

# **Inteligência Artificial: Riscos para direitos humanos e possíveis ações**

Paula Gorzoni

## **RESUMO**

O desenvolvimento de Inteligência Artificial (IA) pode trazer consigo riscos para os direitos humanos sob o ponto de vista de suas implicações sociais, éticas e jurídicas, tanto da perspectiva individual quanto coletiva. Para garantir que o desenvolvimento de IA ocorra para o benefício da sociedade como um todo, de forma inclusiva, é necessário elencar os riscos existentes, compreendê-los e pensar em formas de mitigá-los. Levando em consideração estes riscos, o artigo analisará as possíveis ações para a sua mitigação e a possibilidade do desenvolvimento de IA no âmbito de uma estrutura ética. Para isso, os esforços da União Europeia (UE) no desenvolvimento de diretrizes éticas nesta área serão analisados.

## **INTRODUÇÃO**

Novas tecnologias de inteligência artificial têm se desenvolvido nos mais diferentes campos. Desde o sistema de recomendação de filmes, séries e músicas utilizados por sites ou aplicativos como, por exemplo, YouTube, Netflix e Spotify, ou o sistema de anúncios utilizado pelo Facebook, o mecanismo que identifica o que é spam na caixa de e-mails, passando por assistentes virtuais inteligentes que interagem pelo reconhecimento de voz, até carros inteligentes, todos esses exemplos utilizam técnicas de inteligência artificial em seu funcionamento.

No entanto, apesar de técnicas de IA estarem já espalhadas no dia-a-dia das pessoas, é comum que a sociedade olhe com temor o desenvolvimento destas novas técnicas.<sup>1</sup> O imaginário de uma distopia é visto, por exemplo, na série *Black Mirror*. “Robôs de guerra” ou sistemas de armamento autônomos já são tematizados.<sup>2</sup> Elon Musk, fundador e CEO da Tesla, tem alertado na mídia os perigos levantados pela IA e chamado atenção para a importância de se regular o desenvolvimento deste tipo de tecnologia.<sup>3</sup> Nesse sentido, é necessário reconhecer a velocidade em que IA está se desenvolvendo e examinar os riscos que técnicas de IA podem trazer consigo.<sup>4</sup>

## DIREITOS HUMANOS NA ERA DIGITAL

### Pontos iniciais de debate: definindo cenários e desafios

Ponto inicial para a avaliação de riscos aos direitos humanos que a inteligência artificial pode causar é definir estes direitos no contexto da era digital. O Conselho de Direitos Humanos da Organização das Nações Unidas (ONU) definiu em 2016 que “direitos humanos que as pessoas têm *offline* devem ser também protegidos *online*”.<sup>5</sup> Isso significa que tratados internacionais sobre direitos humanos definidos antes do desenvolvimento destas tecnologias, como por exemplo o Pacto Internacional de Direitos Civis e Políticos (1966), aplicam-se da mesma forma tanto no cenário *online* quanto no *offline*.

No entanto, pergunta-se se as legislações internacionais e nacionais relativas à proteção de direito humanos são suficientes neste cenário de

---

<sup>1</sup> Ver Daniel Cullen, Why Artificial Intelligence is Already a Human Rights Issue, Oxford Human Rights Hub, 2018. Disponível em: <<http://ohrh.law.ox.ac.uk/why-artificial-intelligence-is-already-a-human-rights-issue/>>. Acesso: 26 de janeiro de 2019.

<sup>2</sup> Cf. Lorna McGregor, Why We Need to Stop Talking About ‘Killer Robots’ and Address the AI Backlash, Blog of the European Journal of International Law, EJIL: Talk!, 9 de julho de 2018. Disponível em: <<https://www.ejiltalk.org/why-we-need-to-stop-talking-about-killer-robots-and-address-the-ai-backlash/>>. Acesso em: 24 de janeiro de 2019.

<sup>3</sup> Ver, por exemplo, “Elon Musk: regulate AI to combat 'existential threat' before it's too late”, The Guardian, 17 julho de 2017. Disponível em: <<https://www.theguardian.com/technology/2017/jul/17/elon-musk-regulation-ai-combat-existential-threat-tesla-spacex-ceo>>. Acesso: 24 de janeiro de 2019.

