

## As Diretrizes Éticas em IA

Após o estudo dos filtros constitucionais estrangeiros e a constatação da atual realidade do instituto da repercussão geral no STF no segundo capítulo, de estudar os conceitos primordiais dos mecanismos da inteligência artificial inerentes à área da Ciência da Computação no terceiro capítulo, assim como de identificar as aplicações dos sistemas de IA no meio jurídico. Agora, faz-se necessário analisar-se como os países desenvolvidos estão conduzindo a inserção dos sistemas de IA nos seus respectivos países, bem como o Brasil vem tratando o tema da ética na IA.

A ética, no uso da IA, é um assunto delicado e ultimamente muito tratado pelos pesquisadores, uma vez que muitos indagam: *poderá uma máquina substituir um ser humano em determinadas tomadas decisões da vida cotidiana?* Essa é a pergunta fundamental. Conforme evidências disponíveis, hoje, pode-se afirmar que – assim como a declaração da Conferência Internacional de Comissários de Proteção e Privacidade de Dados (ICDPPC) reconhece - o uso de Inteligência Artificial tem impacto não somente sobre os indivíduos, mas também impacta significativamente grupos da sociedade em geral.<sup>1</sup> Por isso, a ética é tema imprescindível a ser abordado e enfrentado em qualquer projeto de IA.

No tocante, em importantes documentos internacionais, a ética na (IA) já faz parte de um contexto real. A exemplo, o programa do Mercado Único Digital da União Europeia é importante para destacar a preocupação global sobre o tema. Em 2017, a Resolução do Parlamento Europeu recomendou à Comissão Europeia sobre disposições de direito civil incidentes sobre Robótica (2015/2013(INL)), destacando que as definições de “robô” e de “inteligência artificial” sejam flexíveis e não criem obstáculos à inovação.<sup>2</sup> A referida resolução

---

<sup>1</sup> INTERNATIONAL CONFERENCE OF DATA PROTECTION AND PRIVACY COMMISSIONERS, 41., 2019. **XLI International...** Tirana, Albânia: 2019. Disponível em: <https://privacyconference2019.info/>. Acesso em: 16 ago. 2020.

<sup>2</sup> “(...) a humanidade se encontra no limiar de uma era em que robôs, “bots”, androides e outras manifestações de inteligência artificial (IA), cada vez mais sofisticadas, parecem estar preparados para desencadear uma nova revolução industrial, que provavelmente não deixará nenhuma camada da sociedade intacta, é extremamente importante que o legislador pondere as suas implicações e os seus efeitos a nível jurídico e ético, sem pôr entaves à inovação”. UNIÃO EUROPEIA. Parlamento Europeu. **Disposições de Direito Civil sobre Robótica**. Resolução do Parlamento Europeu, de 16 de fevereiro de 2017, que contém recomendações à Comissão sobre disposições de Direito Civil sobre Robótica (2015/2103(INL)). Estrasburgo: Parlamento Europeu,

trata da importância de princípios éticos para uso de aplicações da IA no cenário atual,<sup>3</sup> que, indiscutivelmente, está na tangente entre o risco e o retorno, podendo criar efeitos ainda desconhecidos na área da segurança, da saúde,<sup>4</sup> da

---

16 de fevereiro de 2017. Disponível em: [https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051\\_PT.html](https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_PT.html). Acesso em: 16 ago. 2020.

<sup>3</sup> **Princípio do respeito dos direitos fundamentais:** assegurar que a concessão e a aplicação de instrumentos e serviços de inteligência artificial sejam compatíveis com os direitos fundamentais - **Princípio da não discriminação:** prevenir especificamente o desenvolvimento ou a intensificação de qualquer discriminação entre indivíduos ou grupos de indivíduos. - **Princípio da qualidade e da segurança:** No que respeita ao tratamento de decisões e dados judiciais, utilizar fontes certificadas e dados incorpóreos com modelos concebidos de forma multidisciplinar, num ambiente tecnológico seguro. - **Princípio da transparência, imparcialidade e equidade:** tornar os métodos de tratamento de dados acessíveis e compreensíveis, autorizar auditorias externas. - **Princípio "sob controle do usuário":** impedir uma abordagem prescritiva e garantir que os utilizadores sejam agentes informados e controlem as suas escolhas. UNIÃO EUROPEIA. Comissão Europeia para a Eficácia da Justiça (CEPEJ). Conselho da Europa. **Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente.** Estrasburgo: CEPEJ, 03 de dezembro de 2018. Disponível em:

[https://rm.coe.int/carta-etica-traduzida-para-portugues\\_revista/168093b7e0#\\_Toc530141214](https://rm.coe.int/carta-etica-traduzida-para-portugues_revista/168093b7e0#_Toc530141214). Acesso em: 16 ago. 2020. Grifo nosso.

<sup>4</sup> Os softwares aplicados na área da saúde possuem um desenvolvimento inteligente utilizando Sistemas Especialistas (SE). Não é de hoje que essa aplicação é feita, mas desde quando um dos primeiros sistemas foi desenvolvido, sendo ele o MYCIN. O sistema que marcou na aplicação da Inteligência Artificial na medicina foi o Sistema Especialista MYCIN. Sistema desenvolvido por Edward Shortliffe, onde propôs a construção de um "Especialista Artificial", um conselheiro em antibioticoterapia. Shortliffe desenvolve o sistema especialista como dissertação de doutorado, escrito na linguagem Lisp, tendo como usuários alvo: médicos, paramédicos e estudantes de medicina. Sendo o primeiro programa solucionador de problemas de raciocínio através de informações incertas ou incompletas. MYCIN diagnosticaria patologias com base nos sintomas do paciente e testes médicos, caso necessário, o programa solicitaria informações adicionais como exames laboratoriais, chegando a um provável resultado e recomendaria um tratamento apropriado. Caso necessário o MYCIN relatava como atingiu determinado resultado. **A segunda etapa** do desenvolvimento da IA na medicina ficou marcada pelo surgimento dos seguintes sistemas: EMYCIN, Expert e Age. Os sistemas foram desenvolvidos para aplicação em controle de instrumentos de UTI, interpretação de teste laboratorial para aconselhamento de diagnóstico e tratamento de doenças. De acordo com Westphal (2003) **o terceiro estágio** da IA na medicina período que marca a iniciativa de pesquisadores de IA no desenvolvimento de formas representativas para o raciocínio médico. Iniciando-se os conceitos de Redes Neurais Artificiais (RNA) para a aplicação na informática médica. GUARIZI, Débora D. OLIVEIRA, Eliane V. Estudo da Inteligência Artificial Aplicada na área da Saúde. **Colloquium Exactarum**, Presidente Prudente, v. 6, p. 26-37, jul./dez. 2014. Disponível em: <http://www.unoeste.br/site/enepe/2014/suplementos/area/Exactarum/Computa%C3%A7%C3%A3o/ESTUDO%20DA%20INTELIG%C3%8ANCIA%20ARTIFICIAL%20APLIACADA%20NA%20C3%81REA%20DA%20SA%C3%9ADE.pdf>. Acesso em: 16 ago. 2020.

liberdade,<sup>5</sup> da privacidade,<sup>6</sup> da integridade,<sup>7</sup> da dignidade, da autodeterminação,<sup>8</sup> da não discriminação<sup>9</sup> e da proteção de dados pessoais<sup>10</sup>.

---

<sup>5</sup> As decisões automatizadas baseiam em um método estatístico para análise de grande volume de dados e informações. Assim, conforme alertado por Lyon (2003), isso pode acarretar a perda de chances e oportunidades do indivíduo na sociedade, dar-se-á uma restrição indevida à sua autonomia, limitando a sua liberdade de ação, suas escolhas econômicas e até mesmo existenciais. LYON, David. **Surveillance as social sorting: privacy, risk, and digital discrimination**. Nova York: Routledge, 2003.

<sup>6</sup> A Cambridge Analytica (CA) aplicou técnicas psicométricas somadas ao cruzamento de dados geográficos sem autorização de milhões de usuários do Facebook para persuadi-los a votar de um determinado modo. A partir de um questionário respondido por alguns milhares de usuários do Facebook, a personalidade de cada um deles foi categorizada por uma das empresas ligadas a CA. Em seguida, passou a acompanhar a navegação no Facebook desses usuários. Utilizando algoritmos de aprendizagem de máquina (*machine learning*) estabeleceu um padrão de como cada tipo de personalidade clicava em determinados *links*, memes e postagens na rede social. A violação massiva de privacidade e do direito a não ser vigiado ocorre quando a CA ocorre quando os algoritmos de aprendizagem passam a acompanhar a navegação e a clicagem de mais de 70 milhões de norte-americanos com a finalidade de traçar o perfil psicométrico de cada um deles. CASSIANO, João Francisco; AVELINO, Rodolfo; SILVEIRA, Sérgio Amadeu. **Direitos Humanos, Inteligência Artificial e Privacidade. Monções: Revista de Relações Internacionais da UFGD**, Dourados, v. 8, n. 15, p. 573-596, jan./jun. 2019. Disponível em: [https://www.researchgate.net/publication/341060413\\_Direitos\\_Humanos\\_inteligencia\\_artificial\\_e\\_privacidade](https://www.researchgate.net/publication/341060413_Direitos_Humanos_inteligencia_artificial_e_privacidade). Acesso em: 14 mar. 2020. p. 578.

<sup>7</sup> A integridade da IA guarda relação com a origem dos dados, uma vez que os dados devem ser verdadeiros sem manipulação de modo a desvirtuar sua origem, e assim vacular sua validade. Uma IA íntegra é o espelho, o retrato do passado dos dados que irão construir o aprendizado da máquina. Fonte: o autor.

<sup>8</sup> A dignidade e autodeterminação caminham lado a lado, e inclusive, há razões pelas quais os algoritmos podem produzir resultados discriminatórios e prejudicar determinadas pessoas ou grupos sociais dividem-se em dois grandes grupos. É preciso ficar atento à principal promessa que as decisões baseadas em algoritmos costumam suscitar: a obtenção de maior objetividade das decisões e a possibilidade de fundamentá-las em critérios estatísticos, eliminando o risco dos vieses ("*bias*") típicos das decisões tomadas pelos seres humanos. Ao contrário, grande parte da literatura tem demonstrado o risco de discriminação inerente aos processos de tomada de decisão baseados em algoritmos, a qual ocorre, muitas vezes, de forma involuntária e sem o conhecimento dos tomadores de decisão. BAROCAS, Solon; SELBST, Andrew D. Big Data's Disparate Impact. **California Law Review**, v. 104, n. 671, p. 671-731, 2015. Disponível em: <https://www.semanticscholar.org/paper/Big-Data%27s-Disparate-Impact-Barocas-Selbst/1d174f0e3c391368d0f3384a144a6c7487f2a143>. Acesso em: 14 mar. 2020. p. 673.

