital**. A/HRC/39/29 (03 de agosto de 2018, p. 2.). Disponível em: <https://undocs.org/A/HRC/39/29>. Acesso em: 31 jul 2021.



**Outro temor é que haja troca de dados e informações coletadas através de inteligência artificial entre Estados e empresas** para usos secundários à finalidade a que foram coletados, sem mecanismos para que o titular de dados possa descobrir quem os detém e como vêm sendo utilizados (transparência), e sem uma análise aprofundada da proporcionalidade entre a importância e aplicação dos dados transferidos e a proteção dos direitos humanos.<sup>12</sup> Em se tratando de dados biométricos,<sup>13</sup> os riscos de possíveis violações aumentam, assim como a sua potencial gravidade, uma vez que estes dados possuem uma ligação permanente com pessoas em específico,<sup>14</sup> sofrendo poucas alterações ao longo do tempo.

Uma vez extraído um dado biométrico e estabelecida uma conexão a uma pessoa, esta poderá ser facilmente identificada e dificilmente conseguirá transitar anonimamente em certos espaços, o que pode afetar a sua vida privada e o seu comportamento em público.

**Outro grande direito em risco é a liberdade de expressão.** Inteligências artificiais utilizadas para moderação de conteúdo online, se utilizadas sem a devida diligência em relação aos direitos fundamentais, podem se transformar em instrumentos de censura,<sup>15</sup> tanto pela iniciativa privada quanto pelo poder público.

Mais um aspecto da violação da liberdade de expressão, e também da liberdade de informação, relaciona-se com o modo como as ferramentas de busca online e redes sociais são desenhadas. Tais sistemas definem as informações a que os usuários terão acesso de acordo com prévios interesses e buscas realizadas e, conseqüentemente, acabam selecionando quais conteúdos serão acessados e por quem. Segundo a organização de sociedade civil Access Now, os algoritmos por trás de tais sistemas têm sido responsáveis por criar câmaras de eco (*echo chambers*), espaços em que se propaga apenas uma mesma ideia ou perspectiva, colocando em risco a pluralidade de ideias e a diversidade de opiniões no ambiente online.<sup>17</sup>

<sup>12</sup> *Ibidem*. p. 4-5.

<sup>13</sup> A LGPD não traz definição exata para dados biométricos. Desse modo, utilizaremos a definição do art. 4º [14] do RGPD, o qual indica dados biométricos como aqueles dados pessoais resultantes de processamento técnico específico relacionado às características físicas, fisiológicas ou comportamentais dos indivíduos.

<sup>14</sup> ORGANIZAÇÃO DAS NAÇÕES UNIDAS, *Op. cit.*, p. 5.

<sup>15</sup> ACCESS NOW. **Human rights in the age of human intelligence**. 2018, p. 22. Disponível em: <https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf>. Acesso em: 31 jul 2021.

<sup>16</sup> LATONERO, Mark. **Governing artificial intelligence: upholding human rights & dignity**. Data&Society, 2018, p. 14-15. Disponível em: [https://datasociety.net/wp-content/uploads/2018/10/DataSociety\\_Governing\\_Artificial\\_Intelligence\\_Upholding\\_Human\\_Rights.pdf](https://datasociety.net/wp-content/uploads/2018/10/DataSociety_Governing_Artificial_Intelligence_Upholding_Human_Rights.pdf). Acesso em: 31 jul 2021.

<sup>17</sup> ACCESS NOW. *Op. cit.*, p. 23.

A utilização de tecnologias de IA para moderação de conteúdo online também pode implicar **violações à liberdade de reunião pacífica**. Redes sociais e sítios eletrônicos são meios utilizados por movimentos sociais, grupos e coletivos para organizar reuniões e protestos. A inteligência artificial pode ser usada para retirar do ar total ou parcialmente grupos, páginas ou conteúdos direcionados a organizar encontros e colaborações de grupos.<sup>18</sup>

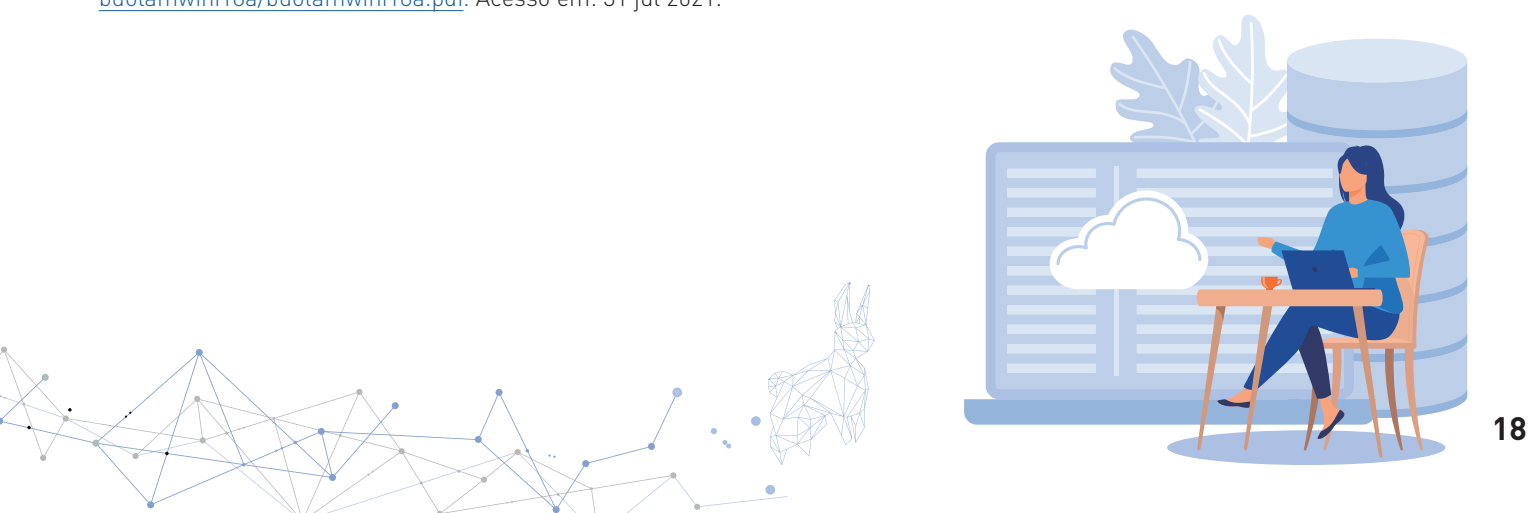
Além disso, tem sido igualmente debatido como o uso de tecnologias de reconhecimento facial também pode promover uma restrição indevida do direito à reunião pacífica. O uso deste tipo de tecnologia em contextos de protestos ou reuniões também pode ter um “**chilling effect**”, um efeito inibitório.<sup>19</sup> Isto é, pessoas podem se sentir menos confortáveis em participar de protestos e reuniões se souberem que poderão ser identificadas. Esse desconforto pode levá-las a deixarem de se reunirem, limitando o exercício da liberdade de reunião e provocando uma redução do espaço de discussão democrática.

A controvérsia a respeito do uso de tecnologias de reconhecimento facial não se restringe ao impacto negativo à liberdade de reunião pacífica. Estas tecnologias vêm sendo discutidas principalmente pelo seu **potencial discriminatório**. Joy Buolamwini e Timnit Gebru demonstraram em 2018 que, apesar de as TRF desenvolvidas pela IBM, Microsoft e Face++ possuírem taxas de acerto expressivas, elas performaram diferente em relação a distintos grupos demográficos. As pesquisadoras identificaram que todas as tecnologias possuíam uma taxa de acerto maior para homens brancos e performavam com menos acertos quando as pessoas avaliadas eram negras e do sexo feminino.<sup>20</sup>

<sup>18</sup> *Ibidem*.

<sup>19</sup> *Ibidem*.

<sup>20</sup> BUOLAMWINI, Joy; GEBRU, Timnit. **Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification**. Proceedings of Machine Learning Research, Nova York, v. 8 p. 1-15, fevereiro, 2018. Disponível em: <http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>. Acesso em: 31 jul 2021.



As inteligências artificiais utilizadas para processos de recrutamento de empregados também estão sob a mira de críticos por viabilizarem, consciente e inconscientemente, práticas discriminatórias.

Em 2018, a Amazon descobriu que a IA que adotava para fazer uma triagem inicial dos candidatos era discriminatória contra mulheres. Os dados utilizados para treinar a inteligência artificial fizeram com que ela desse preferência a candidatos homens e avaliasse negativamente currículos que continham expressões que relacionavam uma atividade ou qualificação a uma mulher, como “*women’s*” (de mulheres) e “*women’s chess club captain*” (capitã do clube feminino de xadrez).<sup>21</sup> Por conta desse viés presente na tecnologia, a Amazon optou por deixar de utilizá-la em seus processos seletivos.

Ainda, no Brasil, diante das evidências de uso abusivo e pouco transparente, foi lançada a campanha “**Tire Meu Rosto da Sua Mira**”,<sup>22</sup> uma mobilização da sociedade civil que requer o banimento das tecnologias de reconhecimento facial na segurança pública brasileira, sob o fundamento de que o reconhecimento nesse contexto viola os direitos de privacidade, proteção de dados, liberdade de expressão e reunião, igualdade e não-discriminação.

Também ressalta-se o risco ao **direito de petição**, estipulado no art. 5º, XXXIV, da Constituição Federal de 1988, que concretiza o direito de petição aos Poderes Públicos em defesa de direitos ou contra ilegalidade ou abuso de poder, independentemente do pagamento de taxas. O risco se caracteriza a partir do comprometimento do eficaz direito de petição por algoritmos não transparentes que dificultam o exercício do direito de revisão de decisões automatizadas, por exemplo.

<sup>21</sup> REUTERS. **Amazon ditched AI recruiting tool that favored men for technical jobs. The Guardian.** 11 out 2018 Disponível em: <https://www.theguardian.com/technology/2018/oct/10/amazon-hiring-ai-gender-bias-recruiting-engine>. Acesso em: 31 jul 2021.

<sup>22</sup> TIRE MEU ROSTO DA SUA MIRA. Disponível em: <https://tiremeurostodasuamira.org.br/>. Acesso em: 07 fev 2023.

Por fim, menciona-se o risco ao **devido processo legal**, estipulado no art. 5º, LIV, da Constituição Federal de 1988, caracterizado especificamente como devido processo informacional. Esse risco resta configurado quando o direito de ampla defesa e contraditório a quem possa ter direitos e liberdades afetados por uma decisão totalmente ou parcialmente automatizada não é garantido em sistemas de inteligência artificial.

Vislumbra-se, então, que o uso de tecnologias de inteligência artificial pode ter sérias implicações no gozo de direitos fundamentais. É nesse contexto, portanto, que **as avaliações de impacto algorítmico surgem como um dos instrumentos mais importantes para promover uma cultura de desenvolvimento tecnológico que garanta o respeito a esses direitos**. A seção a seguir é destinada a entender o que é esse instrumento, assim como suas características principais.

## PARTE II: ENTENDENDO AS AVALIAÇÕES DE IMPACTO ALGORÍTMICO

### Avaliações de Impacto no Brasil

Avaliações, relatórios e diagnósticos do impacto de determinada atividade ou produto são instrumentos que ganham cada vez mais importância em uma sociedade na qual as ações humanas e empresariais podem provocar riscos de difícil ou impossível reparação. **Através da condução de uma avaliação do impacto, é possível identificar e prever possíveis riscos que uma determinada atividade pode causar a direitos e bens protegidos pelo nosso ordenamento jurídico.** Em posse dessas informações, torna-se possível estudar a possibilidade de implementar salvaguardas e medidas de mitigação de riscos, o que permite mitigar ou eliminar os impactos deletérios dessa atividade a um nível aceitável.

Dentre as diversas avaliações de impacto existentes, uma das que mais tem se destacado no Brasil está regulada pela Lei Geral de Proteção de Dados Pessoais (Lei n. 13.709/2018, LGPD).<sup>23</sup> Nomeada como **Relatório de Impacto à Proteção de Dados Pessoais** (RIPD),<sup>24</sup> essa avaliação consiste na análise de processos de tratamento de dados pessoais que podem gerar riscos às liberdades civis e aos direitos fundamentais, tendo ainda como objetivo analisar as salvaguardas e os mecanismos de mitigação de risco (art. 5º, XVII, LGPD).

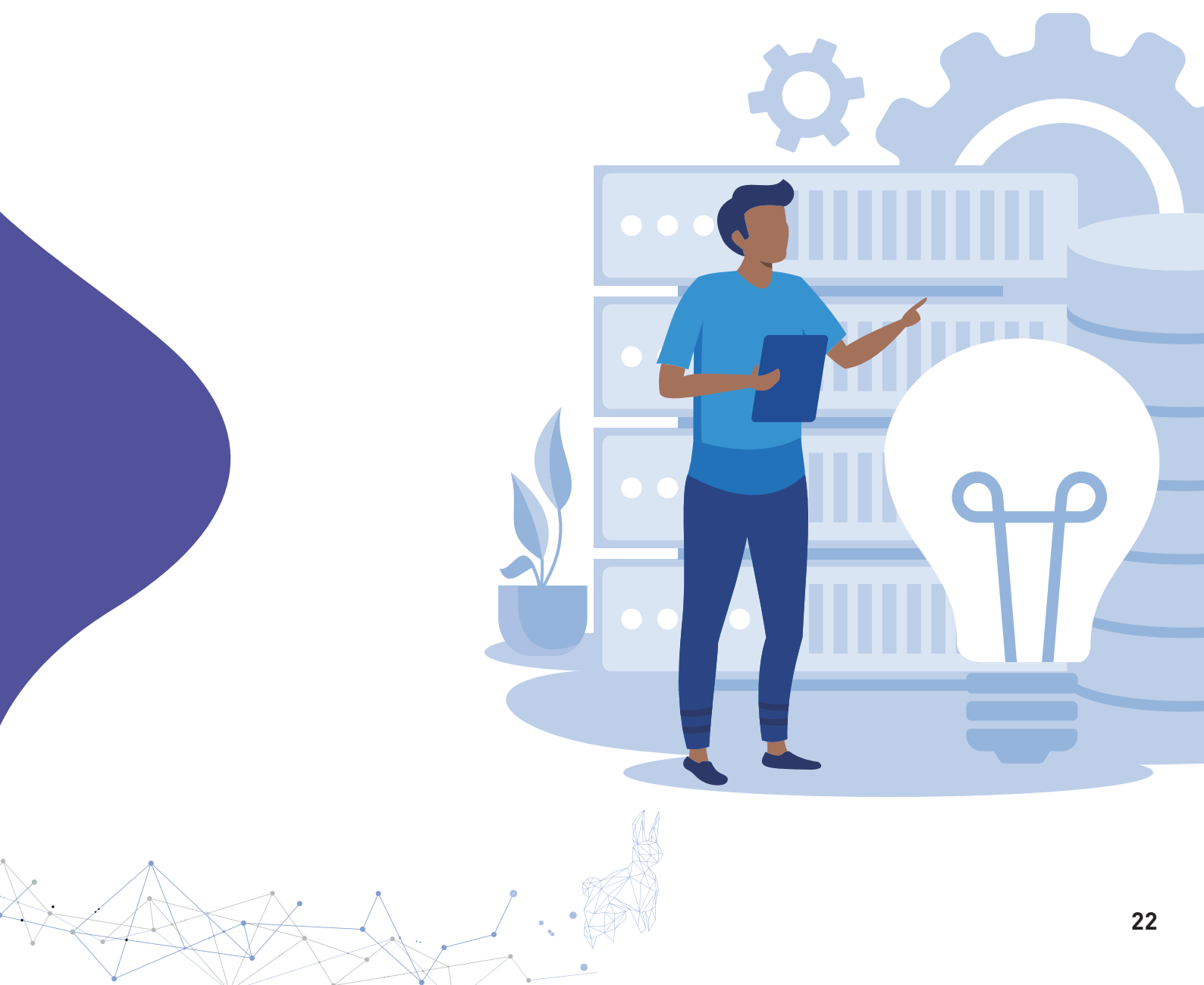
Contudo, esse não é o primeiro instrumento de avaliação de impacto que o Brasil adotou. Em 1981, a Lei nº 6.938 definiu a **Avaliação de Impacto Ambiental** como um dos instrumentos de proteção relativos à Política Nacional de Meio Ambiente (art. 9º, III). Posteriormente, a própria Constituição Federal de 1988 consolidou o instrumento de avaliação de impacto voltado para o meio ambiente em seu art. 225, §1º, inciso IV, que determina que o Poder Público exigirá a realização de um estudo prévio de impacto ambiental antes que se inicie uma “obra ou atividade potencialmente causadora de significativa degradação do meio ambiente”.