<sup>4</sup> Cf. Lorna McGregor, Why We Need to Stop Talking About ‘Killer Robots’ and Address the AI Backlash.

<sup>5</sup> Conselho de Direitos Humanos da ONU, A/HRC/32/L.20, 27 de junho 2016.

desenvolvimento de novas tecnologias. Para analisar esta questão, é necessário primeiramente elencar os riscos trazidos aos direitos humanos por técnicas de inteligência artificial.

## **Riscos para direitos humanos**

Ao analisar o desenvolvimento de inteligência artificial é preciso levar em consideração que as tecnologias são construídas frequentemente por organizações, como empresas específicas e por um grupo de pessoas específicas (por exemplo, determinados cientistas em uma organização).<sup>6</sup> Desta forma, novas tecnologias são influenciadas por e replicam normas sociais ou valores da sociedade atual.<sup>7</sup> Por outro lado, estas novas tecnologias moldam a sociedade e mudam hábitos sociais de forma rápida. Por isso é essencial que o desenvolvimento de IA leve em consideração os possíveis riscos aos direitos humanos.

### Privacidade e proteção de dados

Um dos riscos mais levantados corresponde a violações do direito à privacidade e proteção de dados.<sup>8</sup> O direito à privacidade está elencado no artigo 17 do Pacto Internacional de Direitos Civis e Políticos. Tecnologias de inteligência artificial têm sido utilizadas, por exemplo, para a criação de sistemas de decisões automatizadas. Tais mecanismos são capazes de tomar decisões automaticamente, fazer previsões ou sugerir recomendações de ações, sem intervenção humana. Para a criação desses sistemas é necessário um grande volume de dados pessoais para que um algoritmo possa ser alimentado e seja

---

<sup>6</sup> Cf. European Commission, *Artificial Intelligence: A European Perspective*. Luxembourg: Publications Office of the European Union, 2018, p. 55.

<sup>7</sup> European Commission, *Artificial Intelligence: A European Perspective*, p. 55.

<sup>8</sup> Ver, por exemplo, Conselho da Europa, *Study on the Human Rights Dimensions of Automated Data Processing Techniques (in particular algorithms) and possible regulatory implications*, 2017, p. 12.

capaz de tomar uma decisão com certa precisão.<sup>9</sup> O algoritmo aprende por meio do enorme volume de dados disponível e com isso é capaz de tomar decisões por si mesmo. Tais decisões podem afetar indivíduos, grupos sociais ou a sociedade como um todo. Desta forma, o enorme volume de dados e a forma como estes dados são utilizados por estes sistemas devem ser jurídica e eticamente regulados.

Neste sentido, é necessário pensar: (1) quais dados podem ser utilizados para o desenvolvimento de inteligência artificial e (2) como os dados serão processados pelos sistemas. Existem dados que são considerados como sensíveis, como, por exemplo, dados sobre religião ou orientação sexual. Tais dados não podem ser utilizados a fim de se gerar discriminação em relação a um determinado grupo. Existe, neste sentido, o risco da utilização de dados pessoais para definições de perfis (“*profiling*”).<sup>10</sup> A respeito do processamento de dados, o relatório sobre “inteligência artificial e privacidade” da autoridade norueguesa de proteção de dados (Datatilsynet) aponta os seguintes desafios da inteligência artificial: limitação de propósito do uso dos dados; minimização do uso de dados; e transparência e direito à informação relacionados aos dados usados.<sup>11</sup>

A não limitação do propósito do uso dos dados pode gerar complicações, visto que é possível gerar novos dados por meio do cruzamento ou agregação de dados de diferentes sistemas. Ao se criar tecnologias de inteligência artificial, o propósito do uso de dados deve estar claro desde o início. Além disso, é necessário pensar na segurança dos dados armazenados. Novas tecnologias devem ser desenvolvidas tendo também como objetivo proteger os dados utilizados de terceiros não autorizados a terem acesso ao banco de dados.

Uma possível abordagem para mitigar riscos relacionados a estes direitos é conhecida como “privacidade por desenho” (*privacy by design*), que significa

---

<sup>9</sup> Ver The Norwegian Data Protection Authority, Artificial Intelligence and privacy report, p. 6, Information Commissioner’s Office (ICO), “Big data, artificial intelligence, machine learning and data protection”, p. 7.

<sup>10</sup> Ver CoE, p. 13.