<sup>9</sup> O Princípio da não discriminação, busca prevenir especificamente o desenvolvimento ou a intensificação de qualquer discriminação entre indivíduos ou grupos de indivíduos. Dada a capacidade destes métodos de tratamento para revelar a discriminação existente, através do agrupamento ou da classificação de dados relativos a indivíduos ou grupos de indivíduos, os intervenientes públicos e privados devem garantir que os métodos não reproduzem ou agravam essa discriminação e que não conduzem a análises ou utilizações determinísticas. Deve ser dada especial atenção tanto na fase de desenvolvimento como na de implantação, especialmente quando o tratamento se baseia, direta ou indiretamente, em dados "sensíveis". Tal poderá incluir alegada origem racial ou étnica, antecedentes socioeconômicos, opiniões políticas, convicções religiosas ou filosóficas, filiação sindical, dados genéticos, dados biométricos, dados relativos à saúde ou dados relativos à vida sexual ou à orientação sexual. Quando essa discriminação tiver sido identificada, devem ser consideradas medidas corretivas para limitar ou, se possível, neutralizar esses riscos, bem como a sensibilização das partes interessadas. No entanto, a utilização da aprendizagem automática e de análises científicas multidisciplinares para combater esta discriminação deve ser incentivada. UNIÃO EUROPEIA. Comissão Europeia para a Eficácia da Justiça (CEPEJ). Conselho da Europa. **Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente**. Estrasburgo: CEPEJ, 03 de dezembro de 2018. Disponível em: [https://rm.coe.int/carta-etica-traduzida-para-portugues- revista/168093b7e0#\\_Toc530141214](https://rm.coe.int/carta-etica-traduzida-para-portugues- revista/168093b7e0#_Toc530141214). Acesso em: 16 ago. 2020.

A Resolução do Parlamento Europeu, ao propor uma atualização do arcabouço jurídico da União Europeia sobre disposições de direito civil incidentes na Robótica, na Carta de 4/12/2018 da Comissão Europeia para Eficiência da Justiça (CE-PEJ) no âmbito do Conselho da Europa, trouxe um código de conduta para engenheiros de robótica e comitês de ética de investigação quando analisam protocolos de robótica para os comitês de modelos para criadores e utilizadores.<sup>11</sup>

Outro documento importante na área da ética foi a comunicação da Comissão Europeia ao Parlamento Europeu, ao Conselho Europeu, ao Conselho, ao Comitê das Regiões: Inteligência artificial para a Europa. Na referida comunicação, foram definidos o conceito de inteligência artificial, o modo de aplicação e, principalmente, delinearam-se três pilares de observância para qualquer utilização em matéria de IA, que visam a: (i) reforçar a capacidade industrial e tecnológica da UE e a adoção da IA na economia, tanto no setor público quanto no setor privado incluindo investimentos em investigação e inovação e um melhor acesso aos dados; (ii) preparar a UE para as mudanças socioeconómicas decorrentes da IA, incentivando a modernização dos sistemas de educação e de formação, fomentando o talento, antecipando as mudanças no mercado de trabalho, apoiando as transições no mercado de trabalho e a adaptação dos sistemas de proteção social; (iii) garantir um quadro ético e jurídico apropriado, baseado nos valores da União e em consonância com a Carta dos Direitos Fundamentais da União Europeia. Isso inclui futuras

---

<sup>10</sup> É notório o posicionamento dominante de um punhado de empresas na oferta de serviços digitais, que se alimentam primordialmente de dados dos usuários, como Facebook, Google, Amazon, entre outros. Devido a constatação técnica de que para melhorar sua eficácia a IA necessita de grandes quantidades de dados, resulta preocupante o panorama atual do nível de domínio de mercado que têm essas empresas no concernente a exploração de dados pessoais. SATARIANO, Adam; ISAAC, Mike. Facebook Used People's Data to Favor Certain Partners and Punish Rivals, Document Show. *In*: THE NEW YORK TIMES. Nova York, 05 de dezembro de 2018. Disponível em: <https://www.nytimes.com/2018/12/05/technology/facebook-documents-uk-parliament.html>. Acesso em: 16 ago. 2019. FUNDACIÓN KARISMA. **Ética e proteção de dados na inteligência artificial**: a continuação do debate. Uma contribuição da América Latina e Caribe. Bogotá, Colômbia: Fundación Karisma, 15 de abril de 2019. Disponível em: <https://web.karisma.org.co/etica-e-protecao-de-dados-na-inteligencia-artificial-a-continuacao-do-debate-uma-contribuicao-da-america-latina-e-caribe/>. Acesso em: 16 ago. 2020.

<sup>11</sup> A Carta de 4/12/2018 da Comissão Europeia para Eficiência da Justiça (CE-PEJ) no âmbito do Conselho da Europa no mesmo sentido, apresentou a importância do código de conduta para outros comitês de outras áreas como saúde, transporte, educação, emprego e impactos ambientais. UNIÃO EUROPEIA. Parlamento Europeu. **Sobre o Painel de Avaliação da Justiça na EU de 2017 (2018/2009 (INI))**. Estrasburgo: Parlamento Europeu, 03 de maio de 2018. Disponível em: [https://www.europarl.europa.eu/doceo/document/A-8-2018-0161\\_PT.html](https://www.europarl.europa.eu/doceo/document/A-8-2018-0161_PT.html). Acesso em: 16 ago. 2020.

orientações sobre as regras de responsabilidade decorrente dos produtos em vigor, uma análise pormenorizada dos desafios emergentes e a cooperação com as partes interessadas, por intermédio de uma Aliança Europeia para a IA, com vista à elaboração de orientações sobre as questões éticas associadas à IA.<sup>12</sup>

Resulta afirmar que, além de a comunicação da Comissão Europeia ao Parlamento Europeu demandar o desenvolvimento de regras jurídicas adequadas, de modo a reforçar a capacidade industrial e tecnológica da União Europeia e adoção da IA na economia e, assim, preparar as mudanças socioeconômicas, também destaca a necessidade de garantir procedimentos éticos para as futuras aplicações de IA.<sup>13</sup> Cumpre observar que todos os esforços

---

<sup>12</sup> Na Comunicação está explícito que, assim como aconteceu no passado com o motor a vapor ou a eletricidade, a IA está a transformar o nosso mundo, a nossa sociedade e a nossa indústria. O aumento da capacidade computacional, a disponibilidade de dados e a evolução dos algoritmos fizeram da IA uma das tecnologias mais estratégicas do século XXI. A fasquia não podia estar mais alta. O modo como abordarmos a IA definirá o mundo em que vivemos. Num contexto de feroz concorrência global, é necessário um quadro europeu sólido. UNIÃO EUROPEIA. Comissão Europeia. **Regulation of the European Parliament and of the Council. Concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications).** Bruxelas, Bélgica: Comissão Europeia, 10 de janeiro de 2017. Disponível em: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52017PC0010>. Acesso em: 16 ago. 2020. p. 2.

<sup>13</sup> Segundo a Comunicação é necessário gerar um clima de confiança e responsabilidade em torno do desenvolvimento e da utilização da IA. Os **valores** enunciados no artigo 2.º do Tratado da União Europeia constituem a base dos direitos de que gozam as pessoas que vivem na União. Além disso, a **Carta dos Direitos Fundamentais da União Europeia** reúne num único texto todos os direitos individuais, cívicos, políticos, económicos e sociais gozados pelas pessoas no espaço da UE. A UE dispõe de um quadro regulamentar sólido e equilibrado, que pode estabelecer a norma global para uma abordagem sustentável a esta tecnologia. A União dispõe de **padrões elevados em termos de segurança e de responsabilidade decorrente dos produtos**. As primeiras regras à escala da UE em matéria de **segurança das redes e dos sistemas de informação** e o reforço das regras em **matéria de proteção de dados pessoais** tornar-se-ão realidade em maio de 2018. O **Regulamento Geral sobre a Proteção de Dados** garante um elevado nível de proteção dos dados pessoais, incluindo os princípios da proteção de dados desde a concessão e por defeito. Garante também a livre circulação de dados pessoais no território da União. Contém disposições sobre a tomada de decisões unicamente com base no tratamento automatizado, incluindo a definição de perfis. Nestes casos, os titulares dos dados têm o **direito a receber informações úteis** relativas à lógica subjacente à decisão. O Regulamento Geral sobre a Proteção de Dados também confere aos indivíduos o direito de não ficarem sujeitos a decisões tomadas exclusivamente com base no tratamento automatizado, exceto em determinadas situações. A Comissão acompanhará de perto a aplicação do regulamento no contexto da IA e convida as autoridades nacionais de proteção de dados e o Comité Europeu para a Proteção de Dados a que façam o mesmo. Além disso, a Comissão apresentou uma série de propostas, no âmbito da estratégia para o **Mercado Único Digital**, que serão essenciais para permitir o desenvolvimento da IA, como o regulamento sobre a livre circulação de dados não pessoais, e que reforçarão a confiança no mundo em linha, como o Regulamento Privacidade e Comunicações Eletrónicas e o Regulamento Cibersegurança. Estas propostas devem ser adotadas o mais rapidamente possível. Esse apoio é essencial, visto que **tanto os cidadãos como as empresas devem poder confiar na tecnologia com que interagem**, dispor de um ambiente jurídico previsível e confiar em salvaguardas eficazes que protejam os direitos e liberdades fundamentais. Para reforçar a confiança, as pessoas precisam

de diversas entidades da União Europeia são exatamente no sentido de traçar princípios éticos no uso da IA. A exemplo, foi o Parecer “Inteligência artificial: antecipar o seu impacto no trabalho para assegurar uma transição justa” resultante da 537ª Reunião Plenária do Comitê Econômico e Social Europeu em setembro de 2018, que destacou o necessário direcionamento de responsabilidade dos sistemas autônomos em caso de acidente. Ele exigiu claramente o estabelecimento prévio dos riscos para a saúde em que os trabalhadores eventualmente incorram, de modo a garantir-lhes a segurança plena.<sup>14</sup>

Nesse sentido, foi a proposta que trata os princípios éticos relativos ao uso da inteligência artificial em sistemas judiciais da Comissão Europeia para o Eficiência de Justiça (CEPEJ),<sup>15</sup> que lista cinco princípios norteadores, quais

---

igualmente de compreender como a tecnologia funciona, daí a importância da investigação no domínio dos **sistemas de IA explicáveis**. De facto, os sistemas de IA devem ser desenvolvidos de uma forma que permita ao ser humano ter um entendimento (das bases) das suas ações, a fim de aumentar a transparência e minimizar os riscos de distorção ou erro. Como qualquer tecnologia ou ferramenta, a IA pode ser utilizada de uma forma positiva, mas também pode ser usada com maus intuítos. Apesar de a IA gerar claramente novas oportunidades, apresenta também desafios e riscos, por exemplo, nos domínios da segurança. (utilização criminosa ou ciberataques), da segurança e da responsabilidade relativas a produtos, do preconceito e da discriminação. Será necessária uma reflexão sobre as interações entre a IA e os direitos de propriedade intelectual, da perspectiva dos institutos de propriedade intelectual e dos utilizadores, que vise promover a inovação e a segurança jurídica de forma equilibrada. Fonte: Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comitê Econômico e Social Europeu e ao Comitê das Regiões. Resultados da consulta pública sobre a estratégia Europa 2020 para um crescimento inteligente, sustentável e inclusivo. UNIÃO EUROPEIA. Comissão Europeia. **Regulation of the European Parliament and of the Council**. Concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications). Bruxelas, Bélgica: Comissão Europeia, 10 de janeiro de 2017. Disponível em: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52017PC0010>. Acesso em: 16 ago. 2020. p.4.