<sup>23</sup> BRASIL. Lei nº 13.709/2018. **Lei Geral de Proteção de Dados Pessoais**. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/ato2015-2018/2018/lei/l13709.htm](http://www.planalto.gov.br/ccivil_03/ato2015-2018/2018/lei/l13709.htm). Acesso em: 31 jul 2021.

<sup>24</sup> Importante destacar que o RIPD ainda configura como objeto de futura regulamentação pela ANPD, tendo havido tomada de subsídios, mas sem previsão de publicação da minuta da resolução para consulta pública.

Por fim, recentemente o Brasil também adotou a **Avaliação de Impacto Regulatório (AIR)**, que foi instituída pelo art. 5º da Lei de Liberdade Econômica (Lei nº 13.874/2019), pelo art. 6º da Lei Geral das Agências Reguladoras Federais (Lei nº 13.848/2019), e regulamentada por meio do Decreto nº 10.411/2020. A AIR é um instrumento de gestão e implementação de políticas públicas<sup>25</sup> que tem como objetivo realizar uma avaliação prévia à edição de atos normativos para verificar a razoabilidade do seu impacto, bem como subsidiar a tomada de decisão (art. 2º, I, do Decreto nº 10.411/2020).

<sup>25</sup> BRASIL. Ministério da Economia. Normativos da União deverão ter análise de impacto regulatório para serem editados. Disponível em: <https://www.gov.br/economia/pt-br/assuntos/noticias/2020/julho/normativos-da-uniao-deverao-ter-analise-de-impacto-regulatorio-para-serem-editados>. Acesso em: 31 jul 2021.



## Avaliação de Impacto Algorítmico

Em 2018, o instituto de pesquisa estadunidense *AI Now*<sup>26</sup> publicou uma das primeiras produções literárias relativas a uma avaliação de impacto voltada para sistemas de inteligência artificial conhecida como “*Algorithmic Impact Assessment in the Public Sector*”.<sup>27</sup> O propósito inicial da *AI Now* residia na **criação de uma metodologia de avaliação de impacto que pudesse analisar e assegurar a prestação de contas (accountability) das aplicações algorítmicas** utilizadas por agências governamentais nos Estados Unidos da América.

Um dos elementos-chave da avaliação de impacto algorítmico proposta pelo instituto consiste no dever das agências públicas avaliarem os potenciais impactos das aplicações algorítmicas no tocante à **equidade, justiça, enviesamento, devido processo legal** e outras preocupações das comunidades afetadas.<sup>28</sup> Essa ideia, que inicialmente foi desenvolvida para aplicação no setor público, foi rapidamente encampada pelo meio acadêmico, que começou a sugerir estudos e propostas incentivando a condução de avaliações de impacto para sistemas de IA também pelo setor privado. Contudo, as produções literárias sobre essa avaliação de impacto não seguiram uma nomenclatura única, o que passou a provocar certa confusão ao analisar diversos documentos.

Para ilustrar, citamos **ao menos três nomenclaturas** utilizadas para referir-se a avaliações de impacto voltadas para sistemas de IA. **A primeira** delas consiste na **Avaliação de Impacto Algorítmico**, nome este que foi proposto pela *AI Now* e é também utilizado em uma ferramenta desenvolvida pelo Governo do Canadá.<sup>29</sup> Ressalta-se que a minuta do substitutivo proposta pela Comissão de Juristas adotou esta mesma nomenclatura.

<sup>26</sup> *AI Now Institute* é um instituto de pesquisa multidisciplinar que atua para ajudar a garantir que os sistemas IA sejam responsáveis nos contextos em que são aplicados.

<sup>27</sup> REISMAN, Dillon; *et al.* **Algorithmic Impact Assessment in the Public Sector**. *AI Now*, 2018. Relatório técnico. Disponível em: [https://ec.europa.eu/futurium/en/system/files/ged/ai\\_now\\_2018\\_report.pdf](https://ec.europa.eu/futurium/en/system/files/ged/ai_now_2018_report.pdf). Acesso em: 31 jul 2021.

<sup>28</sup> *Ibidem*, p. 15-16.

<sup>29</sup> CANADÁ. **Algorithmic Impact Assessment Tool**. Disponível em: <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html#toc3-1>. Acesso em: 05 mai 2022.



A **segunda** nomenclatura é a chamada **Avaliação de Impacto para Tomada de Decisões Automatizadas**, que foi apresentada em um artigo científico de autoria de Hellen L. Janssen, acadêmica das Universidades de Cambridge e Amsterdam.<sup>30</sup> Por fim, a terceira nomenclatura é conhecida como **Avaliação de Impacto de Inteligência Artificial**, que foi apresentada em 2018 pela *ECP | Platform for the Information Society*.<sup>31</sup>

Apesar de diferentes terminologias, de modo geral, todas as propostas possuem objetivos semelhantes, sendo instrumentos de governança que oferecem diretrizes para a avaliação, documentação e prestação de contas de determinados sistemas de IA.<sup>32</sup> Com isso, visam permitir a identificação e definição de salvaguardas para mitigar riscos associados ao desenvolvimento ou operação destes sistemas,<sup>33</sup> o que contribui para consolidar a IA de forma responsável e protetiva a direitos.<sup>34</sup> A avaliação de impacto visa impedir que o uso prematuro de um algoritmo acarrete consequências indesejadas como imprecisão, ineficácia ou violação de direitos.<sup>35</sup>

**Este relatório optou-se por utilizar a terminologia Avaliação de Impacto Algorítmico (AIA), para fins de harmonização com a nomenclatura utilizada no substitutivo proposto pela Comissão de Juristas.** Dessa forma, além de consolidar uma única nomenclatura, também evitam-se confusões com outros instrumentos.<sup>36</sup>

<sup>30</sup> JANSSEN, Heleen. Detecting new approaches for a fundamental rights impact assessment to automated decision-making. **International Data Privacy Law**, Oxford, v. 10, n. 1, p. 76-106, fevereiro 2020. Disponível em: <https://doi.org/10.1093/idpl/ipy028>. Acesso em: 31 jul 2021.

<sup>31</sup> ECP | PLATFORM FOR THE INFORMATION SOCIETY. **Artificial Intelligence Impact Assessment** (2018). Disponível em: <https://ecp.nl/wp-content/uploads/2019/01/Artificial-Intelligence-Impact-Assessment-English.pdf>. Acesso em: 07 jun 2022.

<sup>32</sup> METCALF, Jacob *et al.* **Algorithmic impact assessments and accountability: the co-construction of impacts**. In: ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY (FAccT '21), 2021, evento online. Proceedings p.1-3; ECP | PLATFORM FOR THE INFORMATION SOCIETY. Artificial Intelligence Impact Assessment. ECP | Platform for the Information Society, 2018, p. 5. Relatório técnico. Disponível em: <https://ecp.nl/wp-content/uploads/2019/01/Artificial-Intelligence-Impact-Assessment-English.pdf>. Acesso em: 31 jul 2021.

<sup>33</sup> CANADÁ. *Op. cit.*

<sup>34</sup> ECP | PLATFORM FOR THE INFORMATION SOCIETY. *Op. cit.*, p. 9.

<sup>35</sup> HOLANDA. MINISTRY OF THE INTERIOR AND KINGDOM RELATIONS. **Impact Assessment Fundamental rights and algorithms**. Disponível em: <https://www.government.nl/documents/reports/2022/03/31/impact-assessment-fundamental-rights-and-algorithms>. p. 3. Acesso em: 26 mai 2022.

<sup>36</sup> Destacamos que em trabalhos anteriores utilizamos a nomenclatura “Avaliação de Impacto de Inteligência Artificial” (AIIA). Porém, para fins de compatibilização com a minuta proposta pela Comissão de Juristas, optamos pelo termo Avaliação de Impacto Algorítmico (AIA) neste relatório.

## Avaliação de Impacto Algorítmico vs. Relatório de Impacto à Proteção de Dados Pessoais

Um sistema de IA é desenvolvido a partir do tratamento de um grande número de dados, que podem ser de natureza pessoal ou não pessoal. Ao processar dados pessoais, o sistema de IA deve respeitar a Lei Geral de Proteção de Dados Pessoais (LGPD) brasileira e, quando cabível, deve ser elaborado um Relatório de Impacto à Proteção de Dados Pessoais (RIPD). O Relatório de Impacto à Proteção de Dados é um documento que demonstra os dados pessoais que são tratados,<sup>37</sup> bem como as **medidas adotadas para mitigação dos riscos** envolvidos nesse tratamento que possam afetar as liberdades civis e os direitos fundamentais dos titulares.

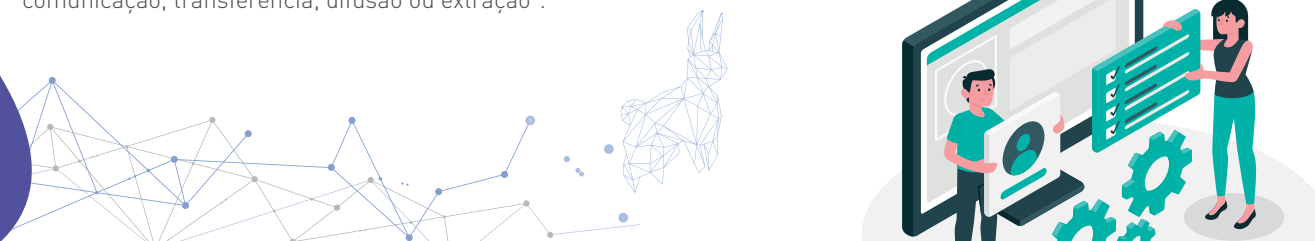
Nesse ponto, o art. 5º, XVII, da LGPD, dispõe que o RIPD é a documentação do controlador que contém a descrição dos processos de tratamento de dados pessoais que podem gerar riscos às liberdades civis e aos direitos fundamentais, bem como medidas, salvaguardas e mecanismos de mitigação de risco.

O art. 38 da LGPD prossegue aduzindo que a Autoridade Nacional de Proteção de Dados (ANPD) poderá determinar ao controlador que elabore o RIPD, que deverá conter, no mínimo, a descrição dos tipos de dados coletados, a metodologia utilizada para a coleta e para a garantia da segurança das informações e a análise do controlador com relação a medidas, salvaguardas e mecanismos de mitigação de risco adotados, observados o segredo comercial e industrial.

A partir desse contexto, é possível que surjam questionamentos sobre o **porquê utilizar uma AIA se já existe um instrumento de governança aplicável a sistemas de IA quando tratam dados pessoais.**

De fato, a AIA é semelhante ao RIPD na medida em que ambos os instrumentos são avaliações de impacto e utilizam lógicas parcialmente semelhantes. Contudo, o RIPD é uma avaliação de impacto que se concentra apenas nos riscos que uma atividade de tratamento de dados pode causar a uma pessoa física ou a um conjunto de pessoas cujos dados pessoais estão sendo tratados. Além disso, por já estar previsto e regulado pela LGPD, o RIPD será de realização obrigatória quando determinado tratamento de dados pessoais puder gerar riscos às liberdades civis e aos direitos fundamentais (art. 5º, inciso XVII, LGPD).

<sup>37</sup> A LGPD em seu art. 5º, inciso X, define tratamento de dados como “toda operação realizada com dados pessoais, como as que se referem a coleta, produção, recepção, classificação, utilização, acesso, reprodução, transmissão, distribuição, processamento, arquivamento, armazenamento, eliminação, avaliação ou controle da informação, modificação, comunicação, transferência, difusão ou extração”.