<sup>11</sup> The Norwegian Data Protection Authority, Artificial Intelligence and privacy report, 2018, p. 4

que a proteção da privacidade ou a proteção de dados deve ser pensada desde a concepção da tecnologia de inteligência artificial, a fim de se minimizar maiores violações aos direitos.<sup>12</sup> Regulamentações como o Regulamento Geral sobre a Proteção de Dados (RGPD) da União Europeia (UE)<sup>13</sup> e a Lei Geral de Proteção de Dados Pessoais brasileira (Lei Federal n. 13.709/2018) representam uma tentativa de se disciplinar o uso de dados nestas situações.

## Discriminação

Sistemas de inteligência artificial, por meio de processamento de dados, podem trazer riscos como discriminação de indivíduos ou grupos.<sup>14</sup> A proibição da discriminação está presente no artigo 26 do Pacto Internacional de Direitos Civis e Políticos. Como a tecnologia é muitas vezes criada por determinados grupos de pessoas ou empresas com interesses específicos, sistemas baseados em inteligência artificial podem reproduzir comportamentos discriminatórios presentes na sociedade.<sup>15</sup> Além disso, a partir do momento em que um algoritmo aprende por ele mesmo a partir de dados que reproduzem comportamentos arraigados na sociedade, condutas discriminatórias podem ser adotadas automaticamente. Estes riscos podem ser percebidos quando a inteligência artificial é utilizada na seleção de candidatos em entrevista de emprego<sup>16</sup> ou concessão de crédito por um banco.

Conforme mencionado anteriormente, dados sensíveis, como informações sobre raça, religião, orientação sexual ou posicionamento político,

---

<sup>12</sup> Ver Lea Gimpel, “Talking about Big Data: The Challenge of Privacy & Data Protection in International Development”, disponível em: <<http://blogs.die-gdi.de/2017/09/08/talking-about-big-data/>>. (Acesso: 13 de maio de 2018).

<sup>13</sup> O RGPD foi criado em 27 de abril de 2016 e entrou em vigor no dia 25 de maio de 2018.

<sup>14</sup> Ver, por exemplo, European Union Agency for Fundamental Rights, #BigData: Discrimination in data-supported decision making, Luxemburgo: Publications Office, 2018, p. 3.

<sup>15</sup> Ver, por exemplo, Gideon Mann, Cathy O'Neil, “Hiring Algorithms Are Not Neutral” Harvard Business Review 2016. Disponível em <<https://hbr.org/2016/12/hiring-algorithms-are-not-neutral>> (Acesso: 13 de março de 2018).

<sup>16</sup> Ver, por exemplo, Oren Danieli, Andrew Hillis, Michael Luca, “How to Hire with Algorithms”, Harvard Business Review 2016, disponível em <[https://hbr.org/2016/10/how-to-hire-with-algorithms?referral=03758&cm\\_vc=rr\\_item\\_page\\_top\\_right](https://hbr.org/2016/10/how-to-hire-with-algorithms?referral=03758&cm_vc=rr_item_page_top_right)> (Acesso: 13 de março de 2018).

podem ser processados de forma discriminatória. Uma das soluções possíveis diz respeito à necessidade de revisão ou verificação constante por pessoas das decisões automatizadas, assim como a avaliação crítica da base de dados e da estrutura de dados disponibilizada para mecanismos de inteligência artificial.<sup>17</sup>

Existe também o risco de que sistemas de inteligência artificial não considerem situações específicas de grupos marginalizados, visto que dados destes grupos podem não estar incluídos no volume de dados a ser processado por determinados sistemas. Este é o caso de um programa de reconhecimento facial que era menos preciso ao reconhecer mulheres negras.<sup>18</sup> Dessa forma, existe o risco de perpetuação da exclusão de grupos na sociedade.<sup>19</sup>

#### Julgamento imparcial e devido processo legal

Sistemas de inteligência artificial que são utilizados para auxiliar tomadas de decisões no âmbito judicial podem trazer riscos relacionados à imparcialidade de um julgamento e ao devido processo legal.<sup>20</sup> Estas garantias estão presentes no artigo 14 do Pacto Internacional de Direitos Civis e Políticos. Estes riscos estão presentes no caso, por exemplo, do Compas (sigla em inglês para *Correctional Offender Management Profiling for Alternative Sanctions*), mecanismo dotado de algoritmo supostamente capaz de prever a probabilidade de reincidência criminal de uma pessoa.<sup>21</sup> Há riscos de um sistema como este

---

<sup>17</sup> Ver Stephan Dreyer e Wolfgang Schulz, *Was bringt die Datenschutz-Grundverordnung für automatisierte Entscheidungssysteme?*, Gütersloh: Bertelsmann Stiftung, 2018, p. 14.