<sup>14</sup> O Parecer destacou a reflexão sobre os riscos da IA, por meio do Princípio da Responsabilidade: “O princípio da responsabilidade jurídica deve ser clarificado. Os robôs industriais ou de serviços colaboram cada vez mais com o ser humano. A IA permite aos robôs «saírem das suas jaulas», podendo provocar acidentes. Como tal, a responsabilidade dos sistemas autônomos em caso de acidente deve ser claramente estabelecida e os riscos para a saúde e segurança em que os trabalhadores incorrem devem poder ser cobertos. A Comissão Europeia está a lançar uma reflexão sobre estes riscos emergentes no quadro da Diretiva em matéria de responsabilidade decorrente dos produtos. Esta abordagem deve ser mais ambiciosa no que diz respeito à segurança no trabalho”. JORNAL Oficial da União Europeia. **Comunicações e Informações**. Luxemburgo: Serviço das Publicações da União Europeia, 06 de dezembro de 2018. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=OJ:C:2018:440:FULL&from=EN>. Acesso em: 17 ago. 2020.

<sup>15</sup> A Comissão Europeia para o Eficiência de Justiça (CEPEJ) – Carta Ética Europeia sobre o uso de inteligência artificial em Sistemas judiciais e seu ambiente, reúne especialistas dos 47 Estados-membros do Conselho da Europa, com objetivo de melhorar a qualidade e a eficiência dos sistemas judiciais europeus, e também, fortalecer a confiança dos usuários dos tribunais europeus. UNIÃO EUROPEIA. European Commission for the Efficiency of Justice (CEPEJ). Council of Europe. **European ethical charter on the use of Artificial Intelligence in judicial**

sejam: (i) princípio do respeito aos direitos fundamentais: garantir que o *design* e implementação de ferramentas de inteligência artificial e os serviços são compatíveis com os direitos fundamentais; (ii) princípio de não discriminação: prevenir especificamente o desenvolvimento ou intensificação de qualquer discriminação entre indivíduos ou grupos de indivíduos; (iii) princípio da qualidade e segurança: no que diz respeito ao tratamento de decisões judiciais e dados, utilizar fontes certificadas e dados intangíveis com modelos elaborados de forma multidisciplinar, em ambiente tecnológico seguro; (iv) princípio da transparência, imparcialidade e justiça: faça métodos de processamento de dados acessíveis e compreensíveis, autorizar auditorias externas; e (v) princípio "sob controle do usuário": dedica uma abordagem prescritiva, de modo a garantir que os usuários sejam atores informados e controle as escolhas feitas.

A Nationale KI-Strategie apresentou o planejamento estratégico alemão sobre a IA, que envolveu, no governo, a pesquisa, a educação, a economia, o trabalho e assuntos sociais.<sup>16</sup> A estratégia de IA do governo federal alemã foi criada em um processo, na primeira fase foi desenvolvido um documento de questões essenciais – publicado em 18 de julho de 2018 –, com base nas recomendações do fórum de especialistas em Sistemas Autônomos da Estratégia de Alta Tecnologia em 20 de março de 2017, uma audiência de especialistas a convite do Chanceler em 29 de maio de 2018 e um importante trabalho preliminar Ministérios federais para desenvolver a estratégia de IA. Os pilares formaram a base para um processo de participação única na fase 2. A segunda fase foi um processo de consulta bem abrangente. Especialistas de política, negócios, ciência e sociedade civil foram convidados para seis fóruns

---

**systems and their environment.** Estrasburgo: CEPEJ, 03 de dezembro de 2020. Disponível em: <https://rm.coe.int/ethicalcharter-en-for-publication-4-december-2018/16808f699c>. Acesso em: 17 ago. 2020.

<sup>16</sup> Com sua estratégia nacional, o governo Alemão há um ano promove a trajetória da Inteligência Artificial Fabricado na Alemanha para o topo do mundo. Trata-se de direitos individuais de liberdade, autonomia, direitos pessoais, liberdade de decisão do indivíduo. Sobre esperanças, medos, potencialidades e expectativas. Mas também se trata de novos mercados para as empresas alemãs, da concorrência global, especialmente com os EUA e a China, e o futuro da Alemanha como local industrial. O Governo Federal está ciente desta enorme responsabilidade: A estratégia de IA para a Alemanha foi criada em um processo democrático abrangente. Muitos especialistas participaram. Com coragem e vontade de criar e sempre com foco na inteligência artificial no atendimento e em benefício das pessoas. ALEMANHA. Nationale Strategie für Künstliche Intelligenz. **Künstliche Intelligenz (KI) ist ein Schlüssel zur Welt von morgen.** Berlim, Alemanha: Die Bundesregierung, [2020?]. Disponível em: <https://www.ki-strategie-deutschland.de/home.html>. Acesso em: 17 ago. 2020.

de especialistas e um processo de consulta *on-line*. Houve diversos fóruns, como de "Mobilidade, Logística e Sustentabilidade", "Saúde & Cuidado", "Transferência para a Economia", "Mundo do Trabalho e Mercado de Trabalho", "Investigação" e "Produção e Indústria 4.0", contando com 140 especialistas falando sobre temas voltados à inteligência artificial. Eles foram discutidos de forma abrangente e houve recomendações para a estratégia de IA desenvolvidas em conjunto. A chamada para o processo de consulta *on-line* foi seguida por 109 especialistas, que enviaram seus comentários por escrito sobre os doze campos de ação formulados no documento de questões-chave.<sup>17</sup>

O Governo Alemão reconhece os avanços globais conquistados pela IA e, com isso, reconheceu como imperioso o investimento em pesquisa e inovação para a construção de uma base sólida para desenvolvimento das tecnologias advindas da IA. O Governo Alemão se mostra agressivo no seu projeto de IA, e os objetivos são de se tornarem os líderes de tecnologias em IA. São as metas alemãs: (i) a Alemanha deve expandir sua forte posição na Indústria 4.0 e se tornar líder em aplicações de IA nesta área. Nossas fortes empresas de médio porte também devem se beneficiar dos aplicativos de IA; (ii) deve-se tornar um local atraente para as mentes de IA mais brilhantes do mundo; (iii) o foco deve estar sempre nos benefícios da IA para os cidadãos.; (iv) na Alemanha, os dados só devem ser usados em benefício da sociedade, do meio ambiente, da economia e do estado; (v) com uma nova infraestrutura para transmissão de dados em tempo real, a base para novos aplicativos de IA deve ser criada; (vi) a IA, na Alemanha, deve ser acompanhada por um alto nível de segurança de TI; (vii) queremos encontrar uma resposta europeia para modelos de negócio baseados em dados e novas formas de criação de valor baseada em dados que corresponda à nossa estrutura económica, de valor e social; (viii) também no mundo do trabalho, deve-se sempre tratar da IA para o benefício de todos os funcionários; (ix) o potencial da IA deve tornar as áreas de moradia e trabalho mais seguras, eficientes e sustentáveis; (x) a IA visa a promover a participação social, a liberdade de ação e a autodeterminação dos cidadãos; (xi) o potencial da IA deve ser usado para o desenvolvimento sustentável e, assim, contribuir para o cumprimento das metas de sustentabilidade da Agenda 2030; (xii) são definidas as condições estruturais para aplicações de IA que criam e mantêm a diversidade e garantem o espaço disponível para o desenvolvimento das liberdades culturais e de mídia.<sup>18</sup>

Nota-se que a estratégia alemã no uso da IA se apoia no próprio anseio democrático voltado para o acesso e desenvolvimento tecnológico do povo alemão. Ela é baseada num

---

<sup>17</sup> ALEMANHA. Nationale Strategie für Künstliche Intelligenz. **Künstliche Intelligenz (KI) ist ein Schlüssel zur Welt von morgen**. Berlim, Alemanha: Die Bundesregierung, [2020?]. Disponível em:

<https://www.ki-strategie-deutschland.de/home.html>. Acesso em: 17 ago. 2020.

<sup>18</sup> Ibid.

contexto como a ética, jurídico, cultural e institucional, com vista a dar acesso a todos os cidadãos, refletindo em áreas sensíveis da vida, com observância aos valores sociais fundamentais, direitos individuais, estruturados numa IA que seja acessível aos cidadãos alemães. E disso, depreende-se que a preocupação do planejamento estratégico do Nationale KI-Strategie é no sentido de antecipar os possíveis riscos oriundos do uso das tecnologias de IA. Contudo, dado o andar do conhecimento das tecnologias de IA, torna-se prematuro antever eventuais riscos que a sociedade venha a se confrontar.

Cumprido destacar que, no planejamento estratégico do Nationale KI-Strategie, apesar de a distinção entre IA fraca e forte ser o meio pelo qual o documento alemão traça o modo de avanço para as tecnologias de IA,<sup>19</sup> essa distinção se mostra inadequada,<sup>20</sup> uma vez que a IA fraca é destinada às soluções de problemas específicos. Ela é, assim, denominada de maneira mais adequada de IA específica ou IA estreita (*narrow*). Essa IA específica, inclusive, em determinadas utilizações, assemelha-se muito do conceito de IA forte utilizada no planejamento estratégico do Nationale KI-Strategie, posto que desempenha funções com maior eficácia em comparação com os humanos.<sup>21</sup> Por isso, a

---

<sup>19</sup> A construção da estratégia alemã exsurtiu da diferenciação de dois tipos de IA, forte e fraca. No planejamento estratégico do Nationale KI-Strategie a IA fraca é destinada a solução de problemas específicos, já a IA forte significa sistemas de IA com as idênticas ou superiores capacidades intelectuais que os seres humanos. ALEMANHA. Nationale Strategie für Künstliche Intelligenz. **Künstliche Intelligenz (KI) ist ein Schlüssel zur Welt von morgen**. Berlim, Alemanha: Die Bundesregierung, [2020?]. Disponível em: <https://www.ki-strategie-deutschland.de/home.html>. Acesso em: 17 ago. 2020.

<sup>20</sup> O professor Ryan Calo da University of Washington School of Law destaca que os estudiosos realizam a distinção daquilo que denominam de IA forte, definida com uma espécie de inteligência artificial geral de capacidade assemelhadas dos humanos em todas as áreas, ao passo que, a IA fraca ou superficial, seria aquelas destinada a tratar dos problemas individualizados. CALO, Ryan. Artificial Intelligence Policy: A Primer and Roadmap. **UCDL Review**, v. 51, p. 399-435, ago. 2017. Disponível em: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3015350](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3015350). Acesso em: 17 ago. 2020. p. 404.