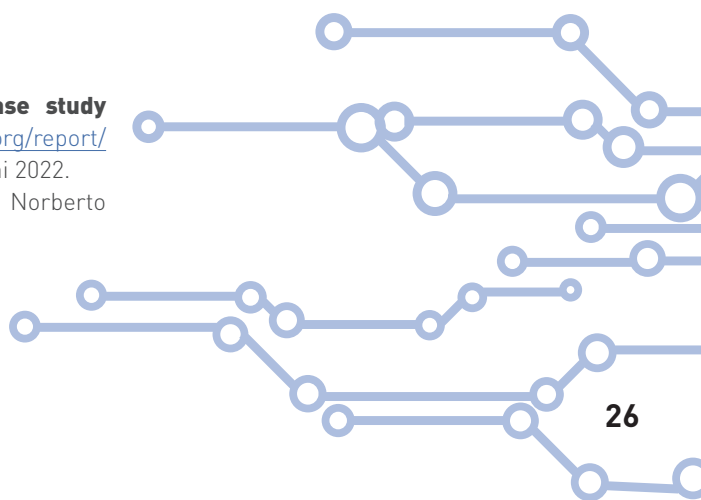
Em contrapartida, **a AIA pode ser considerada um instrumento mais amplo**, já que sua condução exige que se protejam **direitos e interesses** que não se restringem ao tratamento de dados pessoais, mas também envolvem a atividade do próprio algoritmo, da programação e da imprevisibilidade da máquina. Considerando que os dados pessoais são um dos combustíveis utilizados pela IA no desenvolvimento de sua atividade, verifica-se que a AIA abrange uma perspectiva mais ampla, por não se restringir a esse fator, englobando os outros elementos que se unem para caracterizar uma inteligência artificial. Desta forma, a AIA tem como escopo identificar e mitigar externalidades negativas em qualquer campo que a IA atue.

Nesse sentido, uma AIA não se limitará a analisar desdobramentos que sejam apenas provenientes do tratamento de dados pessoais. Além disso, a AIA considera não somente questões legais associadas ao desenvolvimento dos sistemas de IA, como também provoca reflexões sobre preocupações éticas comuns, consideração de consequências intencionais e não intencionais e possíveis medidas para ajudar a mitigar quaisquer danos.<sup>38</sup>

Por fim, vale ressaltar que **a falta de uma regulação própria para uma AIA provoca duas consequências: (i)** discricionariedade quanto à condução de uma AIA, o que nos leva a entender e demonstrar os motivos que tornam a sua realização importante para a proteção de direitos; e **(ii)** falta de orientação legal sobre como conduzir essa avaliação de impacto. Nesse contexto, apesar da execução de alguns estudos e relatórios acerca da AIA,<sup>39</sup> ainda não há um padrão universalmente aceito sobre sua metodologia de execução. Por esta razão, também apresentamos neste relatório (Parte III) uma proposta de metodologia para a condução da AIA.

<sup>38</sup> ADA LOVELACE INSTITUTE. **Algorithmic impact assessment: a case study in healthcare.** p. 5. Disponível em: <https://www.adalovelaceinstitute.org/report/algorithmic-impact-assessment-case-study-healthcare/>. Acesso em: 23 mai 2022.

<sup>39</sup> ECP | PLATFORM FOR THE INFORMATION SOCIETY. *Op. cit.*; ANDRADE, Norberto Nuno Gomes de; KONTSCIEDER, Verena. *Op. cit.*, p. 30, 35-36 e 75-76.



## Por que realizar uma AIA?

Em que pese a previsão contida no anteprojeto de lei apresentado pela Comissão de Juristas, atualmente não há, no ordenamento jurídico brasileiro, previsão legal que determine a condução de uma AIA, eis que o projeto ainda seguirá para apreciação legislativa. Contudo, a sua realização encontra fundamento em panoramas econômicos, reputacionais, regulatórios (inclusive de autorregulação) e de cumprimento de direitos fundamentais.

Apesar da ausência de obrigação legal específica, ressalta-se que, no setor privado, a função social da empresa é norma de natureza cogente no Brasil, caracterizada pelo pensamento do bem-estar coletivo, de modo integral, especialmente no que tange às condutas lesivas de outros entes.<sup>40</sup> Assim, a inobservância de preceitos atinentes à **autonomia e fiscalização humana, prevenção de danos, justiça, robustez, não discriminação, bem-estar social e ambiental, responsabilização e tutela de vulneráveis**<sup>41</sup> pode caracterizar violação ao ordenamento jurídico e acarretar a responsabilidade do infrator.

Como um instrumento de governança, a AIA ajuda a manter a **transparência e a confiança** dos usuários na tecnologia, permitindo expor as capacidades e a finalidade do sistema de IA a todos aqueles que sejam direta ou indiretamente afetados, assumindo especial relevância em uma sociedade com forte tradição consumerista, onde **acesso à informação** figura como direito básico.

Nos casos de sistemas de IA que afetam consumidores, nos termos do art. 6º do **Código de Defesa do Consumidor**, a informação deve ser adequada e clara sobre os diferentes produtos e serviços, com especificação correta de quantidade, características, composição, qualidade, tributos incidentes e preço, bem como sobre os riscos que apresentem.

<sup>40</sup> SANTOS, Adriano Barreto Espíndola. Novos paradigmas para a função social da responsabilidade civil. **Revista Jurídica Luso-Brasileira, Lisboa**, v. 4, n. 3, 2018, p. 15. Disponível em: [https://www.cidp.pt/revistas/rjlb/2018/3/2018\\_03\\_0023\\_0046.pdf](https://www.cidp.pt/revistas/rjlb/2018/3/2018_03_0023_0046.pdf). Acesso em: 31 jul 2021.

<sup>41</sup> Os preceitos referidos são sintetizados no documento sobre Diretrizes Éticas para uma Inteligência Artificial Confiável (2019), documento baseado no trabalho do Grupo Europeu de Ética na Ciência e Novas Tecnologias e outros esforços similares. Disponível em: <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>. Acesso em: 31 jul 2021.



Já em casos em que sistemas de IA tratam dados pessoais, a **LGPD** também impõe como princípios em seu art. 6º, dentre outros:

**A.** a garantia de **transparência**, com informações claras, precisas e facilmente acessíveis sobre a realização do tratamento e os respectivos agentes de tratamento, observados os segredos comercial e industrial (inciso VI);

**B.** a garantia de **segurança**, com utilização de medidas técnicas e administrativas aptas a proteger os dados pessoais de acessos não autorizados e de situações acidentais ou ilícitas de destruição, perda, alteração, comunicação ou difusão (inciso VII);

**C.** a garantia de **prevenção**, com adoção de medidas para prevenir a ocorrência de danos em virtude do tratamento de dados pessoais (inciso VIII);

**D.** e a **responsabilização e prestação de contas**, com a demonstração, pelo agente de tratamento, da adoção de medidas eficazes e capazes de comprovar a observância e o cumprimento das normas de proteção de dados pessoais e, inclusive, da eficácia dessas medidas (inciso X).

Mais à frente neste trabalho, será possível perceber que a realização de uma AIA responde transversalmente aos anseios determinados por estes princípios.

Pode-se também constatar que cada vez mais usuários e consumidores cobram das empresas o uso ético e responsável das tecnologias adotadas a fim de verificar se seus direitos fundamentais estão sendo assegurados. Neste sentido, compete sublinhar o **Princípio 19**, apontado pela **ONU** na carta de **Princípios Orientadores sobre Empresas e Direitos Humanos** de 2011, que afirma que para prevenir e mitigar os impactos negativos sobre os direitos humanos, as empresas devem integrar as conclusões de suas avaliações de impacto em funções e processos internos relevantes e tomar as medidas apropriadas.<sup>42</sup>

Ao considerar razões econômicas e de mercado, um dos projetos aqui analisados, o Open Loop, confirmou os **benefícios da realização de AIAs nos setores europeu e estadunidense de empresas e startups**.<sup>43</sup>

<sup>42</sup> Escritório do Alto Comissariado das Nações Unidas para os Direitos Humanos. **Princípios Orientadores sobre Empresas e Direitos Humanos: Implementando os parâmetros 'Proteger, Respeitar e Reparar' das Nações Unidas**. 2011, p. 20. Disponível em: [https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinesshr\\_en.pdf](https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinesshr_en.pdf). Acesso em: 30 jul 2021.

<sup>43</sup> Open Loop coordenou um experimento de realização de AIA com ao menos 10 empresas e startups que desenvolvem e operam sistemas de IA, como: Allegro.ai (Irlanda), RiAtlas (Itália), NAIX Technology (Alemanhã), Evo Pricing (Reino Unido), Keepler Data Tech (Espanha), Unbabel (Portugal), Feedzai (Estados Unidos), Rogervoice (França), Irida Labs (Grécia) e Reface (Ucrânia). Ver: ANDRADE, Norberto Nuno Gomes de; KONTSCIEDER, Verena. **AI impact assessment: a policy prototyping experiment**. Open Loop, 2021, p. 21-22. Relatório técnico. Disponível em: [https://openloop.org/wp-content/uploads/2021/01/AI\\_Impact\\_Assessment\\_A\\_Policy\\_Prototyping\\_Experiment.pdf](https://openloop.org/wp-content/uploads/2021/01/AI_Impact_Assessment_A_Policy_Prototyping_Experiment.pdf). Acesso em: 31 jul 2021.

Ao final do estudo, constatou-se que todos os participantes eram mais capazes de identificar e mitigar os riscos associados às suas aplicações, assim como de incorporar as melhores práticas e salvaguardas no design de seus produtos,<sup>44</sup> resultando em “maior eficiência e entrega mais rápida ao mercado, reduzindo os custos e riscos de interrupção posterior”.<sup>45</sup>

A avaliação de impacto também ajuda a organização a cumprir a lei, criar confiança com seus clientes e melhorar a sua governança interna.<sup>46</sup> A sociedade também se beneficia com mais informação, experiência de melhores resultados e engajamento nas tomadas de decisões.<sup>47</sup>

A adoção de AIAs pode ser uma resposta a questões legalmente transversais ao desenvolvimento e operação de sistemas de IA, como também pode trazer outros benefícios à própria empresa, **como maior valor reputacional**,<sup>48</sup> aumento da relação de confiança com usuários e consumidores, eficiência, entre outros.

Dessa forma, a realização da AIA não deve ser vista como um fardo administrativo. Ao contrário, a AIA deve ser encarada como uma ferramenta de suporte ao ciclo de vida dos sistemas de IA. Isso porque ela foi desenvolvida para ajudar corporações a entenderem e gerenciarem melhor os riscos associados aos sistemas,<sup>49</sup> protegendo os direitos fundamentais das pessoas físicas envolvidas.

Aclaradas as vantagens e as razões para a adoção da AIA por empresas e organizações, na próxima seção serão indicados quais sistemas devem ser submetidos a este tipo de avaliação.

<sup>44</sup> Centre for Information Policy Leadership (CIPL). **Recommendations on Adopting a Risk-Based Approach to Regulating Artificial Intelligence in the EU**. CIPL, 2021, p. 3-4. Relatório técnico. Disponível em: [https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl\\_risk-based\\_approach\\_to\\_regulating\\_ai\\_22\\_march\\_2021\\_.pdf](https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_risk-based_approach_to_regulating_ai_22_march_2021_.pdf). Acesso em: 31 jul 2021.

<sup>45</sup> *Ibidem*. Conclusão semelhante pode se extrair de BRYNJOLFSSON, Eric; MCAFEE, Andrew. The business of Artificial Intelligence. **Harvard Business Review**, edição especial Artificial Intelligence, Boston, jul 2017. Disponível em: <https://hbr.org/2017/07/the-business-of-artificial-intelligence>. Acesso em: 31 jul 2021.

<sup>46</sup> ICO. **The Alan Turing Institute**. Disponível em: <https://ico.org.uk/for-organisations/guide-to-data-protection/key-dp-themes/explaining-decisions-made-with-artificial-intelligence/part-1-the-basics-of-explaining-ai/benefits-and-risks/>. Acesso em: 27 mai 2022.

<sup>47</sup> *Ibidem*. Acesso em: 27 mai 2022.

<sup>48</sup> O valor reputacional é compreendido como a reputação institucional ou corporativa relacionada com a responsabilidade social da empresa e que se sustenta em eixos vinculados à missão, gestão e valores. Além disso, é um conceito que emerge da imagem global da empresa, como resultado das percepções dos indivíduos nas dimensões de qualidade e responsabilidade. Para Capriotti, a reputação corporativa é um ativo intangível que faz parte das abordagens atuais econômicas, gerido pela empresa para alcançar uma adequada avaliação em seus grupos de interesses. Ver CAPRIOTTI, P. Schulze. **Responsabilidad Social Empresarial**. Barcelona: Colección de libros Executive Business School, 2010, p. 115.

<sup>49</sup> CANADÁ. *Op. cit.*

## Quais sistemas de IA devem ser submetidos a uma AIA?

Atualmente ainda existe uma ampla discussão, a nível internacional, sobre como diferenciar de modo preciso o que seria uma **atividade de baixo ou médio risco de uma atividade de alto risco**. A OECD, por exemplo, publicou o “*The OECD Framework for the Classification of AI systems*”,<sup>50</sup> que traz como dimensões-chave para a delimitação de riscos, em síntese, pessoas e planeta, contexto econômico, dados e modelo de IA. Ainda, o debate sobre a classificação de riscos culminou no desenvolvimento de duas propostas: a prescritiva e a procedimental.<sup>51</sup>

A **proposta prescritiva** é aquela ancorada em classificações rígidas que indicam previamente o **tipo de utilização** e o **tipo de setor** em que a aplicação poderá ser considerada de alto risco, como, por exemplo, identificação de padrões de doenças no setor da saúde.<sup>52</sup> Por exigir uma definição prévia a respeito dos setores em que um sistema de IA pode ser considerado de alto risco, pode-se abrir lacunas em que uma AIA não será conduzida em usos em diferentes e novas áreas de expansão da IA de alto risco.

<sup>44</sup> Centre for Information Policy Leadership (CIPL). **Recommendations on Adopting a Risk-Based Approach to Regulating Artificial Intelligence in the EU**. CIPL, 2021, p. 3-4. Relatório técnico. Disponível em: [https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl\\_risk-based\\_approach\\_to\\_regulating\\_ai\\_22\\_march\\_2021.pdf](https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_risk-based_approach_to_regulating_ai_22_march_2021.pdf). Acesso em: 31 jul 2021.

<sup>45</sup> *Ibidem*. Conclusão semelhante pode se extrair de BRYNJOLFSSON, Eric; MCAFEE, Andrew. The business of Artificial Intelligence. **Harvard Business Review**, edição especial Artificial Intelligence, Boston, jul 2017. Disponível em: <https://hbr.org/2017/07/the-business-of-artificial-intelligence>. Acesso em: 31 jul 2021.

<sup>46</sup> ICO. The Alan Turing Institute. Disponível em: <https://ico.org.uk/for-organisations/guide-to-data-protection/key-dp-themes/explaining-decisions-made-with-artificial-intelligence/part-1-the-basics-of-explaining-ai/benefits-and-risks/>. Acesso em: 27 mai 2022.

<sup>47</sup> *Ibidem*. Acesso em: 27 mai 2022.