<sup>18</sup> Ver Joy Buolamwini e Timnit Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, *Proceedings of Machine Learning Research* 81, 2018, pp. 1-15. Ver também Joy Buolamwini, *How I'm fighting bias in algorithms*, 2017. Disponível em: <[https://www.ted.com/talks/joy\\_buolamwini\\_how\\_i\\_m\\_fighting\\_bias\\_in\\_algorithms](https://www.ted.com/talks/joy_buolamwini_how_i_m_fighting_bias_in_algorithms)>. Acesso: 26 de janeiro de 2019.

<sup>19</sup> Cf. Stephan Dreyer e Wolfgang Schulz, *Was bringt die Datenschutz-Grundverordnung für automatisierte Entscheidungssysteme?*, p. 15.

<sup>20</sup> Nesse sentido, ver estudo da ONG Article 19, “Algorithms and automated decision-making in the context of crime prevention”. Disponível em <<https://www.article19.org/resources/algorithms-and-automated-decision-making-in-the-context-of-crime-prevention/>>. Acesso em: 25 de janeiro de 2019.

<sup>21</sup> Ver “Sent to Prison by a Software Program’s Secret Algorithms”, *The New York Times*, 1 de maio de 2017. Disponível em <<https://www.nytimes.com/2017/05/01/us/politics/sent-to-prison-by-a-software-programs-secret-algorithms.html>>. (Acesso: 13 de março de 2018).

reproduzir comportamentos discriminatórios, conforme salientado no tópico anterior. Além disso, um sistema como este pode carecer de transparência, quando o uso de dados e modo como as decisões são tomadas não são explicados de forma clara.

### Liberdade de expressão, reunião e associação

A Remoção de conteúdo em plataformas de redes sociais como Facebook é feita frequentemente por sistemas automatizados ou semiautomatizados, que utilizam inteligência artificial para seu funcionamento. Neste caso, existe risco à liberdade de expressão (artigo 19 do Pacto Internacional de Direitos Civis e Políticos) por meio de remoções de conteúdo sem maiores justificativas, feitas por agente privado. Publicações com conteúdo humorístico, por exemplo, são com frequência de difícil interpretação e por meio destes mecanismos podem sofrer remoções automatizadas.

Mecanismos de inteligência artificial também podem ser utilizados no combate ao terrorismo ou ao extremismo político por meio de reconhecimento facial, filtros de conteúdo na internet e remoção automática de conteúdo.<sup>22</sup> Nestes casos, existem também riscos à liberdade de expressão por meio de “super bloqueios” (*overblocking*) de conteúdo.

Por fim, mecanismos de remoção automática de conteúdo e reconhecimento facial também podem gerar riscos para defensores de direitos humanos e opositores políticos, como o risco do estabelecimento de uma atmosfera de vigilância dependendo da forma com que governos podem se apropriar e utilizar tais mecanismos.<sup>23</sup> Consequentemente, tal cenário representa também um risco para a liberdade de associação e reunião.

---

<sup>22</sup> Ver, por exemplo, Facebook: Monika Bickert, Brian Fishman, Hard Questions: How We Counter Terrorism, 15 de Junho de 2017. Disponível em: <https://newsroom.fb.com/news/2017/06/how-we-counter-terrorism/>. Acesso: 25 de janeiro de 2019.

<sup>23</sup> Cf. Conselho da Europa, Study on the Human Rights Dimensions of Automated Data Processing Techniques (in particular algorithms) and possible regulatory implications, p. 16.