<sup>21</sup> Há caso na medicina que a ia específica demonstra um desempenho superior ao humano. A inteligência artificial analisa dados médicos complexos. Seu potencial para explorar relações significativas em um conjunto de dados é usado no diagnóstico, tratamento e previsão de resultados em muitos cenários clínicos. Nesse sentido, projetos de ia específica, têm potencial para serem aplicadas em quase todos os campos da medicina. O primeiro desses algoritmos é um dos vários exemplos existentes de um algoritmo que supera os médicos nas tarefas de classificação de imagens. No outono de 2018, pesquisadores do hospital da universidade nacional de Seul e da faculdade de medicina desenvolveram um algoritmo de ia chamado dlad (deep learning-based automatic detection – detecção automática baseada em aprendizado profundo) para analisar radiografias de tórax e detectar crescimento celular anormal, como possíveis cânceres (figura abaixo). O desempenho do algoritmo foi comparado às habilidades de detecção de vários médicos nas mesmas imagens e superou 17 dos 18 médicos testados. Fonte: development and validation of deep learning-based automatic detection algorithm for malignant pulmonary nodules on chest radiographs. JU, Gang Nam *et al.*

distinção entre a IA fraca e forte é definida por este autor como inadequada, porque, além de a IA fraca ser definida mais apropriadamente como IA específica, entende o autor que a especificidade no uso das tecnologias de IA é o caminho para construção ética das modelagens de IA. Inclusive, no documento alemão, o uso das tecnologias de IA está direcionado às funções específicas, quais sejam: sistemas de dedução, sistemas baseados no conhecimento, reconhecimento e análise de padrões, sistemas robóticos autônomos, interações homem-máquina multimodais inteligentes (análise e entendimento da linguagem, imagens, gestos e outras formas de interação humana). Isso resulta afirmar que a nomenclatura mais apropriada seja IA específica, já que as funções predestinadas das tecnologias de IA não são para substituir o raciocínio cognitivo do ser humano e, sim, reduzir o tempo de tarefas procedimentais e repetitivas, precipuamente, tratando-se de matérias de direito constitucional.

Após delinear os procedimentos e funcionalidades da IA acima referenciados, a Alemanha também apresentou a política interna para consubstanciar a inteligência artificial “*by Germany*”. Então, foram traçados tanto os objetivos específicos<sup>22</sup> quanto os objetivos táticos de consolidação da primazia alemã nas tecnologias de IA, assentado num investimento até 2025 da ordem de 3 bilhões de euro<sup>23</sup>.

---

Development and validation of deep learning-based automatic detection algorithm for malignant pulmonary nodules on chest radiographs. **Radiology**, v. 290, n. 1, p. 218-228, jan. 2019. Disponível em: <https://pubs.rsna.org/doi/full/10.1148/radiol.2018180237>. Acesso em: 18 ago. 2020.

<sup>22</sup> O impulso de inovação que vem com a tecnologia deve ser usado extensivamente para o benefício de todos. O excelente centro de pesquisa na Alemanha deve ser assegurado, a competitividade da economia alemã expandida e as diversas possibilidades de aplicação da IA promovidas em todas as áreas da sociedade no sentido de um progresso social perceptível e no interesse dos cidadãos. O foco é o benefício às pessoas e ao meio ambiente. Logo os objetivos específicos são: Orientar o governo a fornecer respostas políticas rápidas aos avanços no campo da IA, assegurar o uso das inovações desencadeadas pela IA para o benefício da sociedade em geral, aumentar a competitividade da indústria alemã, promover as muitas maneiras de usar a IA em todas as partes da sociedade, alcançar um progresso tangível na sociedade, no interesse dos cidadãos, concentrar no uso benéfico às pessoas e ao meio ambiente, e promover um intenso diálogo com todas as seções da sociedade. ALEMANHA. Nationale Strategie für Künstliche Intelligenz. **Künstliche Intelligenz (KI) ist ein Schlüssel zur Welt von morgen**. Berlim, Alemanha: Die Bundesregierung, [2020?]. Disponível em: <https://www.ki-strategie-deutschland.de/home.html>. Acesso em: 17 ago. 2020.

<sup>23</sup> Até 2025 inclusive, o governo federal pretende disponibilizar um total de cerca de 3 bilhões de euros para a implementação da estratégia. O efeito de alavanca desse compromisso nos negócios, na ciência e nos países levará a pelo menos uma duplicação desses fundos. Isso tudo para almejar os 3 objetivos táticos alemão. O primeiro objetivo tático: desenvolver centros de excelência para IA em nível supra regional, formando uma rede de centros com pelo menos doze centros e hubs de aplicação, oferecer condições de trabalho e remuneração internacionalmente atraente e competitiva para pesquisadores e reforçar o ensino acadêmico, criar ao menos 100

No documento alemão, está explícito o caráter de respeito ao direito fundamental do cidadão alemão quando da manipulação dos dados, um ponto fundamental para o progresso da IA *Germany*, uma vez que os algoritmos, para um aprendizado de máquina eficiente, muitas vezes, acessam dados sensíveis dos indivíduos. O programa alemão se mostra interdisciplinar,<sup>24</sup> o que permite que a IA seja acessível a todos os cidadãos alemães, algo muito diferente quando equiparado com o Brasil, onde as tecnologias de IA são acessíveis às classes médias altas, salvo quando a classe baixa utiliza o aplicativos para sair do trânsito engarrafado.

---

pontos de trabalho para professores para IA, assegurando a presença, o mais forte possível, da IA no ensino superior, trabalhar com a França para promover o desenvolvimento de uma rede franco-alemã de pesquisa (“centro virtual”), baseado em estruturas já existentes e nas habilidades específicas, tornar a IA prioritária na agência, que é prevista para atuar em inovações e descobertas, formar um cluster europeu de inovação, fornecendo financiamento para projetos de pesquisa cooperativos, aumentar suporte específico de IA para pequenas e médias empresas, com facilitadores/capacitadores de IA com pelo menos 1.000 empresas por ano (utilizando o *Mittestand 4.0 Centres of Excellence*), ampliar iniciativas orçamentárias e criar oportunidades para capitais de investimento, capitais de risco e fundos, inclusive com a iniciativa do Fundo de Crescimento Tecnológico, expandir serviços de consultoria e financiamento para *startups*, estimular estruturas de compartilhamento voluntário de dados, em conformidade com as normas de proteção de dados, no caminho de formar uma infraestrutura confiável de dados e análise, incluindo a construção de uma plataforma em nuvem com capacidade de armazenamento e processamento atualizável, na qual uma infraestrutura possa ser executada. O segundo objetivo tático guarda relação com a IA voltada a sociedade, estabelecendo um observatório nacional de IA, para atender ao ser humano, e fornecer habilidades ao povo alemão para que possuam habilidades digitais e das tecnologias da IA. Além de financiar espaços de inovação nas empresas para aplicativos de IA no mundo do trabalho, e financiar no mínimo 50 (cinquenta) sistemas de IA para beneficiar o meio ambiente e o clima, bem como desenvolver princípios de avaliação. Por fim, o terceiro objetivo tático se dedica as seguintes ações: focar o desenvolvimento e usos de IA nas pessoas, especialmente quando houver impactos no mundo do trabalho, centrar o desenvolvimento de habilidades, talentos e criatividade, permitindo a autodeterminação, fornecendo segurança e protegendo a saúde, promover a representação adequada da diversidade, considerar as necessidades especiais das pessoas com deficiência para incrementar sua participação na vida profissional, usar os potenciais de IA para melhorar a segurança, a eficiência e a sustentabilidade em campos de aplicação particularmente importantes, promover a participação social e cultural, a liberdade de ação e autodeterminação dos cidadãos, utilizar o potencial da IA para o desenvolvimento sustentável e auxílio no alcance das metas de desenvolvimento da Agenda 2030, e criar um ambiente de políticas para aplicativos de IA que crie e mantenha a diversidade e garanta o escopo necessário para o desenvolvimento de liberdades culturais e de mídia, pelo fato de que em uma era digital, o nível de liberdade de uma sociedade democrática é medido em termos da diversidade cultural e de mídia e, particularmente, da independência da mídia. ALEMANHA. Nationale Strategie für Künstliche Intelligenz. **Künstliche Intelligenz (KI) ist ein Schlüssel zur Welt von morgen**. Berlim, Alemanha: Die Bundesregierung, [2020?]. Disponível em: <https://www.ki-strategie-deutschland.de/home.html>. Acesso em: 17 ago. 2020.

<sup>24</sup> No planejamento estratégico do Nationale KI-Strategie para proliferação das tecnologias de IA foi projetada uma estrutura de clusters e hubs de transferência, que por meio de um conselho consultivo composto de membros internacionais (Ibid.).

Aqui, está sendo descrito o planejamento estratégico do Nationale KI-Strategie, pois entende-se como sofisticado e digno de uma nação de primeiro mundo a preocupação pelo desenvolvimento e olhar ético é notável. O documento já projeta as mudanças nos postos de trabalho que serão afetados pelo avanço da IA específica, esta que, no curto prazo, eliminará as profissões de trabalhadores com tarefas repetitivas. Por isso, com preocupação ética, tanto no ponto de vista do sistema quanto do ponto de vista dos trabalhadores afetados pela tecnologia, o projeto alemão resgata esse trabalhador que poderá ser excluído do mercado de trabalho e oferece o treinamento necessário para retornar ao mercado de trabalho com o conhecimento das tecnologias da IA.

Além disso, o aspecto ético se sobressai quando o planejamento estratégico de IA do governo alemão destaca a necessidade de *accountability* das decisões tomadas pela IA, ou seja, pelo aprendizado de máquina construído sob a égide dos algoritmos. Isso é direito fundamental do cidadão e, por isso, digno de destaque. Muito embora, invariavelmente, sabe-se que é intangível ao usuário de IA o racional explicativo de como a máquina chegou à determinada decisão, pois o desafio da IA é exatamente prestar ao usuário a *accountability* das decisões tomadas pela IA.

No campo jurídico, a referida *accountability* deve ser interpretada e compreendida no tempo, ou seja, primeiro a partir da fundamentação do próprio julgador (não-algoritmo). Isso, ao fim e ao cabo, trata-se da própria preservação da isonomia e de cuidados na manipulação de dados sigilosos ou sensíveis das partes. Notoriamente, na grande parte dos sistemas de IA, a transparência algorítmica é limitada, e isso se dá por dois motivos: primeiro, há limitações técnicas e humanas para descrever o racional dos milhares de algoritmos que constroem o aprendizado de máquinas no mundo e, no campo do direito, a explicabilidade e a transparência algorítmica podem colocar em xeque o próprio direito fundamental de inviolabilidade das partes e paridade de armas, algo que deve ser observado no campo do Direito. As discussões futuras para um consenso, conforme alcançado por meio da difusão segura bizantina na rede de *blockchain* do bitcoin<sup>25</sup>, são indiscutivelmente necessárias.