<sup>48</sup> O valor reputacional é compreendido como a reputação institucional ou corporativa relacionada com a responsabilidade social da empresa e que se sustenta em eixos vinculados à missão, gestão e valores. Além disso, é um conceito que emerge da imagem global da empresa, como resultado das percepções dos indivíduos nas dimensões de qualidade e responsabilidade. Para Capriotti, a reputação corporativa é um ativo intangível que faz parte das abordagens atuais econômicas, gerido pela empresa para alcançar uma adequada avaliação em seus grupos de interesses. Ver CAPRIOTTI, P. Schulze. **Responsabilidad Social Empresarial**. Barcelona: Colección de libros Executive Business School, 2010, p. 115.

<sup>49</sup> CANADÁ. *Op. cit.*

<sup>50</sup> OECD. **OECD framework for the classification of AI systems**. OECD Digital Economy Papers. n. 323, Fevereiro, 2022. Disponível em: <https://oecd.ai/en/classification>. Acesso em: 24 mai 2022.

<sup>51</sup> ANDRADE, Norberto Nuno Gomes de; KONTSCHIEDER, Verena. *Op. cit.*, p. 11-12.

<sup>52</sup> A proposta prescritiva de definição de atividades de alto risco pode ser encontrada no “Livro Branco sobre Inteligência Artificial” da Comissão Europeia. Bruxelas: Comissão Europeia, 2020. Disponível em: <https://op.europa.eu/pt/publication-detail/-/publication/ac957f13-53c6-11ea-aece-01aa75ed71a1>. Acesso em: 19 mai 2022.

A Comissão Europeia classifica uma aplicação em alto risco observando se o setor e a utilização prevista envolvem **riscos significativos**, notadamente do ponto de vista da proteção da segurança, dos direitos dos consumidores e dos direitos fundamentais, tais como os setores de saúde, transportes e energia, por exemplo.<sup>53</sup>

Por outro lado, **a proposta procedimental** defende que a definição de alto risco depende de um conjunto de etapas e indagações destinadas a alcançar uma identificação do risco mediante **diálogo, reflexão e uma análise qualitativa** de informações associadas ao sistema de IA em si, e não somente ao setor e tipo de utilização a que se propõe.<sup>54</sup> A definição do risco, portanto, não se consagra aprioristicamente, dependendo de uma **análise do caso concreto**.

<sup>53</sup> UNIÃO EUROPEIA. *Op. cit.*, p. 20.

<sup>54</sup> ANDRADE, Norberto Nuno Gomes de; KONTSCIEDER, Verena. *Op. cit.*, p. 9.

## PROPOSTAS

## PRESCRITIVA

## PROCEDIMENTAL

|                                 |                                                                                                                                                                                                                                                                                                                                                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|---------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Abordagem                       | Indicação prévia em <b>abstrato</b> do tipo de utilização e do setor que pode ser considerado de alto risco.                                                                                                                                                                                                                                               | Indicação de um conjunto de etapas e/ou questões destinadas a chegar a <b>determinações de risco</b> por meio de análise qualitativa, diálogo e reflexão.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| Forma de classificação de risco | Setores enumerados de <b>forma exaustiva e específica</b> em quadro regulamentar que é periodicamente revisto e atualizado (ex: saúde, transportes e energia). Dadas as características das atividades tipicamente realizadas, esperam-se riscos significativos.                                                                                           | O nível de <b>incerteza e complexidade</b> dos tipos de risco requer abordagens procedimentais robustas para avaliação de risco, em vez de uma abordagem ancorada em classificações rígidas com base no setor em que a IA está sendo utilizada.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| Operacionalidade                | A proposta também enumera <b>casos excepcionais</b> em que, a despeito de não constarem nos setores indicados como de alto risco, serão assim considerados em razão dos elevados riscos em causa, tais como utilização de IA para recrutamento de emprego, matéria de relações de trabalho, identificação biométrica à distância e vigilâncias intrusivas. | <b>Aplicação equilibrada de requisitos em resposta aos riscos identificados:</b> supervisão humana, explicabilidade, direitos de reparação, monitoramento, requisitos de divulgação, entre outros. Além disso, as organizações devem avaliar a própria <b>ética corporativa e valores</b> ; estruturas de governança interna e medidas institucionais e de segurança; funções internas, equipes e responsabilidades; gerenciamento de operações; e estratégias para se comunicar com partes interessadas externas para determinar os riscos apresentados por sistemas de IA. Por tal abordagem, <b>os requisitos são atribuídos de acordo com a IA específica</b> em questão e o nível e extensão dos riscos avaliados, juntamente com o cálculo dos benefícios que a aplicação traz. |

Há um alto risco quando há uma chance significativa de que as decisões automatizadas tomadas pelo sistema de IA possam resultar ou influenciar em efeitos negativos nos direitos e liberdades das pessoas físicas, tais como os elencados abaixo. A **compreensão do que é risco inclui**, mas não se limita a, efeitos adversos que podem causar:<sup>55</sup>

- Interferência nos direitos fundamentais, como os direitos à igualdade e à não discriminação, à privacidade, e à liberdade de expressão;
- Potenciais danos à saúde ou segurança do sujeito, como perda da vida ou danos corporais;
- Potenciais danos psicológicos, como a autocensura, a perda de autoestima e a perda de autonomia pessoal;
- Potenciais danos sociais ou econômicos, como danos financeiros, perda de propriedade ou restrição a serviços públicos ou privados;
- Potenciais danos reputacionais ou de estigmatização;
- Potenciais preconceitos ou discriminações injustas em relação ao sujeito, incluindo discriminação de preços, discriminação no emprego ou acesso diferenciado ou discriminatório aos serviços;
- Potencial perda de controle ou supervisão para o sujeito, incluindo manipulação econômica ou psicológica;
- Potenciais danos coletivos, como perda de liberdade ou instabilidade econômica ou política;
- Aplicação em larga escala de tomada de decisões automatizada, incluindo a definição de perfis e o monitoramento sistemático, que podem afetar as comunidades ou a sociedade como um todo.

Sob o **critério procedimental**, será sempre recomendada a condução de uma AIA quando for identificado que um sistema de IA, cuja aplicação ocorre em larga escala de tomada de decisões automatizada (incluindo a definição de perfis e o monitoramento sistemático), possa causar um dos riscos acima indicados. Em outras palavras, sugere-se um **requisito duplo**<sup>56</sup> para a condução de uma AIA. Destaca-se, novamente, que essa se trata de uma lista exemplificativa, e que o responsável pelo sistema de IA deve chegar à conclusão sobre o risco de seu sistema mediante diálogo, reflexão e análise qualitativa de informações associadas à aplicação.

<sup>55</sup> ANDRADE, Norberto Nuno Gomes de; KONTSCIEDER, Verena. *Op. cit.*, p. 30, 35-36 e 75-76.

<sup>56</sup> A definição do requisito duplo para a realização de uma AIA é fruto das discussões dos autores deste relatório.



Este relatório sugere a abordagem conjunta dos dois critérios procedimentais, assim como parece ter sugerido a Comissão de Juristas na minuta do substitutivo, já que:

- O art. 17 enumera de forma específica tipos de utilização e setores previamente considerados como de alto risco, demonstrando a aplicação da proposta prescritiva;
- O art. 18 permite a atualização periódica da lista de sistemas pré-classificados como de alto risco, tratando-se, também, de traço da proposta prescritiva;
- Entretanto, o art 24, ao definir uma metodologia para a avaliação de impacto, traz consigo um ferramental típico da proposta procedimental.

Essa abordagem conjunta assegura que, **independente da utilização ou do setor, devem ser objeto de uma AIA todas as aplicações que possam resultar em um alto risco para os direitos e liberdades de pessoas físicas**, principalmente quando o potencial de dano for considerado especialmente grave ou irreversível.<sup>57</sup> Nesse contexto, verifica-se que ambas as abordagens estão presentes na proposta da Comissão de Juristas, no intuito de assegurar a verificação do risco.

## Quais são os atores associados a sistemas de IA?

Por estarem no centro do desenvolvimento econômico e cultural da sociedade, os sistemas de IA provocam constantemente efeitos jurídicos através do apoio/tomada de decisões em si. Por isso, é importante identificar os principais e diferentes atores que atuam dentro do ciclo de vida dos sistemas de IA.

Os atores de IA são aqueles que desempenham um **papel ativo** em todo o ciclo de vida do sistema de IA e podem incluir organizações e indivíduos que implantam ou operam o sistema.<sup>58</sup> Pensando nisso, é possível apontar ao menos quatro atores:<sup>59</sup>

<sup>57</sup> MARTINS, Guilherme Magalhães. **Riesgo, solidaridad y responsabilidad civil**. In: MORE, César E. Moreno (coord.). Estudios sobre la responsabilidad civil. Legales: Lima, 2015, p. 189.

<sup>58</sup> OECD. **OECD framework for the classification of AI systems**. OECD Digital Economy Papers. n. 323, Fevereiro, 2022. Disponível em: <https://www.oecd-ilibrary.org/docserver/cb6d9eca-en.pdf?expires=1652284916&id=id&accname=guest&checksum=D1B84F556282520301213485FA41E9CE>. p. 22. Acesso em: 11 mai. 2022.

<sup>59</sup> ANDRADE, Norberto Nuno Gomes de; KONTSCIEDER, Verena. *Op. cit.*, p. 74-75; OECD. **OECD framework for the classification of AI systems**. OECD Digital Economy Papers. n. 323, Fevereiro, 2022. Disponível em: <https://www.oecd-ilibrary.org/docserver/cb6d9eca-en.pdf?expires=1652284916&id=id&accname=guest&checksum=D1B84F556282520301213485FA41E9CE>. Acesso em: 11 mai. 2022. p. 22; NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY. AI Risk Management Framework. Disponível em: <https://www.nist.gov/system/files/documents/2022/03/17/AI-RMF-1stdraft.pdf>. Acesso em: 11 mai. 2022.

- **Desenvolvedor:** pessoa física ou jurídica responsável pelo desenvolvimento do sistema de IA. Como desenvolvimento, entende-se tanto: (i) a codificação do algoritmo de aprendizagem que irá basear o sistema de IA, como também (ii) a seleção dos dados (de treinamento) e a subsequente criação e/ou treinamento do modelo;<sup>60</sup>
- **Fornecedor:** pessoa física ou jurídica, de natureza pública ou privada, responsável pela disponibilização de sistema de IA para que terceiros o operem, mediante a colocação do sistema no mercado nacional ou mediante a aplicação em serviço por ela fornecido, a título oneroso ou gratuito;
- **Operador de sistema de IA:**<sup>61</sup> pessoa física ou jurídica responsável pela implementação e operação de um sistema de IA para atingir um determinado objetivo;
- **Sujeito afetado:** pessoa física que é direta ou indiretamente sujeita ou impactada pela decisão de um sistema de IA.

Importa mencionar que há casos em que o **fornecedor, desenvolvedor e/ou o operador do sistema de IA podem figurar na mesma pessoa/organização.**

Vale também ressaltar a classificação de atores da OCDE, que identifica cinco categorias que se diferem em razão da dimensão: (i) *end-users*; (ii) *system operators*; (iii) *data collectors and processors*; (iv) *developers and modellers* e (v) *system integrators*. Esta classificação se baseia no seguinte cenário:<sup>62</sup>

- **Pessoas e Planeta (*end-users*):** Usuários finais e partes interessadas que usam ou são afetados por sistemas de IA. Abrange todas as organizações e indivíduos envolvidos ou afetados por sistemas de IA, direta ou indiretamente.

<sup>60</sup> No projeto Open Loop desenvolvedor foi definido como “a pessoa física ou jurídica que desenvolveu o sistema automatizado de tomada de decisão. Este ator pode apenas fornecer o algoritmo de aprendizagem, mas é mais provável que seja a pessoa ou organização responsável pela seleção dos dados (de treinamento) e algoritmos de aprendizagem relevantes e a subsequente criação e/ou treinamento do modelo” (tradução livre). *Op. cit.* p. 74.

<sup>61</sup> Vale ressaltar que em nossos trabalhos anteriores nominamos este agente como “Aplicador da tecnologia”.

<sup>62</sup> OECD. *Op. cit.* p. 22.



- **Contexto Econômico (*system operators*):** Operadores de sistema que planejam, projetam, operam e monitoram a IA aplicada a sistemas.

- **Dados e Entrada (*data collectors and processors*):** Coletores e processadores de dados, incluindo coleta e limpeza de dados, rotulagem, realização de verificações de integridade e qualidade, e documentação das características do conjunto de dados.

- **Modelo de IA (*developers and modellers*):** Desenvolvedores e modeladores que constroem e usam modelos, bem como os verificam e os validam.

- **Tarefa e Saída (*system integrators*):** Integradores de sistemas que implantam sistemas de IA.

Além disso, o *AI Risk Management Framework* do *National Institute of Standards and Technology* (NIST) define quatro grupos: (i) *AI system stakeholders*; (ii) *operator and evaluators*; (iii) *external* e (iv) *general public*.<sup>63</sup>

- **Partes interessadas no sistema de IA (*AI system stakeholders*):** são aqueles que têm mais controle e responsabilidade sobre o design, desenvolvimento, implantação e aquisição de sistemas de IA, bem como sobre a implementação de práticas de gestão. Pode incluir indivíduos ou equipes de organizações com responsabilidades de financiar, adquirir, desenvolver ou implantar um sistema de IA: equipes de negócios, de design e desenvolvimento, equipes internas de gerenciamento de riscos e de *compliance*.

- **Operadores e avaliadores (*operators and evaluators*):** no intuito de avaliar seu desempenho, fornecem monitoramento e testes formais/informais de avaliação, validação e verificação do desempenho do sistema, tanto em termos técnicos quanto sociotécnicos. Este grupo pode incluir usuários que interpretam ou incorporam a saída de sistemas de IA em ambientes com alto potencial de impactos adversos, como pesquisadores do setores acadêmico, público e privado, avaliadores e auditores profissionais, operadores do sistema e usuários finais especializados.

- **Partes interessadas externas (*external public*):** fornecem normas ou orientações para especificar e abordar os riscos da IA, bem como limitar ou equilibrar a operação observando valores relacionados às liberdades e direitos civis, economia e segurança. Pode incluir grupos comerciais, organizações de desenvolvimento de padrões e organizações da sociedade civil.

- **Público em geral (*general public*):** população que experimenta diretamente os impactos positivos e negativos da IA. Podem fornecer a motivação para as ações tomadas pelas outras partes interessadas, incluindo indivíduos, comunidades e consumidores no contexto em que um sistema de IA é desenvolvido ou implantado.