## Eleições

Mecanismos de inteligência artificial podem também trazer riscos às eleições por meio dos “*social media bots*”.<sup>24</sup> *Bots* são robôs que executam tarefas automaticamente. *Bots* em redes sociais podem assumir identidades por meio de contas falsas e divulgar opiniões a respeito de candidatos ou temas sensíveis politicamente. Por meio de mecanismos de inteligência artificial, estes *bots* conseguem interagir com outros usuários e divulgar informações conforme a situação. Os *bots* nas redes sociais podem, portanto, imitar comportamentos humanos relacionados a padrões de comunicação. Deste modo, tais mecanismos podem influenciar de forma considerável debates em eleições. Existe a possibilidade de *bots* manipularem opiniões e espalharem informações falsas nas redes sociais, constituindo um risco para a democracia.<sup>25</sup> Existe um risco aos direitos relacionados ao estabelecimento de eleições livres, conforme o disposto no artigo 25 do Pacto Internacional de Direitos Civis e Políticos e do artigo 3 do Protocolo 1 da Convenção Europeia de Direitos Humanos.

Além disso, por meio dos filtros nas redes sociais, há a possibilidade da criação de “bolhas ideológicas”, o que significa que o usuário somente visualizará publicações que se assemelhem com seus interesses e opiniões.<sup>26</sup>

## Direito ao trabalho

Mecanismos de inteligência artificial podem trazer riscos relacionados ao direito ao trabalho, disposto no artigo 6 do Pacto Internacional dos Direitos Econômicos, Sociais e Culturais. É possível que existam riscos para a empregabilidade.<sup>27</sup> Sistemas de inteligência artificial podem vir a substituir mão de obra humana e, deste modo, provocar desemprego em certos setores. Um

---

<sup>24</sup> Conselho da Europa, Study on the Human Rights Dimensions of Automated Data Processing Techniques (in particular algorithms) and possible regulatory implications, p. 31.

<sup>25</sup> Cf. Alessandro Bessi e Emilio Ferrara, “Social bots distort the 2016 U.S. Presidential election online discussion”, First Monday 21, 2016.

<sup>26</sup> Conselho da Europa, Study on the Human Rights Dimensions of Automated Data Processing Techniques (in particular algorithms) and possible regulatory implications, p. 29-30.

<sup>27</sup> Cf. European Commission, Artificial Intelligence: A European Perspective, p. 77.



dos setores que mais apontado como possivelmente afetado pela inteligência artificial no futuro é o de telemarketing.<sup>28</sup> Para evitar este risco, é essencial pensar em como trabalhadores e sistemas de inteligência artificial podem interagir, pensando em novas formas de trabalho e uma educação consciente das mudanças que estarão por vir, tendo como objetivo aumentar o entendimento da população em relação às novas tecnologias.<sup>29</sup>

## **POSSÍVEIS AÇÕES PARA MITIGAÇÃO DE RISCOS**

Após ter analisado os possíveis riscos da inteligência artificial aos direitos humanos, possíveis ações para a mitigação destes riscos serão aqui elencadas. Importante perceber que no âmbito do desenvolvimento de novas tecnologias, a proteção aos direitos humanos não deve ser só pensada segundo a visão tradicional, na qual tais direitos foram concebidos como direitos oponíveis somente ao Estado, com a função de proteger os indivíduos contra os abusos daquele. Agentes privados, como empresas, desempenham papel importante no desenvolvimento de inteligência artificial e podem, portanto, também restringir direitos. É necessário adotar perspectiva desses direitos como princípios que atuam em todos os âmbitos do direito - inclusive o direito privado - e que proporcionam diretrizes e impulsos para a legislação, administração e justiça.<sup>30</sup> Por outro lado, também é preciso pensar quem poderia regulamentar e fiscalizar o emprego de inteligência artificial por setores públicos ou pelo poder judiciário.

Nesse sentido, possíveis ações para mitigação de riscos devem ser pensadas a partir de uma abordagem multissetorial, que envolvam poder público (legislativo, executivo e judiciário), setor privado, academia e sociedade civil.<sup>31</sup>

---

<sup>28</sup> Cf. McKinsey Global Institute, *Skill Shift Automation and the Future of the Workforce*, Discussion Paper, 2018 p. 37.

<sup>29</sup> Cf. European Commission, *Artificial Intelligence: A European Perspective*, p. 78.

<sup>30</sup> Sobre direitos fundamentais nas relações entre particulares, ver Juan Maria Bilbao Ubillos, *La eficacia de los derechos fundamentales frente a particulares: análisis de la jurisprudencia del Tribunal Constitucional*. Madrid: Centro de Estudios Políticos y Constitucionales, 1997, p. 262.