---

<sup>25</sup> Muitos algoritmos baseados em blockchain, como Bitcoin, implementam um sistema de transferência de ativos descentralizado, frequentemente referido como criptomoeda. Conforme afirmado no artigo original de Nakamoto, no cerne desses sistemas está o problema de evitar

O programa alemão destaca ainda que guarda observância a transparência e explicabilidade das tecnologias de IA a serem implantadas no futuro, pois devem proteger os interesses e privacidade dos usuários. Isso ocorrerá por meio do controle interno e externo e a curadoria de dados, com objetivo central de reforçar o caráter ético do avanço tecnológico em IA. O *German AI observatory* é o exemplo pelo qual a Alemanha busca garantir sua primazia tecnológica. O observatório terá a função de consubstanciar o treinamento daquele cidadão que terá seu posto de trabalho repetitivo superado pelo *software* de IA, para que ele tenha a oportunidade de reinserção no mercado de trabalho, até superar o desafio de garantir o binômio da transparência e da privacidade no uso da IA.

O Canadá foi outro país que preocupado com o avanço das tecnologias; realizou a *Montreal Declaration for the Responsible Development of Artificial Intelligence Launched*,<sup>26</sup> um conjunto de diretrizes éticas para o desenvolvimento da inteligência artificial, resultado de mais de um ano de trabalho, pesquisas e consultas com cidadãos, especialistas, formuladores de políticas públicas e partes interessadas da indústria, bem como de organizações da sociedade civil

---

gastos duplos; isso geralmente é resolvido por meio do consenso sobre a ordem das transferências entre os participantes. Neste artigo, tratamos o problema de transferência de ativos como um objeto concorrente e determinamos seu número de consenso, mostrando que o consenso, de fato, não é necessário para evitar gastos duplos. Primeiro consideramos o problema conforme definido por Nakamoto, onde apenas um único processo --- o proprietário da conta --- pode retirar de cada conta. Segurança e vivacidade precisam ser garantidas para os proprietários corretos da conta, ao passo que proprietários de contas com comportamento inadequado podem não ser capazes de realizar transferências (1). Em seguida, consideramos um mais geral k-objeto de transferência de ativos compartilhados onde até (k) processos podem retirar atômica da mesma conta, e mostrar que este objeto tem número de consenso (k). Estabelecemos nossos resultados no contexto de memória compartilhada com falhas benignas, o que nos permite entender adequadamente o nível de dificuldade do problema de transferência de ativos. Também traduzimos esses resultados no cenário de troca de mensagens com jogadores bizantinos, um modelo que é mais relevante na prática. Neste modelo, descrevemos uma implementação de transferência de ativos tolerante a falhas bizantina assíncrona que é mais simples e mais eficiente do que soluções baseadas em consenso de última geração. Nossos resultados são aplicáveis às configurações com permissão (privado) e sem permissão (público), já que normalmente sua diferenciação é oculta pelas abstrações nas quais nossos algoritmos são baseados. Com o requisito de consenso fora do caminho, o sistema de Difusão Segura Bizantina pode garantir transações seguras de criptomoedas em larga escala com um custo energético de praticamente zero - aproximadamente equivalente ao da troca de e-mails, e gerando apenas alguns gramas de CO<sub>2</sub>, em comparação com os estimados 300 kg para cada transação de Bitcoin. GUERRAOUI, Rachid *et al.* The Consensus Number of a Cryptocurrency. *ArXiv*, p. 1-11, jun. 2019. Disponível em: <https://arxiv.org/pdf/1906.05574.pdf>. Acesso em: 18 ago. 2020.

<sup>26</sup> THE MONTREAL Declaration for the Responsible Development of Artificial Intelligence Launched. In: CAB Canada-ASEAN Business Council. Montreal: Universidade de Montreal, 12 de abril de 2018. Disponível em: <https://www.canasean.com/the-montreal-declaration-for-the-responsible-development-of-artificial-intelligence-launched/>. Acesso em: 19 ago. 2020.

e de ordens profissionais com olhar nas tecnologias de IA.<sup>27</sup> A Declaração destaca que, para lê-los corretamente, é importante manter os seguintes pontos em mente: embora sejam princípios, eles devem ser interpretados de forma consistente para evitar qualquer conflito que poderia impedir sua aplicação. Os princípios como regra geral, seus limites são definidos por outro princípio no campo da aplicação; embora reflitam a moral e a política cultural da sociedade em que foram desenvolvidos, eles fornecem a base para um diálogo intercultural e internacional; embora possam ser interpretados em diferentes formas, eles não podem ser interpretados em qualquer maneira; é imperativo que a interpretação seja coerente<sup>28</sup> e, embora sejam princípios éticos, eles podem ser traduzidos em linguagem política e interpretados na forma legal.

---

<sup>27</sup> Vejamos a manifestação de alguns atores envolvidos no *Montreal Declaration for the Responsible Development of Artificial Intelligence Launched*. Disse Remi Quirion, cientista-chefe de Quebec: “Além de apoiar o desenvolvimento de conhecimentos científicos em inteligência artificial, tornou-se importante para o Fonds de recherche du Québec levar em consideração os aspectos éticos e sociais do progresso neste campo de pesquisa. Com isso em mente, saúdo o trabalho que levou ao lançamento da Declaração de Montreal para o Desenvolvimento Responsável da Inteligência Artificial, desenvolvida em colaboração com a sociedade civil. O novo Observatório Internacional sobre os Impactos Societais da Inteligência Artificial e Tecnologias Digitais, anunciado ontem, certamente será baseado nos princípios contidos na Declaração.” O reitor da Université de Montréal, Guy Breton assinalou: “A Université de Montréal tem estado no centro de muitas descobertas científicas em inteligência artificial, e estou satisfeito que uma vasta rede de pesquisadores e líderes das esferas científica, jurídica, econômica e política agora se uniu em torno dos desafios sociais e éticos relacionados a AI”. Yoshua Bengio, membro do Comitê Diretivo da Declaração de AI Responsável de Montreal, acrescentou: “Como cientista, sinto uma grande responsabilidade em garantir que os resultados de minha pesquisa e dos desenvolvimentos em todo o espectro da AI sejam usados em benefício de como tantas pessoas quanto possível. Ao integrar os princípios da responsabilidade social no início do desenvolvimento tecnológico em laboratórios universitários e industriais, podemos ter um impacto verdadeiramente positivo e concreto. Essa é a característica distintiva desta Declaração: ela incorpora as preocupações de todas as partes interessadas no campo e reúne toda a sociedade civil em torno de princípios sólidos”. THE MONTREAL Declaration for the Responsible Development of Artificial Intelligence Launched. In: CABC Canada-ASEAN Business Council. Montreal: Universidade de Montreal, 12 de abril de 2018. Disponível em: <https://www.canasean.com/the-montreal-declaration-for-the-responsible-development-of-artificial-intelligence-launched/>. Acesso em: 19 ago. 2020.

<sup>28</sup> Nesse sentido, deve ser destacado a importância da coerência da Teoria Dworkiniana. Em suma, a teoria dworkiniana mostra-se inexoravelmente coerente com a atribuição de sentidos derivada dos valores, que como se pôde notar, sempre andam em duplicidade de modo que se tenha sempre uma paridade de valores colocados em discussão, exatamente onde repousa a coerência no uso da moral. O direito para Dworkin é visto como um ramo da moral, apesar dos atributos que o particularizam. Já em termos de filosofia política, os direitos fundamentais são vistos como espécies de trunfos contramajoritários, e vinculam-se à ideia de dignidade de que Dworkin lança mão, da qual parte o direito a ser tratado com igual consideração e respeito, e divide a dignidade em dois princípios: (i) o direito a ser tratado com igual consideração e respeito e (ii) o de que os indivíduos têm a responsabilidade pessoal de definir o sucesso em suas próprias vidas, Trata-se de uma concepção nitidamente liberal quanto à concepção dos valores morais e, por conseguinte, do direito, na medida em que Dworkin o vincula a essa ideia de dignidade. DWORKIN, Ronald. **A raposa e o porco-espinho**: justiça e valor. Tradução de Marcelo Brandão Cipolla. São Paulo: Martins Fontes, 2014. p. 513.

Esse conjunto de diretrizes éticas traçou três propósitos basilares, para que os princípios sejam norteados: (1) desenvolver uma estrutura ética para o desenvolvimento e implantação da IA; (2) orientar a transição digital para que então todos se beneficiem da revolução tecnológica; (3) abrir um fórum nacional e internacional para discussão, de modo a alcançar coletivamente a equidade, com inclusão ecologicamente sustentável no desenvolvimento da IA.

Cumprido destacar que a Declaração deixa claro que os princípios apresentados a seguir em lista não devem ser lidos em hierarquia. O último princípio não é menos importante do que o primeiro. No entanto, é possível, dependendo nas circunstâncias, dar mais peso a um princípio do que a outro, ou até mesmo, considerar um princípio mais relevante do que outro.<sup>29</sup> O Princípio do Bem-Estar destaca que o desenvolvimento e o uso de sistemas de inteligência artificial devem permitir o crescimento do bem-estar de todos os seres humanos;<sup>30</sup> o Princípio da Autonomia traz que deve ser desenvolvido e

---

<sup>29</sup> Aqui mais uma vez, embora os princípios em questão estejam sendo discutidos no âmbito da interdisciplinaridade, pois trata do avanço da IA em todas as áreas do conhecimento canadense, e não a partir de uma objetividade principiológica entre Direito e IA, no âmbito da Teoria do Direito. E Dworkin, ao tratar dos conceitos interpretativos, destaca que as práticas elaboradas em determinadas comunidades, frequentemente, estão baseadas em regras carregadas de juízos de valor, e funcionariam como estímulos à reflexão sobre aquilo que é exigido por essas práticas. O objetivo de Dworkin estaria em alcançar a melhor interpretação sobre o que essas práticas pressupõem ou exigem, pois entende que os conceitos, tais como justiça, liberdade, igualdade, democracia, e o próprio Direito, eminentemente carregados de valores poderiam ser classificados como interpretativos. DWORKIN, Ronald. **A justiça de toga**. Tradução de Jefferson Luis Camargo. São Paulo: Martins Fontes, 2010. p. 17-20.