<sup>63</sup> National Institute of Standards and Technology. *Op. cit.*



Para os fins do presente relatório, será utilizada a primeira classificação: **desenvolvedor, fornecedor, operador da tecnologia e sujeito afetado**. Isso porque tal nomenclatura também é citada em outros frameworks de renome, tais como os da OCDE e do NIST e, ao mesmo tempo, traduz a proposta de uma maneira simplificada, já que outros modelos consideram demasiados atores.

Ao considerar os incontáveis modelos de negócio que a inteligência artificial pode promover, inclusive aqueles que ainda serão criados, poderão ocorrer casos específicos em que não haja compatibilidade ou identificação dos atores aqui definidos com a situação real e complexa associada ao ciclo de vida do sistema de IA. Por isso, essa é uma definição que estará em contínuo amadurecimento.

## Quem deve realizar uma AIA em cada uma das etapas do ciclo de vida de um sistema de IA?

Existem **ao menos cinco momentos em que é recomendável realizar uma AIA**, dependendo se o responsável pelo sistema de IA estiver no papel de desenvolvedor ou operador da tecnologia:

**1) No início da fase de planejamento e design de um projeto:** ao definir pela inclusão de técnicas de inteligência artificial em um determinado projeto, o **desenvolvedor** e sua equipe devem se preparar para conduzir uma AIA no momento do design do sistema de IA.<sup>64</sup>

Isso permite que os desenvolvedores incluam considerações técnicas e legais ainda no design da aplicação, aumentando a viabilidade de desenvolvimento, de comercialização do produto e reduzindo custos, uma vez que posteriormente pode se tornar impossível ou muito custoso fazer qualquer alteração.<sup>65</sup>

**2) Após finalizar o desenvolvimento do sistema de IA:** a realização da AIA uma segunda vez, agora ao final da fase de desenvolvimento, permitirá que o **desenvolvedor** e sua equipe validem os resultados do sistemas de IA, isto é, verifiquem se eles refletem exatamente aquilo que foi projetado e desenvolvido sem que traga efeitos negativos, planejados ou não, sobre direitos.<sup>66</sup>

<sup>64</sup> CANADÁ. *Op. cit.*

<sup>65</sup> ECP | PLATFORM FOR THE INFORMATION SOCIETY. *Op. cit.*, p. 29.

<sup>66</sup> CANADÁ. *Op. cit.*



**3) Após uma modificação substancial do sistema de IA:** apesar de o sistema de IA ter sido desenvolvido de forma responsável, e mesmo já tendo sido conduzidas AIAs, é recomendável que se realize uma nova AIA sempre que houver:

- a)** Uma modificação substancial acerca da forma de operação do sistema de IA em suas bases de dados de treinamento;
- b)** Uma mudança/ampliação da sua finalidade originalmente planejada;<sup>67</sup> ou
- c)** Caso o sistema de IA atue com aprendizado de máquina de modo que a sua tomada de decisão possa ser influenciada por novos dados e informações ao longo do tempo.<sup>68</sup>

**4) Antes da aquisição/utilização do sistema de IA:** é recomendável que o **operador da tecnologia** realize uma AIA quando, ao adquirir um sistema de IA, o operacionalize a partir da sua própria base de dados e informações. Essa é uma medida que promove transparência e faz com o aplicador da tecnologia tenha ganho reputacional perante parceiros e sujeitos afetados.

**5) Avaliação periódica:** independentemente se o responsável é o **desenvolvedor, fornecedor ou o operador** da tecnologia do sistema de IA, é recomendável que se defina um intervalo de tempo para avaliar periodicamente a aplicação. Há recomendações para que a avaliação seja conduzida anualmente ou a cada dois anos.<sup>69</sup>

Por fim, os responsáveis por sistemas de IA podem também implementar a AIA como parte dos processos de *due diligence* (diligência prévia),<sup>70</sup> que são processos que envolvem gestão de riscos e *compliance* (conformidade), auditoria, previsão de cenários de negócios e identificação de oportunidades. Inserir a AIA dentro deste ciclo de diligências faz com que a AIA se torne parte de um processo contínuo, que será regularmente revisto e atualizado.<sup>71</sup>

<sup>67</sup> ECP | PLATFORM FOR THE INFORMATION SOCIETY. *Op. cit.*, p. 61.

<sup>68</sup> REISMAN, Dillon; SCHULTZ, Jason; CRAWFORD, Kate; WHITTAKER, Meredith. *Op. cit.*, p. 10.

<sup>69</sup> *Ibidem*; ANDRADE, Norberto Nuno Gomes de; KONTSCHIEDER, Verena. *Op. cit.*, p. 76.

<sup>70</sup> De acordo com a empresa Legal Ethics Compliance - LEC, "o due diligence é o processo estruturado de estudo, auditoria, investigação e avaliação de riscos e oportunidades em operações empresariais. Pode ser traduzido como 'diligência prévia' ou 'diligência devida'". Ver: LEC. **Due Diligence: entenda o que é e quais são as suas aplicações**. Setembro, 2019. Disponível em: <https://lec.com.br/due-diligence-entenda-o-que-e-e-quais-sao-as-suas-aplicacoes/>. Acesso em: 03 nov 2021.

<sup>71</sup> SHERPA. Recommendation: **Develop baseline model for AI impact assessments**. Disponível em: <https://www.project-sherpa.eu/ai-impact-assessment/>. Acesso em: 24 de nov 2021.



## Consideração de consulta a terceiros

Ao identificar a necessidade da AIA, recomenda-se que a sua realização tenha o aconselhamento de uma equipe jurídica especializada e o apoio de especialistas independentes de áreas como tecnologia da informação, sociologia e ética,<sup>72</sup> de modo a **identificar os riscos a partir de uma perspectiva multidisciplinar**.

Contudo, além do envolvimento da equipe interna da organização, é altamente recomendado que terceiros externos à organização possam participar da AIA a depender da amplitude,<sup>73</sup> da finalidade<sup>74</sup> ou do risco<sup>75</sup> associado ao desenvolvimento ou operação do sistema de IA - em outras palavras, aplicações que afetem significativamente a sociedade, ainda que em grupos específicos ou minorias.

Ainda, **permitir que a sociedade civil, a academia e os próprios sujeitos afetados** pela aplicação acessem e se manifestem a respeito do sistema tendo a AIA como base, demonstrará transparência e responsabilidade social, podendo agregar em **valor reputacional**.<sup>76</sup> Essa participação pode ser viabilizada, por exemplo, por meio de webinários, audiências públicas, fóruns, listas de discussão, elaboração de parecer e nota técnica ou qualquer outra via que permita o amplo debate de ideias.

Para além disso, ressalta-se que a consulta a terceiros dependerá do contexto, da amplitude, da finalidade, do risco associado e da estrutura corporativa. Isso porque microempresas e empresas de pequeno porte, por exemplo, podem não contar com especialistas internos disponíveis e nem orçamento para custear consultorias externas, o que torna mais desafiadora a elaboração da avaliação de impacto. Nesse ponto, embora o núcleo da AIA não possa ser suprimido, recomenda-se que sua elaboração seja feita em compatibilidade e proporcionalidade com a estrutura, porte e impacto da tecnologia a ser analisada.

<sup>72</sup> Information Commissioner 's Office (ICO). *Op. cit.*

<sup>73</sup> Por exemplo, um sistema de IA privado que pretende utilizar reconhecimento facial em locais públicos, como o metrô ou outra zona de grande movimento de pessoas.

<sup>74</sup> Por exemplo, o desenvolvimento de hardwares ou softwares baseados em inteligência artificial que possibilitem grande automação de processos e procedimentos e que, consequentemente, possam causar impacto no mercado de trabalho.

<sup>75</sup> Por exemplo, o desenvolvimento de um sistema de IA voltado para moderação de conteúdo de redes sociais.

<sup>76</sup> ECP | PLATFORM FOR THE INFORMATION SOCIETY. *Op. cit.*, p. 59.



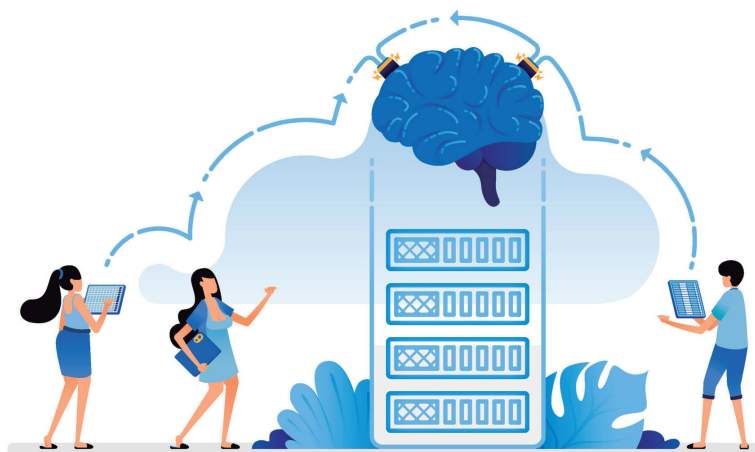
## Publicação do resultado

É recomendável optar por publicizar o resultado de uma AIA, principalmente para demonstrar para a sociedade que a empresa desenvolve e opera a tecnologia de maneira **responsável**. Contudo, a realização de uma AIA pode envolver informações relevantes e que estão diretamente relacionadas com o modelo de negócio e segredos comerciais da empresa. Essa é uma situação em que a publicação do resultado de uma AIA pode ser feita de forma não integral.<sup>77</sup>

Nesses casos, é recomendável que se apresente à sociedade apenas um resumo qualificado da AIA ou sua conclusão, contendo informações mínimas para que quem a analise compreenda os riscos envolvidos no sistema.<sup>78</sup> Tal publicação também tem o escopo de contribuir para o fortalecimento de uma relação de **transparência e confiança**.

<sup>77</sup> O ordenamento jurídico brasileiro protege o segredo comercial e industrial, existindo diversos artigos da própria LGPD que colocam tais segredos como um limite à própria transparência e ao direito à informação do titular de dados, como o artigo 6º, VI, LGPD.

<sup>78</sup> GRUPO DE TRABALHO DO ARTIGO 29º PARA A PROTEÇÃO DE DADOS. **Orientações relativas à Avaliação de Impacto sobre a Proteção de Dados (AIPD) e que determinam se o tratamento é «suscetível de resultar num elevado risco» para efeitos do Regulamento (UE) 2016/679**, abril de 2017, p. 21. Disponível em: [https://www.cnpd.pt/media/f0ide5i0/aipd\\_wp248rev-01\\_pt.pdf](https://www.cnpd.pt/media/f0ide5i0/aipd_wp248rev-01_pt.pdf). Acesso em: 31 jul 2021.



## Benefícios esperados pela adoção da AIA no setor empresarial

Seguidos os passos descritos acima, a AIA aparece, neste cenário, como um instrumento que viabiliza **desenvolvimento econômico e tecnológico empresarial, primando pela preservação dos direitos**. Dada esta premissa, a realização da AIA no âmbito empresarial visa o alcance dos seguintes resultados:

1. Previsão e entendimento a respeito dos efeitos de curto e longo prazo, e dos riscos trazidos pelo uso de sistemas de decisões automatizadas;<sup>79</sup>
2. Previsão e mensuração do impacto dos sistemas de decisões automatizadas em diferentes grupos demográficos e sociais, assim como entendimento a respeito das razões que levam a esta diferença;<sup>80</sup>
3. Prevenção de erros, imprecisões e discriminação;
4. Melhor entendimento a respeito da tecnologia em questão e se ela é apropriada para atingir os fins pretendidos;
5. Demonstração de respeito ao cliente e ao usuário e aos seus direitos ao publicizar quais sistemas de decisões automatizadas estão sendo utilizados e qual o impacto gerado a nível individual e coletivo;<sup>81</sup>
6. Elevação dos índices de confiança da empresa dentre (potenciais) clientes ou usuários e estabelecimento de uma boa reputação no mercado, tornando-a mais competitiva;
7. Aumento da capacidade e expertise interna a respeito dos sistemas de decisões automatizadas, permitindo à empresa avaliar e escolher melhor os sistemas de decisões automatizadas que utiliza, além de antecipar potenciais problemas;<sup>82</sup>
8. Prevenção de litígios relacionados ao uso de sistemas de IA;
9. Identificação de informações úteis que vão auxiliar a concepção e criação de mecanismos internos para apoiar decisões automatizadas e mitigar os eventuais impactos negativos por elas produzidos;<sup>83</sup>
10. Maior competitividade para empresas do setor tecnológico que fornecem apoio para a realização de AIA e que priorizam equidade, transparência e prestação de contas em relação ao tipo de tecnologia que comercializam.

Discutidos os benefícios e as características de uma AIA, apresenta-se a seguir uma possível metodologia para sua condução pelo setor privado no Brasil.

<sup>79</sup> REISMAN, Dillon; SCHULTZ, Jason; CRAWFORD, Kate; WHITTAKER, Meredith. *Op. cit.*, p. 3.

<sup>80</sup> *Ibidem*.

<sup>81</sup> *Ibidem*.

<sup>82</sup> *Ibidem*.

<sup>83</sup> *Ibidem*.

## PARTE III: CONDUZINDO UMA AVALIAÇÃO DE IMPACTO ALGORÍTMICO

Nesta seção, é apresentada uma metodologia que pretende orientar a condução de uma AIA pelas empresas e organizações através de uma **perspectiva de direitos fundamentais**. Isso não significa, porém, que esta será uma estrutura imperativa, uma vez que organizações podem adaptar o modelo em face das peculiaridades de seu cotidiano de trabalho.

Compete sublinhar que a condução da AIA é um processo contínuo, que segue em fluxo de monitoramento e fiscalização mesmo após a conclusão e publicação, uma vez que novos riscos podem surgir à medida em que o sistema atua. Nesse mesmo sentido, o art. 25 da minuta de substitutivo proposta pela Comissão de Juristas, determina que a avaliação consistirá em processo interativo contínuo, executado ao longo do todo o ciclo de vida dos sistemas de alto risco, requeridas atualizações periódicas e com participação pública, a partir de procedimento de consulta a partes interessadas, ainda que de maneira simplificada.

A estrutura a seguir foi inspirada em relatórios e artigos científicos que defendem o uso do relatório de impacto à proteção de dados pessoais em sistemas de IA, bem como nas propostas de avaliações de impacto voltadas para decisões automatizadas tomadas por sistemas de IA. A seção 'Metodologia' apresenta o processo de desenvolvimento da estrutura desta AIA de forma detalhada. Ao final de cada etapa será adicionado um exemplo, especialmente voltado para a hipótese de um **sistema de moderação de conteúdo**, com o objetivo de demonstrar em uma perspectiva mais concreta como se desenvolve a aplicação da avaliação de impacto.