<sup>30</sup> O princípio da Proteção da Privacidade e Intimidade busca disponibilizar espaços pessoais em que as pessoas não estejam sujeitas a vigilância ou avaliação digital, esta que deve ser protegida do avanço da IA e aquisição de dados para a construção dos sistemas. A intimidade de pensamentos e emoções deve ser estritamente protegido dos usos da IA e aquisições e construção de dados para os sistemas de IA capazes de causar danos, especialmente em usar julgamentos morais às pessoas ou suas escolhas de estilo de vida. As pessoas devem sempre ter o direito à desconexão digital em suas vidas privadas, e o sistemas de IA devem oferecer explicitamente a opção desligar em intervalos regulares, sem encorajar as pessoas para ficar conectado. As pessoas devem ter amplo controle sobre as informações relacionadas suas preferências. O sistema de IA não deve criar preferência individual perfis para influenciar o comportamento dos indivíduos sem seu livre consentimento. O sistema de aquisição e arquivamento de dados deve garantir a confidencialidade dos dados pessoais e o anonimato do perfil. Cada pessoa deve ser capaz de exercer amplo controle sobre seus dados pessoais, especialmente quando se trata de sua coleta, uso e disseminação. O acesso ao sistema de inteligência artificial e serviços digitais por indivíduos não devem ser condicionados ao abandono de propriedade dos seus dados pessoais. Os indivíduos devem ser livres para doar seus dados pessoais para organizações de pesquisa, a fim de contribuir para o avanço do conhecimento. A integridade de sua identidade pessoal deve ser garantida. O sistema de inteligência artificial não deve ser usado para imitar ou alterar a aparência de uma pessoa, voz, ou outras características individuais, a fim de prejudicar sua reputação ou manipular outras pessoas. THE MONTREAL Declaration for the Responsible Development of Artificial Intelligence Launched. In: CAB Canada-ASEAN Business Council. Montreal: Universidade de Montreal, 12 de abril de 2018.

usado com respeito pela autonomia das pessoas e com o objetivo de aumentar o controle das pessoas sobre suas vidas e seus arredores;<sup>31</sup> o Princípio da Privacidade e a Intimidade diz que devem ser protegidas contra invasões por sistemas de IA e por sistemas de aquisição e arquivamento de dados;<sup>32</sup> o Princípio da Solidariedade é o desenvolvimento do sistema de IA que deve ser compatível com a manutenção dos laços de solidariedade entre pessoas e gerações;<sup>33</sup> o Princípio da Participação Democrática descreve que o sistema da

---

Disponível em: <https://www.canasean.com/the-montreal-declaration-for-the-responsible-development-of-artificial-intelligence-launched/>. Acesso em: 22 ago. 2020.

<sup>31</sup> O Princípio da Autonomia permitir que os indivíduos cumpram sua própria moral objetiva e sua concepção de uma vida digna de ser vivida, que não deve ser desenvolvida ou usado para impor um determinado estilo de vida dos indivíduos, seja direta ou indiretamente, implementando vigilância e avaliação opressiva ou mecanismos de incentivo. As instituições públicas não devem usar a IA para promover ou desacreditar uma concepção particular da vida boa. É fundamental capacitar os cidadãos em relação ao digital tecnologias, garantindo o acesso aos formulários relevantes de conhecimento, promovendo a aprendizagem de competências fundamentais (alfabetização digital e midiática) e fomento ao desenvolvimento de pensamento crítico. A IA não deve ser desenvolvido para espalhar informações não confiáveis informações, mentiras ou propaganda e devem ser planejadas a fim de conter a sua divulgação. E, ainda, o desenvolvimento da IA deve evitar a criação de dependências por meio de técnicas de captura de atenção ou a imitação de características humanas (aparência, voz e outras formas) de maneiras que pode causar confusão entre a IA e humanos (Ibid.).

<sup>32</sup> O princípio da Proteção da Privacidade e Intimidade busca disponibilizar espaços pessoais em que as pessoas não estejam sujeitas a vigilância ou avaliação digital, esta que deve ser protegida do avanço da IA e aquisição de dados para a construção dos sistemas. A intimidade de pensamentos e emoções deve ser estritamente protegido dos usos da IA e aquisições e construção de dados para os sistemas de IA capazes de causar danos, especialmente em usar julgamentos morais às pessoas ou suas escolhas de estilo de vida. As pessoas devem sempre ter o direito à desconexão digital em suas vidas privadas, e o sistemas de IA devem oferecer explicitamente a opção desligar em intervalos regulares, sem encorajar as pessoas para ficar conectado. As pessoas devem ter amplo controle sobre as informações relacionadas suas preferências. O sistema de IA não deve criar preferência individual perfis para influenciar o comportamento dos indivíduos sem seu livre consentimento. O sistema de aquisição e arquivamento de dados deve garantir a confidencialidade dos dados pessoais e o anonimato do perfil. Cada pessoa deve ser capaz de exercer amplo controle sobre seus dados pessoais, especialmente quando se trata de sua coleta, uso e disseminação. O acesso ao sistema de inteligência artificial e serviços digitais por indivíduos não devem ser condicionados ao abandono de propriedade dos seus dados pessoais. Os indivíduos devem ser livres para doar seus dados pessoais para organizações de pesquisa, a fim de contribuir para o avanço do conhecimento. A integridade de sua identidade pessoal deve ser garantida. O sistema de inteligência artificial não deve ser usado para imitar ou alterar a aparência de uma pessoa, voz, ou outras características individuais, a fim de prejudicar sua reputação ou manipular outras pessoas. THE MONTREAL Declaration for the Responsible Development of Artificial Intelligence Launched. In: CAB Canada-ASEAN Business Council. Montreal: Universidade de Montreal, 12 de abril de 2018. Disponível em:

<https://www.canasean.com/the-montreal-declaration-for-the-responsible-development-of-artificial-intelligence-launched/>. Acesso em: 22 ago. 2020.

<sup>33</sup> O Princípio da Solidariedade impõe que o sistema de inteligência artificial não deve ameaçar a preservação das relações humanas morais e emocionais satisfatórias e deve ser desenvolvido com o objetivo de promover esses relacionamentos e reduzir a vulnerabilidade e o isolamento das pessoas. O sistema de inteligência artificial deve ser desenvolvido com o objetivo de colaborar com humanos em tarefas complexas e deve promover o trabalho colaborativo entre humanos. O sistema de inteligência artificial não deve ser implementado para substituir pessoas em funções que requeiram relações humanas de qualidade, mas deve ser desenvolvido para

IA deve atender aos critérios de inteligibilidade, justificabilidade e acessibilidade, e deve ser submetido ao escrutínio, debate e controle democráticos;<sup>34</sup> o Princípio da Equidade assegura o desenvolvimento e uso do sistema de IA que deve contribuir para a criação de uma sociedade justa e equitativa;<sup>35</sup> o Princípio da Inclusão da Diversidade descreve que o desenvolvimento e o uso de sistemas de IA devem ser compatíveis com a manutenção da diversidade social e cultural

---

facilitar essas relações. Os sistemas de saúde que usam a IA devem levar em consideração a importância do relacionamento do paciente com a família e a equipe de saúde. O desenvolvimento da IA não deve encorajar o comportamento cruel em relação a robôs projetados para se assemelhar a seres humanos ou animais não humanos em aparência ou comportamento. A IA deve ajudar a melhorar a gestão de riscos e promover condições para uma sociedade com uma distribuição mais justa e mútua dos riscos individuais e coletivos (Ibid.).

<sup>34</sup> O Princípio da Participação Democrática narra que os processos de IA que tomam decisões que afetam a vida, qualidade de vida ou reputação de uma pessoa devem ser inteligíveis para seus criadores. As decisões tomadas pelo sistema de IA que afetam a vida, a qualidade de vida ou a reputação de uma pessoa devem ser sempre justificáveis em uma linguagem que seja entendida pelas pessoas que os usam ou que estão sujeitas às consequências de seu uso. A justificação consiste em tornar transparentes os fatores e parâmetros mais importantes que moldam a decisão, e deve assumir a mesma forma que a justificação que exigiríamos de um ser humano tomando o mesmo tipo de decisão. O código para algoritmos, sejam públicos ou privados, deve estar sempre acessível às autoridades públicas relevantes e partes interessadas para fins de verificação e controle. A descoberta de erros operacionais do sistema de IA, efeitos inesperados ou indesejáveis, violações de segurança e vazamentos de dados devem ser obrigatoriamente relatados às autoridades públicas relevantes, partes interessadas e aqueles afetados pela situação. De acordo com o requisito de transparência das decisões públicas, o código dos algoritmos de tomada de decisão utilizados pelas autoridades públicas deve ser acessível a todos, com exceção dos algoritmos que apresentam alto risco de perigo grave se utilizados indevidamente. Para os sistemas de IA públicos com impacto significativo na vida dos cidadãos, os cidadãos devem ter oportunidade e competências para deliberar sobre os parâmetros sociais desses sistemas, os seus objetivos e os limites da sua utilização. Devemos ser sempre capazes de verificar se os AIS estão fazendo o que foram programados e para que são usados. Qualquer pessoa que utilize um serviço deve saber se uma decisão a seu respeito ou que os afeta foi tomada por um sistema de IA. Qualquer usuário de um serviço com *chatbots* deve ser capaz de identificar facilmente se está interagindo com um sistema de IA ou com uma pessoa real. A pesquisa de inteligência artificial deve permanecer aberta e acessível a todos (Ibid.).

<sup>35</sup> O Princípio da Equidade foi projetado e treinado de modo a não criar, reforçar ou reproduzir a discriminação com base - entre outras coisas - em diferenças sociais, sexuais, étnicas, culturais ou religiosas nos sistemas da IA. O desenvolvimento da IA deve ajudar a eliminar as relações de dominação entre grupos e pessoas com base nas diferenças de poder, riqueza ou conhecimento. O desenvolvimento da IA deve produzir benefícios sociais e econômicos para todos, reduzindo as desigualdades e vulnerabilidades sociais. O desenvolvimento da IA industrial deve ser compatível com as condições de trabalho aceitáveis em todas as etapas do seu ciclo de vida, desde a extração de recursos naturais até a reciclagem e incluindo o processamento de dados. A atividade digital dos utilizadores da IA e dos serviços digitais deve ser reconhecida como mão de obra que contribui para o funcionamento dos algoritmos e cria valor. O acesso a recursos, conhecimentos e ferramentas digitais fundamentais deve ser garantido a todos. Devemos apoiar o desenvolvimento de algoritmos comuns - e de dados abertos necessários para treiná-los - e expandir seu uso, como um objetivo socialmente justo. THE MONTREAL Declaration for the Responsible Development of Artificial Intelligence Launched. In: CABIC Canada-ASEAN Business Council. Montreal: Universidade de Montreal, 12 de abril de 2018. Disponível em: <https://www.canasean.com/the-montreal-declaration-for-the-responsible-development-of-artificial-intelligence-launched/>. Acesso em: 22 ago. 2020.

e não devem restringir o escopo das escolhas de estilo de vida e experiência pessoal;<sup>36</sup> o Princípio da Prudência dispõe que cada pessoa envolvida no desenvolvimento do AIS deve ter cuidado antecipando tanto quanto possível, as consequências adversas potenciais do uso do AIS e tomar as medidas apropriadas para evitá-las;<sup>37</sup> o Princípio da Responsabilidade traz que o desenvolvimento e o uso do sistema de IA não devem contribuir para diminuir a responsabilidade do ser humano na tomada de decisões;<sup>38</sup> e, por fim, o Princípio

---

<sup>36</sup> O Princípio da Inclusão da Diversidade prega que o desenvolvimento e uso da IA não devem levar à homogeneização da sociedade por meio da padronização de comportamentos e opiniões. A partir do momento em que os algoritmos são concebidos, o desenvolvimento e implantação do sistema de IA deve levar em consideração a multiplicidade de expressões de diversidade social e cultural presentes na sociedade. Ambientes de desenvolvimento de IA, seja em pesquisa ou indústria, devem ser inclusivos e refletir a diversidade dos indivíduos e grupos da sociedade. A IA deve evitar o uso de dados adquiridos para bloquear indivíduos em um perfil de usuário, corrigir sua identidade pessoal ou confiná-los a uma bolha de filtragem, o que restringiria e confinaria suas possibilidades de desenvolvimento pessoal - especialmente em campos como educação, justiça ou o negócio. A IA não deve ser desenvolvido ou utilizado com o objetivo de limitar a liberdade de expressão de ideias ou a oportunidade de ouvir opiniões diversas, sendo ambas condições essenciais de uma sociedade democrática. Para cada categoria de serviço, a oferta da IA deve ser diversificada para evitar que monopólios de fato formem e minem as liberdades individuais (Ibid.).

<sup>37</sup> O Princípio da Prudência narra que é necessário desenvolver mecanismos que considerem o duplo uso potencial, benéfico e prejudicial, da pesquisa em IA e desenvolvimento de sistema de IA (seja público ou privado), a fim de limitar os usos prejudiciais. Quando o uso indevido de um sistema de IA põe em perigo a saúde ou segurança pública e tem alta probabilidade de ocorrência, é prudente restringir o acesso aberto e a divulgação pública ao seu algoritmo. Antes de serem colocados no mercado e sejam oferecidos gratuitamente ou gratuitamente, o sistema de IA deve atender a rígidos requisitos de confiabilidade, segurança e integridade e ser submetido a testes que não coloquem em risco a vida das pessoas, prejudiquem sua qualidade de vida, ou impactar negativamente sua reputação ou integridade psicológica. Esses testes devem ser abertos às autoridades públicas relevantes e às partes interessadas. O desenvolvimento do sistema de IA deve prevenir os riscos de uso indevido de dados do usuário e proteger a integridade e confidencialidade dos dados pessoais. Os erros e falhas descobertos no sistema e armazenamento de dados da IA devem ser compartilhados publicamente, em escala global, por instituições públicas e empresas em setores que representam um perigo significativo para a integridade pessoal e organização social (Ibid.).

<sup>38</sup> O Princípio da Responsabilidade assegura que, somente seres humanos podem ser responsabilizados pelas decisões decorrentes das recomendações da IA e pelas ações delas decorrentes. Em todas as áreas onde uma decisão que afete a vida, qualidade de vida ou reputação de uma pessoa deve ser tomada, onde o tempo e as circunstâncias permitirem, a decisão final deve ser tomada por um ser humano e essa decisão deve ser livre e informada. A decisão de matar deve ser sempre tomada por seres humanos, não devendo a responsabilidade por esta decisão ser transferida para um sistema de IA. As pessoas que autorizam a IA a cometer um crime ou uma infração, ou demonstram negligência ao permitir que o sistema de IA os cometas, são responsáveis por este crime ou infração. Quando um dano ou dano foi infligido por um sistema de IA, e o sistema de IA é provado ser confiável e ter sido usado conforme planejado, não é razoável colocar a culpa nas pessoas envolvidas em seu desenvolvimento ou uso. THE MONTREAL Declaration for the Responsible Development of Artificial Intelligence Launched. In: CAB Canada-ASEAN Business Council. Montreal: Universidade de Montreal, 12 de abril de 2018. Disponível em: <https://www.canasean.com/the-montreal-declaration-for-the-responsible-development-of-artificial-intelligence-launched/>. Acesso em: 22 ago. 2020.

do Desenvolvimento do Sustentável, que assinala que o desenvolvimento e a utilização do AIS devem ser realizados de forma a garantir uma forte sustentabilidade ambiental do planeta.<sup>39</sup> Assim, encerra-se uma série de dez princípios da *Montreal Declaration for the Responsible Development of Artificial Intelligence Launched*.

A Declaração canadense é inovadora do ponto de vista ético, pois observa a transparência não como algo necessário, mas obrigatório. Os princípios norteadores da Declaração destacam a importância dos sistemas de IA serem construídos sob a égide ética, isto é, a governança algorítmica observa o modus de captação e armazenamento dos dados para construção algorítmica. A declaração destaca três pilares para construção do sistema de IA canadense: (1) desenvolver uma estrutura ética para o desenvolvimento da IA; (2) orientar a transação digital para um benefício coletivo; e (3) criar um fórum nacional e internacional para a discussão dos avanços coletivos e desenvolvimento inclusivo e sustentável da IA.

No caminho de destacar os modelos estrangeiros mais relevantes, os Estados Unidos são outro país que observa a ética na utilização dos sistemas de IA e, assim como Alemanha, têm pretensões de assumir a liderança nas tecnologias de IA.<sup>40</sup> As principais diretrizes norte-americanas encontram-se no *American AI Initiative*. Um documento realizado pelo *Office of Science and*

---

<sup>39</sup> O Princípio do Desenvolvimento Sustentável assinala que o hardware do sistema de IA, sua infraestrutura digital e os objetos relevantes dos quais depende, como *data centers*, devem ter como objetivo a maior eficiência energética e mitigar as emissões de gases de efeito estufa ao longo de todo o seu ciclo de vida. O hardware de IA, sua infraestrutura digital e os objetos relevantes nos quais se baseia, devem ter como objetivo gerar a menor quantidade possível de resíduos elétricos e eletrônicos e proporcionar procedimentos de manutenção, reparação e reciclagem de acordo com os princípios da economia circular. O hardware de IA, sua infraestrutura digital e os objetos relevantes nos quais depende, devem minimizar nosso impacto sobre os ecossistemas e a biodiversidade em todas as fases de seu ciclo de vida, principalmente no que diz respeito à extração de recursos e ao descarte final do equipamento quando atingiu o fim da sua vida útil. Os atores públicos e privados devem apoiar o desenvolvimento ambientalmente responsável pelo sistema de IA, a fim de combater o desperdício de recursos naturais e bens produzidos, construir cadeias de abastecimento e comércio sustentáveis e reduzir a poluição global (Ibid.).

<sup>40</sup> Nas palavras do Presidente Donald Trump: “A liderança americana contínua em IA é de suma importância para manter a segurança econômica e nacional dos Estados Unidos e para moldar a evolução global de IA de uma maneira consistente com a de nossa nação valores, políticas e prioridades”. ESTADOS UNIDOS DA AMÉRICA. The White House. Office of Science and Technology Policy. **American Artificial Intelligence Initiative: Year One Annual Report**. Washington, DF, EUA: The White House, fevereiro de 2020. Disponível em: <https://www.whitehouse.gov/wp-content/uploads/2020/02/American-AI-Initiative-One-Year-Annual-Report.pdf>. Acesso em: 25 ago. 2020.

*Technology Policy* definiu os rumos do governo americano a respeito da Inteligência Artificial enfatizada em cinco políticas e práticas principais para o governo em IA, quais sejam: (i) investir em pesquisa e desenvolvimento de IA<sup>41</sup>; (ii) liberar recursos de IA<sup>42</sup>; (iii) remover as barreiras para a inovação em IA<sup>43</sup>; (iv) promover uma ambiente internacional de apoio à inovação em IA americana<sup>44</sup>; e (v) adotar uma IA confiável para serviços e missões do governo.<sup>45</sup>

Na linha da característica sintética da Constituição Americana, assim é o *American AI Initiative*, que trata objetivamente do tratamento dos sistemas de IA. Ao fim e ao cabo, também se assemelha ao modo característico negocial

---

<sup>41</sup> Os Estados Unidos devem promover o investimento federal em AI P&D em colaboração com a indústria, academia, parceiros e aliados internacionais e outras entidades não federais para gerar avanços tecnológicos em IA. O presidente Trump pediu um Duplicação de 2 anos de P&D de IA de não defesa em sua proposta de orçamento para o ano fiscal (AF) de 2021, e em 2019 a administração atualizou seu plano estratégico de P&D de IA, desenvolveu o primeiro relatório de progresso descrevendo o impacto dos investimentos federais em P&D, e publicou o primeiro relato de Gastos do governo em P&D de IA sem defesa (Ibid.).

<sup>42</sup> Os Estados Unidos devem melhorar o acesso a dados federais de alta qualidade, modelos e recursos de computação para aumentar seu valor para I&D de IA mantendo estendendo as proteções de segurança, privacidade e confidencialidade. The American AI Initiative a pelou às agências federais para identificar novas oportunidades para aumentar o acesso e uso de dados e modelos. Em 2019, o Escritório de Gestão e Orçamento da Casa Branca estabeleceu o Federal Data Strategy como uma estrutura para princípios operacionais e melhores práticas sobre como Agências federais usam e gerenciam dados (Ibid.).

<sup>43</sup> Os Estados Unidos devem reduzir as barreiras para o cofre desenvolvimento, teste, implantação e adoção de tecnologias de IA fornecendo orientação para a governança da IA consistente com os valores de nossa nação e impulsionando o desenvolvimento de padrões técnicos de IA apropriados. Como parte da American AI Initiative, The White House publicou para comentário os Princípios Regulatórios da AI dos Estados Unidos, o primeiro AI política regulatória que promove a inovação sustentada pelos valores e bens americanos práticas regulatórias. Além disso, o Instituto Nacional de Padrões e Tecnologia (NIST) emitiu a estratégia inédita para o envolvimento federal no desenvolvimento de técnicas de IA padrões (Ibid.).

<sup>44</sup> Promova um ambiente internacional de apoio à inovação em IA americana: The United Os Estados devem se engajar internacionalmente para promover um ambiente global que apoie a IA americana pesquisa e inovação e abre mercados para as indústrias americanas de IA ao mesmo tempo em que protege nossa vantagem tecnológica em IA. No ano passado, os Estados Unidos lideraram esforços históricos na Organização para a Cooperação e Desenvolvimento Econômico (OCDE) para desenvolver o primeiro consenso internacional acordos sobre princípios fundamentais para a administração de IA confiável. Estados Unidos também trabalhou com seus parceiros internacionais no G7 e G20 para adotar princípios de IA semelhantes (Ibid.).

<sup>45</sup> Os Estados Unidos devem adotar tecnologias como inteligência artificial para melhorar o fornecimento e a eficiência de serviços governamentais ao povo americano e garantir que sua aplicação mostre o devido respeito os valores de nossa nação, incluindo privacidade, direitos civis e liberdades civis. Os Serviços Gerais A administração estabeleceu um AI Center of Excellence para permitir que as agências federais determinassem práticas recomendadas para incorporar IA em suas organizações. ESTADOS UNIDOS DA AMÉRICA. The White House. Office of Science and Technology Policy. **American Artificial Intelligence Initiative: Year One Annual Report**. Washington, DF, EUA: The White House, fevereiro de 2020. Disponível em: <https://www.whitehouse.gov/wp-content/uploads/2020/02/American-AI-Initiative-One-Year-Annual-Report.pdf>. Acesso em: 26 ago. 2020

americano e não se afasta da busca incansável pela liderança geopolítica e econômica americana.