### Metodologia para Avaliação de Impacto Algorítmico

- **Etapa 1.** Descrição detalhada do sistema de IA
- **Etapa 2.** Avaliação da legalidade da finalidade do sistema de IA
- **Etapa 3.** Avaliação da necessidade de participação de terceiros
- **Etapa 4.** Identificação e avaliação dos riscos envolvidos
- **Etapa 5.** Definição de salvaguardas, medidas de segurança e medidas de proteção dos direitos e liberdades dos sujeitos
- **Etapa 6.** Avaliação da proporcionalidade entre finalidade e resultado
- **Etapa 7.** Riscos residuais e a necessidade de consulta prévia a uma Autoridade de Supervisão
- **Etapa 8.** Documentação e publicação



## Etapa 1. Descrição detalhada do sistema de IA

A descrição do sistema de IA, etapa prevista em todas as propostas de AIA levantadas neste relatório,<sup>84</sup> deve responder a pelo menos seis questões:

- (I) qual o **contexto** no qual o sistema de IA será utilizado;
- (II) qual a respectiva **finalidade** do sistema de IA;
- (III) qual a **lógica de operação** do sistema de IA;
- (iv) qual **tipo de dado** será utilizado e como será realizado o seu **treinamento**;
- (v) quais são os **sujeitos afetados** de forma direta, indireta, intencional ou não intencional pelo sistema de IA;
- (vi) quais são os **benefícios** do sistema de IA;
- (vii) quais as **leis** que o sistema de IA deve respeitar.

O **contexto** está ligado ao **setor** em que o sistema de IA será utilizado, sendo a sua indicação importante na medida em que se torna possível perceber, de antemão, se o sistema de IA será utilizado em um setor sensível, como saúde, marketing, financeiro, educação, entre outros.

A **finalidade**, por sua vez, refere-se ao exato objetivo que o sistema de IA quer alcançar, como, por exemplo, determinar o tipo e a quantidade de remédio que se deverá administrar em um paciente ou recomendar filmes em uma plataforma de *streaming*. Ressalta-se que a indicação da finalidade viabiliza a realização de um primeiro juízo de valor a respeito dos possíveis riscos que o uso do sistema de IA pode trazer consigo. A este respeito, é importante destacar que é possível que um sistema de IA possua mais de uma finalidade. Neste caso, deve o responsável pela condução da AIA informar eventual hierarquia entre elas.

<sup>84</sup> REINO UNIDO. Information Commissioner's Officer - ICO. **Guidance on AI and data protection**. Mar 2023, p. 17. Disponível em: <https://ico.org.uk/media/for-organisations/guide-to-data-protection/key-dp-themes/guidance-on-ai-and-data-protection-2-0.pdf>. Acesso em 28 mar 2023; ECP | PLATFORM FOR THE INFORMATION SOCIETY. *Op. cit.*, p. 42-43.; ANDRADE, Norberto Nuno Gomes de; KONTSCIEDER, Verena. *Op. cit.*, p. 81.



Também é necessário explicar como o sistema de IA irá funcionar, indicando a sua **lógica de operação**. Essa informação reflete o direito de informação e o consequente requisito de explicabilidade *ex ante* de decisões automatizadas,<sup>85</sup> comumente encontrado em leis de proteção de dados.<sup>86</sup> Nesse sentido, como lógica de operação, o responsável pela condução da AIA deve concentrar-se em especificar quais técnicas e tecnologias que serão utilizadas,<sup>87</sup> explicar o funcionamento e as funcionalidades do sistema de IA e indicar as suas respectivas consequências esperadas.<sup>88</sup>

Os dados ocupam uma parte essencial dos sistemas de IA, por isso, deve-se indicar qual **tipo de dado** e informação serão utilizados, se pessoais ou não pessoais, bem como descrever como se deu o treinamento.<sup>89</sup> Caso sejam utilizados dados pessoais para treinamento de um modelo, é importante a indicação da sua respectiva base legal, de acordo com a LGPD.<sup>90</sup> É importante notar que o uso de dados sensíveis ou de dados referentes a crianças e adolescentes é uma condição que exige maiores cuidados.<sup>91</sup>

Isso porque os dados sensíveis, nos termos do art. 5º, II, da LGPD, são aqueles que se referem à origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de caráter religioso, filosófico ou político, à saúde ou vida sexual e genético ou biométrico, razão pela qual podem expor intimamente o titular e submeter-lhe a discriminações. No mesmo sentido, os dados referentes a crianças e adolescentes merecem cautela por dizerem respeito a titulares em condição de vulnerabilidade e desenvolvimento de personalidade, devendo o tratamento ser feito em atenção ao seu melhor interesse.

<sup>85</sup> É importante fazer uma distinção entre explicabilidade *ex ante* e *ex post*: "Uma explicação *ex ante* é aquela que ocorre antes de uma tomada de decisão automatizada. Isso posto, uma decisão *ex ante* somente pode fazer referência às funcionalidades de um sistema de IA. Por outro lado, uma explicação *ex post* é aquela que é apresentada após a realização de uma decisão automatizada. Assim sendo, uma explicação *ex post* pode referir-se tanto às funcionalidades de um sistema quanto à lógica e as razões individuais de uma decisão em si." LEAL, Victor Moreira Mulin. **O Regulamento Geral sobre Proteção de Dados como instrumento de regulação de uma Inteligência Artificial de Confiança à luz das orientações éticas da Comissão Europeia**. Universidade do Minho. Dissertação de Mestrado em Direito e Informática, 2021.

<sup>86</sup> A título exemplificativo: Regulamento Geral de Proteção de Dados europeu, artigos 132(f) e 14, 2(g) e Lei Geral de Proteção de Dados (LGPD), art. 6º, IV e VI, art. 19 e art. 20.

<sup>87</sup> ECP | PLATFORM FOR THE INFORMATION SOCIETY. *Op. cit.*, p. 44.

<sup>88</sup> WACHTER, Sandra; MITTELSTADT, Brent; FLORIDI, Luciano. *Op. cit.*, p. 76-99, p. 15.

<sup>89</sup> ANDRADE, Norberto Nuno Gomes de; KONTSCHIEDER, Verena. *Op. cit.*, p. 78.

<sup>90</sup> De acordo com o artigo 7º, da LGPD. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2015-2018/2018/lei/l13709.htm](http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm). Acesso em: 31 jul 2021.

<sup>91</sup> REINO UNIDO. Information Commissioner's Officer - ICO. *Op. cit.*, p. 17. D



A **compreensão sobre os sujeitos que serão afetados pela aplicação**, seja de forma direta, indireta, intencional ou não intencional, é parte imprescindível para posterior identificação dos possíveis riscos e impactos negativos provocados pelo sistema de IA. Além disso, o responsável pela AIA deve ter consideração especial caso os sujeitos afetados sejam crianças, populações vulneráveis e populações que precisam de tecnologias assistivas.<sup>92</sup>

Deve-se também indicar **quais são os benefícios do uso do sistema de IA** por meio de uma perspectiva tripla, ou seja, do ponto de vista do desenvolvedor/aplicador da tecnologia, do sujeito afetado pela decisão automatizada e, quando aplicável, da própria sociedade como um todo ou grupos específicos que também podem ser impactados com o sistema.<sup>93</sup>

Por fim, é interessante que se destaque, desde o início da fase de design, **quais as leis e regulamentos específicos o sistema de IA deverá respeitar**. Se houver algum elemento internacional ao longo do processo de desenvolvimento e aplicação da tecnologia, é importante também estar atento aos regulamentos, legislações e tratados e acordos internacionais. Isso evitará surpresas que podem fazer com que os desenvolvedores e aplicadores da tecnologia onerem ainda mais o projeto, tanto em tempo quanto em recursos financeiros.

A título exemplificativo,<sup>94</sup> na hipótese de um sistema de moderação de conteúdo, **(i) o contexto** de utilização seria uma rede social de uso amplo; **(ii) a finalidade** seria o compartilhamento e a transmissão de informações; a **(iii) lógica de operação** irá depender do uso da tecnologia aplicada, contendo, por exemplo, a explicação acerca do uso do machine learning e seus resultados esperados; **(iv) os tipos de dados utilizados** seriam os dados pessoais, indicando-se a base normativa da LGPD e suas hipóteses normativas aplicáveis; **(v) os sujeitos afetados** pela aplicação seriam aqueles usuários cadastrados na plataforma, podendo também atingir sujeitos que não estão cadastrados, sendo pertinente indicar as hipóteses de riscos envolvidos para cada grupo; **(vi) os benefícios do sistema** podem ser compreendidos como a coibição do discurso de ódio e a tutela de sujeitos vulneráveis que são ameaçados pelo abuso da liberdade de expressão; **(vii) e as leis e os regulamentos específicos** conterão os tratados internacionais aplicáveis, a política da rede social e a legislação nacional, tais como o Marco Civil da Internet, a Lei Geral de Proteção de Dados, o Decreto n. 7962/2013 sobre o comércio eletrônico, o Código Civil, a Constituição Federal, entre outros.

<sup>92</sup> IEEE STANDARDS ASSOCIATION - IEEE SA. IEEE 7010-2020 - **IEEE Recommended practice for assessing the impact of autonomous and intelligent systems on human well-being**, p. 26. Disponível em: <https://standards.ieee.org/standard/7010-2020.html>. Acesso em: 31 jul 2021.

<sup>93</sup> ECP | PLATFORM FOR THE INFORMATION SOCIETY. *Op. cit.*, p. 46-47.

<sup>94</sup> O exemplo é apenas um ponto de partida para a elaboração de uma AIA, sendo imprescindível a complementação e detalhamento das ideias preliminares apontadas, com descrição detalhada de seus aspectos e funcionamento.

## Etapa 2. Avaliação da legalidade da finalidade do sistema de IA

Com base nas informações já apresentadas, é importante que se identifique a legalidade ou ilegalidade da finalidade do sistema e da maneira pela qual essa finalidade será alcançada.<sup>95</sup> Essa análise pode recair tão somente na questão legal, o que inclui o Código de Defesa do Consumidor, a Lei de Proteção de Dados Pessoais e leis e regulamentações específicas de cada setor, como saúde, trabalho, financeiro ou educação, por exemplo. Contudo, recomenda-se fortemente que também se realize uma avaliação ética do sistema de IA. Para isso, pode-se levar em consideração quadros de princípios éticos internacionalmente reconhecidos, como o da Organização para a Cooperação e Desenvolvimento Econômico (OCDE),<sup>96</sup> da Comissão Europeia,<sup>97</sup> entre outros.

### QUADRO DE PRINCÍPIOS ÉTICOS DA OCDE <sup>98</sup>

Crescimento inclusivo, desenvolvimento sustentável e bem-estar

Valores centrados no ser humano e equidade

Transparência e explicabilidade

Robustez, segurança e proteção

Responsabilidade e prestação de contas<sup>99</sup>

Na hipótese de um sistema de moderação de conteúdo, por exemplo, verifica-se que a finalidade de transmissão e compartilhamento de informações são práticas lícitas, devendo observar as diretrizes da Lei Geral de Proteção de Dados e os princípios éticos que disciplinam a matéria.

<sup>95</sup> *Ibidem*, p. 48.

<sup>96</sup> ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **Recommendation of the Council on Artificial Intelligence** (2019). **OECD/LEGAL/0449**. Disponível em: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>. Acesso em: 20 mai 2022.

<sup>97</sup> UNIÃO EUROPEIA. Grupo de Peritos de Alto Nível sobre a Inteligência Artificial. **Orientações éticas para uma IA de confiança**. Disponível em: <https://op.europa.eu/pt/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1/language-pt>. Acesso em: 31 jul 2021.

<sup>98</sup> OECD. *Op. cit*, p. 7-8.

<sup>99</sup> Tradução brasileira do termo em inglês “*accountability*”.



### Etapa 3. Avaliação da necessidade de participação de terceiros

Poderão existir casos em que a amplitude e a finalidade de um sistema de IA desperte um interesse na sociedade e em outros agentes externos à organização. Por isso, sugere-se que se analise a possibilidade de participação da sociedade civil, da academia, do poder público e dos próprios sujeitos afetados na realização da AIA, mesmo que de forma restrita ou somente opinativa.<sup>100</sup>

Ainda, é recomendável que a avaliação de impacto conte com especialistas que estudam como a compreensão de raça, gênero e outros marcadores sociais estão ligadas às desigualdades dos sistemas.<sup>101</sup> Caso o responsável pela AIA entenda ser adequada a participação de terceiros, é importante que se documente as opiniões e manifestações apresentadas. Por outro lado, se a conclusão for contrária à participação externa, recomenda-se que essa decisão seja devidamente fundamentada e documentada junto com a AIA.

Por fim, considerando a complexidade dos efeitos e dos resultados que alguns sistemas de IA podem causar no dia a dia da sociedade, caso seja possível, sugere-se que a execução da AIA tenha o apoio e o aconselhamento de especialistas independentes de áreas como direito, tecnologia da informação, sociologia e/ou ética.<sup>102</sup>

Na hipótese de um sistema de moderação de conteúdo, por exemplo, recomenda-se o apoio de especialistas jurídicos, de tecnologia e ética, além do amplo debate com a sociedade civil e com os sujeitos afetados pela adoção do sistema, isto é, os usuários.

<sup>100</sup> GRUPO DE TRABALHO DO ARTIGO 29 PARA A PROTEÇÃO DE DADOS. **Orientações relativas à Avaliação de Impacto sobre a Proteção de Dados (AIPD) e que determinam se o tratamento é «susceptível de resultar num elevado risco» para efeitos do Regulamento (UE) 2016/679**, abril de 2017, p. 8. Disponível em: [https://www.cnpd.pt/media/f0ide5i0/aipd\\_wp248rev-01\\_pt.pdf](https://www.cnpd.pt/media/f0ide5i0/aipd_wp248rev-01_pt.pdf). Acesso em: 31 jul 2021.

<sup>101</sup> DATA&SOCIETY. **Assembling Accountability: Algorithmic Impact Assessment for the Public Interest**. Disponível em: <https://datasociety.net/wp-content/uploads/2021/06/Assembling-Accountability.pdf>. Acesso em: 10 fev. 2022.

<sup>102</sup> REINO UNIDO. Information Commissioner's Officer - ICO. *Op. cit.*, p. 14.



## Etapa 4. Identificação e avaliação dos riscos envolvidos

Determinar os valores e os interesses que devem ser protegidos é a principal atividade para viabilizar a identificação dos riscos. **Os valores definidos representarão a amplitude da avaliação de impacto**, de modo que se pode orientar uma AIA através de uma **abordagem legal e/ou baseada em princípios éticos**.