Logo, após a revisão das diretrizes da Resolução do Parlamento Europeu, da *Nationale KI-Strategie*, da *Montreal Declaration for the Responsible Development of Artificial Intelligence Launched*, e da *American AI Initiative*, resulta afirmar um rico arcabouço regulatório para fins de construção de sistema de IA no estrangeiro, que poderão servir de rumo ou de base para o desenvolvimento de diretrizes nacional para construção ética na IA. Nos modelos estrangeiros, visitou-se um objetivo comum das nações, o de conciliar o inafastável avanço da inteligência artificial nas rotinas dos seres humanos.

Nos modelos de diretrizes éticas estrangeiras, observou-se a possibilidade de inserção da IA por meio dos algoritmos sem a identificação dos construtores, de modo a evitar uma eventual etiquetagem dos profissionais da ciência da computação, bem como uma notória e necessária supervisão humana em observância aos direitos fundamentais.

No âmbito do Direito brasileiro, observando as diretrizes éticas globais no que tange os sistemas de IA, destaca-se a presença da Lei de Proteção de Dados (LGPD), que, no artigo 2º disciplina a proteção de dados, guarda relação direta com a transparência algorítmica. No referido artigo, está expresso a observância à privacidade e, também, à inviolabilidade da intimidade, da honra e da imagem, o que se relaciona diretamente com a ética nos sistemas de IA. Ainda, em suma, conclui-se que os avanços dos sistemas de IA poderão, com coerência e integridade,<sup>46</sup> auxiliar no bom desempenho do Direito.

Dito isso, no decorrer desta tese, exsurgiu, no Brasil, um importante marco para definição de diretrizes que disciplina a ética, transparência e governança na

---

<sup>46</sup> Em suma, a teoria dworkiniana mostra-se inexoravelmente coerente com a atribuição de sentidos derivada dos valores, que como se pôde notar, sempre andam em duplicidade de modo que se tenha sempre uma paridade de valores colocados em discussão, exatamente onde repousa a coerência no uso da moral. O direito para Dworkin é visto como um ramo da moral, apesar dos atributos que o particularizam. Já em termos de filosofia política, os direitos fundamentais são vistos como espécies de trunfos contramajoritários, e vinculam-se à ideia de dignidade de que Dworkin lança mão, da qual parte o direito a ser tratado com igual consideração e respeito, e divide a dignidade em dois princípios: (i) o direito a ser tratado com igual consideração e respeito e (ii) o de que os indivíduos têm a responsabilidade pessoal de definir o sucesso em suas próprias vidas. Trata-se de uma concepção nitidamente liberal quanto à concepção dos valores morais e, por conseguinte, do direito, na medida em que Dworkin o vincula a essa ideia de dignidade. DWORKIN, Ronald. **A raposa e o porco-espinho**: justiça e valor. Tradução de Marcelo Brandão Cipolla. São Paulo: Martins Fontes, 2014. p. 513

produção de sistemas de IA.<sup>47</sup> Foi publicada, no dia 25 de agosto de 2020, a Resolução do Conselho Nacional de Justiça (CNJ) n. 332/2020, que tratou de diversos aspectos importantes.

Inicialmente, a Resolução n. 332/2020 do CNJ já trata da capacidade da Inteligência Artificial contribuir na diminuição de tempo em tarefas repetitivas.<sup>48</sup> E, de modo relevante, ao contribuir com a agilidade e coerência do processo de tomada de decisão, destaca-se que os tribunais deverão observar os Direitos Fundamentais.<sup>49</sup> Essa tomada de decisão deve atender a critérios éticos,<sup>50</sup> ou seja, à decisão fruto da aprendizagem de máquina deve ser transparente e previsível, oferecendo as condições de possibilidades para eventual auditoria.

No tocante, tais decisões, além de guardarem a possibilidade de auditoria, também não devem ser discriminatórias ou desiguais, justamente para evitar eventuais julgamentos errôneos ou discriminatórios.<sup>51</sup> O afastamento desses

---

<sup>47</sup> Com o avanço dos sistemas de inteligência artificial (IA), cada vez mais onipresentes, a governança de IA para a tomada de decisão ética pela IA passou a ser de interesse público da comunidade. No tocante, tornou-se indiscutível a descoberta de soluções técnicas para governança de IA, visto que a IA demanda preocupações éticas atuais exatamente por resultar situações limites que ultrapassam a velocidade de processamento, capacidade de armazenamento, busca de informações, padronizações típicas de sistemas de automação. Para Hartmann, há dois fatores fundamentais: a execução de atividades cognitivas fruto de sistemas de aprendizagem de máquina, ou no mínimo, a delimitação do conteúdo sobre o qual a cognição humana irá atuar para promover decisões e a interconexão do raciocínio jurídico como o raciocínio exato. HARTMANN, Fabiano Peixoto; SILVA, Roberta Zumblick M. da. **Inteligência artificial e direito**. Curitiba: Alteridade Editora, 2019. p. 20.

<sup>48</sup> A Inteligência Artificial, ao ser aplicada no Poder Judiciário, pode contribuir com a agilidade e coerência do processo de tomada de decisão. BRASIL. Conselho Nacional de Justiça. **Resolução Nº 332 de 21 de agosto de 2020**. Dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário e dá outras providências. Brasília, DF: Conselho Nacional de Justiça, 21 de agosto de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 06 ago. 2020.

<sup>49</sup> No desenvolvimento e na implantação da Inteligência Artificial, os tribunais deverão observar sua compatibilidade com os Direitos Fundamentais (Ibid.).

<sup>50</sup> A Inteligência Artificial aplicada nos processos de tomada de decisão deve atender a critérios éticos de transparência, previsibilidade, possibilidade de auditoria e garantia de imparcialidade e justiça substancial (Ibid.).

<sup>51</sup> As decisões judiciais apoiadas pela Inteligência Artificial devem preservar a igualdade, a não discriminação, a pluralidade, a solidariedade e o julgamento justo, com a viabilização de meios destinados a eliminar ou minimizar a opressão, a marginalização do ser humano e os erros de julgamento decorrentes de preconceitos (Ibid.).

limites, de modo a observar os direitos fundamentais, deve estar amparado numa base de dados seguros<sup>52</sup>, e também passíveis de auditoria.<sup>53</sup>

Conforme destacado anteriormente, a LGPD, em seu artigo 2º, já disciplina a proteção de dados. Assim, a Resolução n. 332 do CNJ explica que os sistemas de IA devem respeitar a privacidade dos usuários, com controle e proteção sobre os usos destes dados pessoais.<sup>54</sup>

A utilização da IA deve atuar não para gerar um aumento de produtividade de resoluções processuais, mas promover a igualdade, a liberdade e a justiça, fomentando a dignidade humana. Aqui cabe uma observação: muito embora o haja caráter louvável e em tempo da Resolução n. 332 do CNJ, essas diretrizes de igualdade, liberdade e justiça são indiscutivelmente subjetivas. Portanto, abre-se um leque muito grande para atuação da IA, que deve ser limitada em razão das diretrizes específicas iniciais da referida Resolução do CNJ, uma vez que são tangíveis.<sup>55</sup>

Por fim aqui, em vez de se propor um modelo ético no uso da IA no Direito, algo pretencioso diante da ausência de um marco regulatório, optou-se em revisar as diretrizes dos países estrangeiros que já desenvolveram intenções e documentos factíveis de observância ética na IA, de modo a estudar os

---

<sup>52</sup> A Resolução do Parlamento Europeu traz que os dados e /ou os procedimentos de baixa qualidade em que se baseiam os processos de tomada de decisão e os instrumentos analíticos podem traduzir-se em algoritmos parciais, correlações ilegítimas, erros, numa subestimação das implicações jurídicas, sociais e éticas, no risco de utilização de dados para fins discriminatórios ou fraudulentos e na marginalização do papel dos seres humanos nestes processos, podendo resultar em processos imperfeitos de tomada de decisão, com um impacto nocivo nas vidas e nas oportunidades dos cidadãos, mormente nos grupos marginalizados, bem como em consequências negativas para as sociedades e as empresas. UNIÃO EUROPEIA. Parlamento Europeu. **Implicações dos grandes volumes de dados nos direitos fundamentais.** Resolução do Parlamento Europeu, de 14 de março de 2017, sobre as implicações dos grandes volumes de dados nos direitos fundamentais: privacidade, proteção de dados, não discriminação, segurança e aplicação da lei (2016/2225(INI)). [S.l.]: Parlamento Europeu, 14 de março de 2017. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52017IP0076&from=DA>. Acesso em: 07 nov. 2020.

<sup>53</sup> Os dados utilizados no processo de aprendizado de máquina deverão ser provenientes de fontes seguras, preferencialmente governamentais, passíveis de serem rastreados e auditados. No tratamento, os dados utilizados devem ser eficazmente protegidos contra riscos de destruição, modificação, extravio, acessos e transmissões não autorizadas. BRASIL. Conselho Nacional de Justiça. **Resolução Nº 332 de 21 de agosto de 2020.** Dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário e dá outras providências. Brasília, DF: Conselho Nacional de Justiça, 21 de agosto de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 25 ago. 2020.

<sup>54</sup> O uso da Inteligência Artificial deve respeitar a privacidade dos usuários, cabendo-lhes ciência e controle sobre o uso de dados pessoais. Os dados coletados pela Inteligência Artificial devem ser utilizados de forma responsável para proteção do usuário (Ibid.).

<sup>55</sup> A utilização da Inteligência Artificial deve se desenvolver com vistas à promoção da igualdade, da liberdade e da justiça, bem como para garantir e fomentar a dignidade humana (Ibid.).

referidos modelos para eventual adaptação e auxílio no marco regulatório da IA no Direito brasileiro. Contudo, no decorrer desta tese, exsurgiu a Resolução n. 332 do CNJ exatamente em virtude da ausência de normas específicas quanto à governança e aos parâmetros éticos para o desenvolvimento dos sistemas de IA no Poder Judiciário.

Enfim, convém observar que é fundamental a discussão das diretrizes éticas dos países estrangeiros, de modo a embasar vindouras e necessárias construções de diretrizes éticas no arcabouço brasileiro para uma construção de sistemas de IA integras e coerentes. Destaca-se a fundamental revisão dessas diretrizes éticas na construção dos sistemas IA no Brasil, exatamente pelo fato da constante construção de algoritmos que se faz necessária, em virtude do potencial de dados produzidos em tempo real no mundo e no direito. No Direito, em especial, posto que a fundamentação é direito fundamental das partes, o que guarda relação direta com a transparência algorítmica e ética na construção dos sistemas de IA. Nesse sentido, a questão ética, que guarda relação direta com a transparência algorítmica, será importante para responder à pergunta motor desta tese no próximo e último subcapítulo.