É fortemente aconselhado que se adote uma metodologia em que os **direitos fundamentais** sejam reconhecidos **como os valores objeto de proteção**,<sup>103</sup> identificando, a partir deles, os riscos aos sujeitos afetados por um sistema de IA.

Nessa perspectiva, pode-se citar, **de forma não exaustiva**, ao menos **quinze valores** que podem sofrer efeitos negativos provenientes de sistemas de IA, dentre eles:

|                       |                         |                                 |
|-----------------------|-------------------------|---------------------------------|
| Autonomia Humana      | Liberdade de Expressão  | Estado de Direito               |
| Dignidade Humana      | Liberdade de Reunião    | Devido Processo Legal           |
| Igualdade e Equidade  | Liberdade de Pensamento | Privacidade e Proteção de Dados |
| Saúde Física e Mental | Emprego                 | Asilo                           |
| Educação              | Democracia              | Meio Ambiente                   |

A depender da finalidade do sistema de IA, e dos respectivos dados utilizados, estes são alguns valores que podem sofrer efeitos negativos relacionados às decisões automatizadas. Nesse sentido, ao perceber que um desses valores está em risco, o responsável pela condução da AIA deverá identificar, respectivamente e de forma concreta, os possíveis danos/impactos negativos que o sistema de IA pode causar ao determinado valor.

<sup>103</sup> Getting the future right – **Artificial intelligence and fundamental rights**. European Union Agency for Fundamental Rights. 2020, p. 87. Disponível em: <https://fra.europa.eu/en/publication/2020/artificial-intelligence-and-fundamental-rights>; JANSSEN, Heleen. Detecting new approaches for a fundamental rights impact assessment to automated decision-making. **International Data Privacy Law**. Oxford, v. 10, n. 1, p. 76-106, fevereiro 2020. Disponível em: <https://doi.org/10.1093/idpl/ipz028>. Acesso em: 31 jul 2021.

De modo a facilitar o entendimento, podemos citar alguns exemplos<sup>104</sup> que indicam as categorias de possíveis danos que podem afetar o sujeito ou grupo:

(I) **danos econômicos;**

(II) **danos psicológicos;** e,

(III) **danos coletivos.**

Nesse ponto, **sistemas de recrutamento** podem, conscientemente ou não, avaliar e tratar grupos específicos de forma diferente ou discriminatória, acarretando danos como perdas de oportunidade econômica e desenvolvimento profissional. Isso ocorreu, por exemplo, em ferramentas de recrutamento que discriminavam mulheres na avaliação de seleções de emprego, replicando em massa discriminações de gênero.<sup>105</sup> Essa prática pode ferir valores como equidade, dignidade humana e emprego.

Dados utilizados para **perfilar indivíduos** e categorizá-los em grupos específicos também podem acarretar danos econômicos no acesso diferenciado a bens e serviços de ordem pública ou privada, ferindo valores como equidade, dignidade humana, trabalho, educação, privacidade e proteção de dados. Demonstra-se tal situação, por exemplo, nos sistemas de **credit score** que permite e limita o seu acesso a serviços e direitos, como concessão de empréstimos, financiamentos e até mesmo aposentadoria.<sup>106</sup>

Na esfera psicológica, o aumento da **vigilância na vida pessoal** pode acarretar o risco da autocensura, ferindo valores como autonomia e dignidade humana, liberdade de reunião, privacidade e proteção de dados. A percepção da exposição à vigilância e a intromissão do Estado e de empresas privadas em espaços de confiança que nos permitem ser nós mesmos causa medo, podendo fazer com que as pessoas passem a se autocensurar com receio de invasão da sua privacidade e intimidade.<sup>107</sup>

<sup>104</sup> Inspiradas no projeto Open Loop. ANDRADE, Norberto Nuno Gomes de; KONTSCIEDER, Verena. *Op. cit.*

<sup>105</sup> ELIS, Nick. Ferramenta de recrutamento da Amazon com AI discriminava candidatas mulheres. **Meio Bit**. Disponível em: <https://tecnoblog.net/meiobit/391571/ferramenta-de-recrutamento-amazon-ai-discriminava-mulheres/>. Acesso em: 06 mar 2022.

<sup>106</sup> VINAYAK, Vrinda. **The Human Rights Implications of China's Social Credit System**. Oxford Human Rights Hub. Disponível em: <https://ohrh.law.ox.ac.uk/the-human-rights-implications-of-chinas-social-credit-system/>. Acesso em: 04 abr 2022.

<sup>107</sup> TOSAS, Gina. **Autocensura, la consecuencia a la vigilancia masiva**. La Vanguardia. Disponível em: <https://www.lavanguardia.com/tecnologia/internet/20141107/54419653179/autocensura-consecuencia-vigilancia-masiva.html>. Acesso em: 04 abr 2022.

Ademais, a **delegação de decisões humanas aos sistemas automatizados** traz o risco de perda do controle, ferindo valores como autonomia e dignidade humana, colocando em xeque a independência, a privacidade e o poder sobre a escolha das pessoas.<sup>108</sup> Com base em conhecimentos específicos, um “grupo-alvo” pode ser influenciado sem o seu conhecimento/consentimento a agir de uma determinada forma, evidenciando o risco de manipulação e ferindo valores como autonomia humana, equidade, privacidade e proteção de dados. Como exemplo, aponta-se a utilização de **microtargeting** baseado em uma economia comportamental para que uma empresa identifique vulnerabilidades em um grupo, valendo-se dessa informação para oferecer produtos e serviços em momentos em que os indivíduos estão mais propensos a uma tomada de decisão favorável à empresa.

Na esfera coletiva da sociedade, os **filtros ou moderação de conteúdo** e informação com base na perfilização de grupos de indivíduos pode acarretar a polarização e divisão da sociedade com redução do efetivo exercício das liberdades de opinião e acesso à informação, ferindo valores como democracia, Estado de Direito e liberdade de opinião. Como exemplo, ressalta-se a hiper perfilização de grupos de indivíduos e a utilização de **microtargeting** com objetivo político de fazer **marketing** negativo sobre um candidato adversário, visando deslegitimá-lo ou criar situações que desestimulem o seu eleitorado.<sup>109</sup>

Ressalta-se também o aumento da capacidade de vigilância do Estado mediante perfilização de indivíduos e uso de **reconhecimento facial**, ensejando danos no que tange à liberdade e autonomia, invasão de privacidade e *chilling effects*. Essa prática pode ferir valores como liberdade de reunião, autonomia humana, privacidade, proteção de dados, devido processo legal e asilo. Esses danos são verificados especialmente no uso de sistemas de reconhecimento facial, não somente para identificar um indivíduo, mas também para distinguir emoções.<sup>110</sup>

Por fim, também ressalta-se a **falta de transparência e de explicabilidade em decisões tomadas exclusivamente de forma automatizada**, que pode trazer a impossibilidade de compreensão das razões e justificativas de decisão tomada por um sistema automatizado utilizado pelo setor público, ferindo valores como equidade e devido processo legal.

<sup>108</sup> ANDERSON, Janna e RAINIE, Lee. **Artificial Intelligence and the Future of Humans**. Pew Research Center. Disponível em: <https://www.pewresearch.org/internet/2018/12/10/artificial-intelligence-and-the-future-of-humans/>. Acesso 04 abr 2022.

<sup>109</sup> Época Negócios. O fim da privacidade: como o micro-targeting mudou a política (2017). Disponível em: <https://epocanegocios.globo.com/Tecnologia/noticia/2017/06/o-fim-da-privacidade-como-o-micro-targeting-mudou-politica.html>. Acesso em: 04 abr 2022.

<sup>110</sup> COECKELBERGH, Mark. **Ética da IA**. MIT Pressione. Edição Kindle, p. 4-5. Acesso em: 04 abr 2022.



Como exemplo desse ponto, o uso de um algoritmo no setor da justiça criminal americano chamado COMPAS, realizava uma avaliação de risco sobre cada condenado com a finalidade de prever a probabilidade de reincidência criminal desse indivíduo, interferindo diretamente na possibilidade de liberdade condicional sem que houvesse qualquer intervenção humana ou que apresentasse justificativa ou fundamentação para a decisão.<sup>111</sup> Assim, identificado, de forma concreta, os possíveis danos/impactos negativos que o sistema de IA pode causar a um determinado valor a ser protegido, **o responsável pela AIA deverá avaliar a respectiva probabilidade de ocorrência e de gravidade do dano caso tal evento adverso efetivamente ocorra.**<sup>112</sup>

Vale ainda lembrar da existência de possíveis impactos negativos não intencionais, especialmente pela possibilidade de um sistema de IA desenvolvido para uma tarefa ou finalidade específica acabar sendo utilizado em situações diferentes daquelas originalmente previstas em seu projeto. Um exemplo clássico é o do **robô Gaak**, desenvolvido na Inglaterra, no *Magna Science Center*. Tratou-se de experimento que atribuía aos robôs os papéis de “caçador” e “presa”, colocando-os em uma arena apenas para que, respectivamente, caçassem e fugissem.

O experimento visava verificar a aplicabilidade do princípio da sobrevivência do mais apto aos robôs dotados de inteligência artificial, bem como verificar se eles poderiam se beneficiar do conhecimento adquirido. O robô, no entanto, fora deixado sem vigilância por 15 minutos, conseguindo fugir da arena, atravessar o muro da sede e encontrar uma saída, sendo posteriormente atingido por um carro no estacionamento. O detalhe é que a sua programação não abrangia toda essa atividade.<sup>113</sup> Por isso, **recomenda-se que o responsável pela condução da AIA esteja atento para possíveis riscos e impactos negativos não intencionais.**

Para que se alcance o máximo proveito desta etapa, é aconselhável que seu estudo envolva um juízo de previsões, elaborações de hipóteses e projeção, e utilização de possíveis cenários.<sup>114</sup> Destaca-se também **que esta etapa deve ser conduzida pelo ponto de vista do sujeito que sofrerá os efeitos da decisão**, e não pela visão do desenvolvedor ou aplicador da tecnologia.

<sup>111</sup> ANGWIN, J. et al. **Machine Bias**. ProPublica, 2016. Disponível em: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Acesso em: 04 abr 2022.

<sup>112</sup> Institute of Electrical and Electronics Engineers (IEEE). **Recommended practice for assessing the impact of autonomous and intelligent systems on human well-being**. IEEE 7010-2020, p. 22. Disponível em: <https://standards.ieee.org/standard/7010-2020.html>. Acesso em: 31 jul 2021; ANDRADE, Norberto Nuno Gomes de; KONTSCHIEDER, Verena. *Op. cit.*, p. 79-82.

<sup>113</sup> THE GUARDIAN. **Robot fails to find a place in the sun (2002)**. Disponível em: <https://www.theguardian.com/uk/2002/jun/20/engineering.highereducation>. Acesso em: 08 fev 2022.

<sup>114</sup> IEEE. *Op. cit.* p. 21.

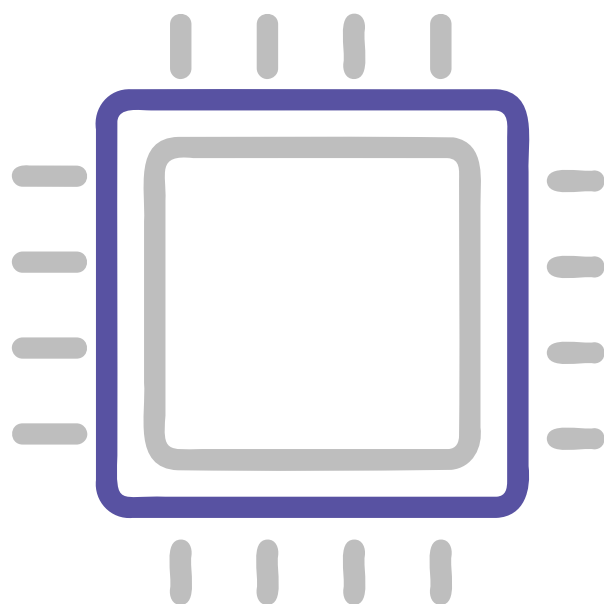


## Etapa 5. Definição de salvaguardas, medidas de segurança e medidas de proteção dos direitos e liberdades dos sujeitos

A fim de definir medidas adequadas a mitigar ou eliminar os riscos, o responsável pela condução da AIA deverá, primeiramente, registrar a fonte de cada risco identificado na etapa anterior. A partir do momento em que se identifica o risco e sua respectiva origem, deve-se analisar e implementar medidas adequadas para sua eliminação ou mitigação a um nível aceitável.

A depender do caso específico, podem existir inúmeras intervenções para mitigar os riscos identificados. A escolha da medida mais adequada recai sobre a própria equipe responsável pelo desenvolvimento ou operação do sistema. Por isso, **é imprescindível que essa equipe tenha uma composição multidisciplinar e multicultural**, de modo que diversos pontos de vista possam ser levados em consideração.

Um exemplo pode ser ilustrado através de um sistema de IA de moderação de conteúdo que remove automaticamente publicações reconhecidas como discurso de ódio em redes sociais. Essa atividade pode trazer um risco à liberdade de expressão. Apesar de a proposta ser benéfica para a sociedade, o sistema de IA poderá cometer erros e remover injustamente o conteúdo publicado. Por isso, é importante que o responsável pela condução da AIA esteja atento para salvaguardas que possam reduzir esses erros, bem como às medidas para remediar o dano causado. Neste caso, uma medida de salvaguarda seria a revisão da decisão de remoção do conteúdo por um agente humano.



## Etapa 6. Avaliação da proporcionalidade entre finalidade e resultado

A presente etapa é uma inspiração lógica do **princípio legal da proporcionalidade**, onde a sua essência está diretamente relacionada com valores e ideais como **justiça, equidade, bom senso, prudência, moderação, justa medida e proibição de excesso**.<sup>115</sup>

Por mais que este seja um princípio frequentemente evocado na busca pelo equilíbrio entre poderes, privilégios, benefícios e restrições, a proporcionalidade também pode ser aplicada a fim de promover certo equilíbrio na relação entre desenvolvedores e aplicadores de sistemas de IA e os respectivos direitos e interesses dos sujeitos afetados por essa tecnologia.

Nesse contexto, a avaliação da proporcionalidade<sup>116</sup> irá exigir que responsável pela AIA identifique **(i)** se a forma como o sistema de IA foi planejado definitivamente possibilita e contribui para o alcance dos propósitos pretendidos; e **(ii)** se não há um meio alternativo, razoável e menos intrusivo de se conseguir os mesmos resultados.<sup>117</sup>

No caso de um sistema de moderação de conteúdo, a avaliação da proporcionalidade recairá sobre a relação entre a restrição da liberdade de expressão dos usuários e os direitos de personalidade de eventuais vítimas atingidas pelo discurso combatido.



<sup>115</sup> LENZA, Pedro. **Direito constitucional esquematizado**. 12ª ed. rev., atual. e amp. São Paulo: Saraiva, 2008, p. 75.

<sup>116</sup> A busca pela proporcionalidade pode ser encontrada no Regulamento Geral Proteção de Dados da União Europeia, em seu artigo 35(7)(b).

<sup>117</sup> ECP | PLATFORM FOR THE INFORMATION SOCIETY. *Op. cit.*, p. 47.

## Etapa 7. Riscos residuais e a necessidade de consulta prévia a uma Autoridade de Supervisão

O sucesso na condução de uma AIA não passa, necessariamente, pela completa eliminação de todo e qualquer risco proveniente de um sistema de IA. Apesar deste resultado ser o melhor possível, a realização de uma avaliação de impacto também deve ser orientada a reduzir os riscos e garantir que o impacto negativo desses efeitos seja minimizado a um nível aceitável.<sup>118</sup> Por isso, mesmo após identificar os riscos, definir salvaguardas e implementar medidas de proteção, **é importante perceber se ainda assim existem riscos residuais que possam ser causados pelo sistema de IA.**<sup>119</sup>

Quando riscos residuais forem considerados de nível aceitável,<sup>120</sup> o responsável pela condução da AIA deverá identificá-los, fundamentar por que o risco residual é aceitável e indicar quais medidas serão tomadas para limitar e/ou reparar possíveis danos se o risco se manifestar.<sup>121</sup>

No caso de um sistema de moderação, podem restar riscos residuais atinentes a má interpretação do algoritmo de remoção de conteúdo, que podem ser mitigados com o auxílio de supervisão humana. Caso o risco residual ainda apresente em um alto risco para os direitos e liberdades dos sujeitos afetados, não podendo, portanto, ser mitigado, além de ser devidamente identificado na AIA, **é recomendável que o responsável pela condução da AIA consulte previamente uma autoridade supervisora**<sup>122</sup> **antes de implementar e operar o respectivo sistema.**<sup>123</sup>

<sup>118</sup> REINO UNIDO. Information Commissioner's Officer - ICO. **Data protection impact assessment.** Versão 1.0.124, de maio de 2018, p. 14. Disponível em: <https://ico.org.uk/media/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr-1-1.pdf>. Acesso em: 31 jul 2021.

<sup>119</sup> ECP | PLATFORM FOR THE INFORMATION SOCIETY. *Op. cit.*, p. 57-58.

<sup>120</sup> Não há uma definição clara e objetiva do que será considerado risco aceitável. Por se tratar de uma avaliação procedimental, cada responsável pela AIA chegará a sua própria conclusão sobre o que é ou não um elevado risco, assim como o que será um risco aceitável. Como exemplo pode-se citar a tecnologia de reconhecimento facial. Isso porque, enquanto empresas como a IBM reconheceram a gravidade dos riscos atrelados ao uso dessa tecnologia, e por isso interromperam o seu desenvolvimento, outras empresas continuam a apostar comercialmente em sua expansão.

<sup>121</sup> ECP | PLATFORM FOR THE INFORMATION SOCIETY. *Op. cit.* 57-58.

<sup>122</sup> No que tange à autoridade reguladora da IA, consideramos necessário existir um regulador central que trabalhe como coordenador das tantas agências reguladoras setoriais que dividem competências sobre a miríade de setores em que se aplica IA, de modo a evitar decisões regulatórias difusas e por vezes contraditórias sobre esse tema. Essa autoridade terá, como competências primárias, a conciliação desses atores, bem como a interpretação e o desenho de políticas comuns a setores diversos, o que inclui tanto atores governamentais quanto entidades auditoras, certificadoras e de padronização. Não havendo tal entidade, indica-se subsidiariamente a autoridade de fiscalização setorial.

<sup>123</sup> ANDRADE, Norberto Nuno Gomes de; KONTSCIEDER, Verena. *Op. cit.*, p. 55.



## Etapa 8. Documentação e publicação

A documentação é considerada uma etapa de encerramento da AIA, onde são registradas as respostas das etapas 1 a 7, de modo a manter registro que **comprove a forma responsável que a empresa ou organização lida com a tecnologia.**

A documentação será útil caso o responsável pelo sistema de IA precise apresentar explicações relativas ao modo de funcionamento ou aos possíveis efeitos provocados pela respectiva aplicação.

Sobre a publicação, o art. 26 da minuta do substitutivo aos Projetos de Lei nº. 5.051/2019, 21/2020, e 872/2021 dispõe que, garantidos os segredos industrial e comercial, as conclusões da avaliação de impacto serão públicas e conterão ao menos as seguintes informações: I - descrição da finalidade pretendida para a qual o sistema será utilizado, assim como de seu contexto de uso e escopo territorial e temporal; II - medidas de mitigação dos riscos, bem como o seu patamar residual, uma vez implementada tais medidas; III - descrição da participação de diferentes segmentos afetados, caso tenha ocorrido.

Esse é o momento em que se deve deliberar sobre a publicação de um resumo da AIA, protegendo informações sigilosas e segredos comerciais. Essa decisão é recomendável no sentido de demonstrar **maior responsabilidade social** e criar um sentimento de **confiança e transparência** com parceiros, clientes e utilizadores.



## CONCLUSÃO

Avaliações, relatórios e diagnósticos do impacto são instrumentos que têm ganhado cada vez mais importância em uma sociedade na qual as ações humanas e empresariais podem provocar riscos de difícil ou impossível reparação. Neste contexto, a avaliação de impacto algorítmico é um instrumento de governança que pode ser adotado pelo responsável pelo desenvolvimento e operação de um sistema de IA para avaliar, documentar e prestar contas de determinadas aplicações que possam causar um alto risco para os direitos das pessoas físicas.

A AIA pode ser feita por meio de algumas etapas, caracterizadas pela descrição detalhada do sistema, avaliação da legalidade da finalidade do sistema, avaliação da necessidade de consulta a terceiros, identificação e avaliação dos riscos envolvidos, definição de salvaguardas, medidas de segurança e de proteção dos direitos e liberdades, avaliação da proporcionalidade entre finalidade e resultado, riscos residuais e necessidade de consulta a uma autoridade de supervisão, documentação e publicação.

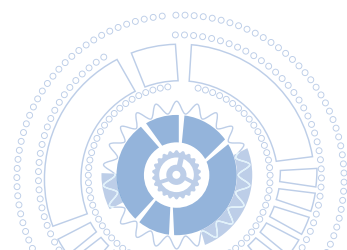
A relevância desse instrumento foi reconhecida no debate regulatório brasileiro, especialmente a partir de sua previsão contida na minuta de substitutivo recentemente proposta pela Comissão de Juristas do Senado Federal. Assim, torna-se necessário acompanhar o debate e os rumos tomados pela futura autoridade competente no que tange à essa regulamentação no Brasil.<sup>124</sup> De toda forma, os apontamentos sobre a avaliação de impacto no presente relatório são expectativas de destaque no cenário que podem auxiliar no desenvolvimento desse relevante instrumento, contribuindo na definição de salvaguardas para eliminar ou mitigar riscos associados ao desenvolvimento ou operação de sistemas de IA e para a consolidação desses sistemas na sociedade de uma forma responsável.

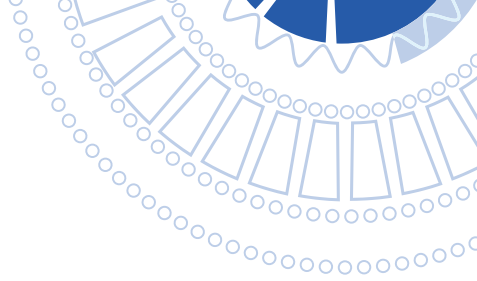
<sup>124</sup> A minuta do substitutivo da Comissão de Juristas indica que a autoridade competente será o órgão ou entidade da Administração Pública Federal, a ser designada pelo Poder Executivo, responsável por zelar, implementar e fiscalizar o cumprimento da futura lei para inteligência artificial no Brasil. Relatório Legislativo, Relatório Final aprovado em 06 de dezembro de 2022. Disponível em: <https://legis.senado.leg.br/comissoes/comissao?codcol=2504>. Acesso em: 23 jan 2023.

## ANEXO I: MODELO VISUAL DA AIA

| <b>ETAPA 1</b><br><b>DESCRIÇÃO</b><br><b>DETALHADA DO</b><br><b>SISTEMA DE IA</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | <b>ETAPA 2</b><br><b>AValiação DA</b><br><b>LEGALIDADE DA</b><br><b>FINALIDADE DO</b><br><b>SISTEMA DE IA</b>                                                                                                                                                                                                  | <b>ETAPA 3</b><br><b>AValiação DA</b><br><b>NECESSIDADE DE</b><br><b>PARTICIPAÇÃO DE</b><br><b>TERCEIROS</b>                                                                                                                                                                                                                                                                                                                                                                 | <b>ETAPA 4</b><br><b>IDENTIFICAÇÃO E</b><br><b>AValiação DOS RISCOS</b><br><b>ENVOLVIDOS</b>                                                                                                                                                                                                                                                                                                      |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Descreva:</p> <ul style="list-style-type: none"> <li>- qual o <b>contexto</b> no qual o sistema de IA será utilizado;</li> <li>- qual a respectiva <b>finalidade</b> do sistema de IA;</li> <li>- qual a <b>lógica de operação</b> do sistema de IA;</li> <li>- qual <b>tipo de dado</b> será utilizado e, em caso de dados pessoais, apresente a respectiva <b>base legal</b>;</li> <li>- como ocorreu o <b>treinamento</b> dos dados;</li> <li>- quais são os <b>sujeitos afetados</b> de forma direta, indireta, intencional ou não intencional pelo sistema de IA;</li> <li>- quais são os <b>benefícios</b> do sistema de IA;</li> <li>- quais as <b>leis e as regulações</b> específicas e setorializadas que o sistema de IA deve respeitar.</li> </ul> | <p>De acordo com o ordenamento jurídico em geral e as regulamentações específicas e setorializadas:</p> <ul style="list-style-type: none"> <li>- Indique se a <b>finalidade</b> do sistema de IA é <b>legal</b>;</li> <li>- Indique se a <b>lógica de operação</b> do sistema de IA é <b>legal</b>.</li> </ul> | <ul style="list-style-type: none"> <li>- Identifique se <b>amplitude e a finalidade</b> de um sistema de IA pode despertar o interesse da sociedade e de outros <b>agentes externos</b> à organização;</li> <li>- Delibere sobre a adequação de oportunizar a <b>participação da sociedade civil</b>, da <b>academia</b>, do <b>poder público</b> e dos próprios <b>sujeitos afetados</b> na realização da AIA, mesmo que de forma restrita ou somente opinativa.</li> </ul> | <ul style="list-style-type: none"> <li>- Identifique possíveis <b>riscos e impactos negativos</b> que o sistema de IA pode causar para os direitos fundamentais;<sup>125</sup></li> <li>- Avalie a <b>probabilidade de ocorrência</b> do risco ou impacto negativo;</li> <li>- Avalie a <b>gravidade do risco</b> ou do impacto negativo caso o dano aos direitos fundamentais ocorra.</li> </ul> |

<sup>125</sup> Exemplos de valores que podem ser considerados: Autonomia Humana; Dignidade Humana; Igualdade e Equidade; Saúde Física e Mental; Educação; Liberdade de Expressão; Liberdade de Reunião; Liberdade de Pensamento; Emprego; Democracia; Estado de Direito; Devido Processo Legal; Privacidade e Proteção de Dados; Asilo; e Meio Ambiente.





| ETAPA 5<br>DEFINIÇÃO DE<br>SALVAGUARDAS, MEDIDAS<br>DE SEGURANÇA E MEDIDAS<br>DE PROTEÇÃO DOS<br>DIREITOS E LIBERDADES<br>DOS SUJEITOS                                                                                                                                                                                                 | ETAPA 6<br>AVALIAÇÃO DA<br>PROPORCIONALIDADE<br>ENTRE FINALIDADE E<br>RESULTADO                                                                                                                                                                                                                                                                                                               | ETAPA 7<br>RISCOS RESIDUAIS<br>E A NECESSIDADE DE<br>CONSULTA PRÉVIA A<br>UMA AUTORIDADE DE<br>SUPERVISÃO                                                                                                                                                                                                                                                                                                                                                                                 | ETAPA 8<br>DOCUMENTAÇÃO E<br>PUBLICAÇÃO                                                                                                                                                                                                                 |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"><li>- Identifique e registre a <b>fonte de cada um dos riscos</b> anteriormente identificados;</li><li>- Identifique e implemente, respectivamente para cada um dos <b>riscos encontrados, salvaguardas e medidas adequadas</b> para a eliminar o risco ou mitigá-lo a um nível aceitável.</li></ul> | <ul style="list-style-type: none"><li>- Indique se a lógica de operação do sistema de IA, bem como se os dados e informações utilizados para o seu treinamento, definitivamente possibilitam e contribuem no <b>alcance das finalidades</b> pretendidas;</li><li>- Indique se não há um <b>meio alternativo</b>, razoável e menos intrusivo de se conseguir o resultado pretendido.</li></ul> | <p>Existem <b>riscos residuais</b>?<br/>Eles estão minimizados a um <b>nível aceitável</b>?</p> <ul style="list-style-type: none"><li>- Se <b>sim</b>, identifique-os, fundamente por que o risco residual é admissível e indique quais medidas serão tomadas para limitar e/ou reparar possíveis danos se o risco se manifestar;</li><li>- Se <b>não</b>, identifique-os e consulte previamente uma autoridade supervisora antes de implementar e operar o respectivo sistema.</li></ul> | <ul style="list-style-type: none"><li>- <b>Registre</b> as respostas das etapas 1 a 7;</li><li>- Delibere sobre a conveniência e oportunidade de se <b>publicar um resumo da AIA</b>, protegendo informações sigilosas e segredos comerciais.</li></ul> |



# LABORATÓRIO DE POLÍTICAS PÚBLICAS E INTERNET



ST SRTV/SUL 701, Cj. D, Bl. A, Sala 818, Pt.  
C, Ed. Centro Empresarial. Brasília (DF),  
70340-907