<sup>31</sup> Ver, neste sentido, Urs Gasser, *AI and the Law: Setting the Stage*, 2017. Disponível em: <<https://medium.com/berkman-klein-center/ai-and-the-law-setting-the-stage-48516fda1b11>>. Acesso: 26

Existe também a possibilidade de se pensar em um órgão independente, formado por pessoas de todos os setores, que possa regular e monitorar o desenvolvimento de IA. Desta forma, os mais diferentes interesses da sociedade estariam representados.

### **Abordagens baseadas em direitos humanos**

Considerando uma abordagem multissetorial, outro pressuposto essencial para este debate sobre mitigação de riscos é pensar no desenvolvimento de novas tecnologias por meio de uma abordagem baseada em direitos humanos:

*“A human rights-based approach (HRBA) is a conceptual framework for the process of human development. It is based on international human rights norms and standards and is dedicated to promoting and protecting human rights. It puts human rights at the center of the preparation, design, implementation, monitoring, and evaluation of policies, regulatory measures, and spending programmes. International human rights norms and standards, as well as principles such as participation, non-discrimination, and accountability should guide all stages of policies, regulatory measures, and spending programmes.”<sup>32</sup>*

Isso significa que direitos humanos devem ser levados em consideração desde o início do desenvolvimento da tecnologia de inteligência artificial.

Além disso, instrumentos já existentes como, por exemplo, princípios orientadores sobre empresas e direitos humanos da ONU,<sup>33</sup> podem ser úteis para a mitigação desses riscos. Os princípios orientadores sugerem que

---

de janeiro de 2019 e Luciano Floridi, Artificial Intelligence and the 'Good Society': the US, EU, and UK approach, *Science and engineering ethics* 24 (2), 2018, p. 523.

<sup>32</sup> United Nations Human Rights Regional Office for Europe, *Making a Difference: An Introduction to Human Rights*, 2018, p. 14. Disponível em: <[https://europe.ohchr.org/Documents/Publications/MakeADifference\\_EN.pdf](https://europe.ohchr.org/Documents/Publications/MakeADifference_EN.pdf)>. Acesso: 25 de janeiro de 2018.

<sup>33</sup> Princípios orientadores sobre empresas e direitos humanos da ONU, 2011. Disponível em: <[https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR\\_EN.pdf](https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf)>. Acesso: 25 de janeiro de 2019.

empresas dos mais diversos setores façam uso de “*due diligence* em direitos humanos”, que corresponde a um processo de administração de riscos a direitos humanos. Este processo deve incluir avaliações de impactos em direitos fundamentais, integração dos resultados das avaliações nos processos relevantes da empresa por meio de medidas de acordo com o seu envolvimento no impacto, acompanhamento da eficácia dessas medidas para verificar sua efetividade, comunicação de como os impactos estão sendo abordados e mostrar às partes interessadas - em particular às partes interessadas afetadas - que existem políticas e processos adequados em vigor.<sup>34</sup> Atores privados poderiam fazer uso deste mecanismo no processo de desenvolvimento de inteligência artificial e, após o desenvolvimento, no processo de monitoramento da performance dos sistemas. Nesse sentido, a ICANN (Sigla para “corporação da internet para atribuição de nomes e números”) vem discutindo a possibilidade de traçar modelos de avaliações de impacto a direitos humanos no setor de tecnologia.<sup>35</sup> Dentre as iniciativas da ICANN, destaca-se o esboço de um questionário com critérios a serem avaliados no âmbito de tais avaliações.<sup>36</sup>

Em uma abordagem baseada em direitos humanos, é preciso também considerar a possibilidade de regulação própria de atores privados neste sentido. Um exemplo deste tipo de iniciativa são os princípios que a Google estabeleceu para a criação de inteligência artificial.<sup>37</sup> Dentre os princípios, a empresa se

---

<sup>34</sup> Princípios orientadores sobre empresas e direitos humanos da ONU, p. 17 (versão em inglês). Mais informações, ver OHCHR, Corporate human rights due diligence – identifying and leveraging emerging practice, disponível em: <<https://www.ohchr.org/EN/Issues/Business/Pages/CorporateHRDueDiligence.aspx>>. Acesso: 25 de janeiro de 2019.

<sup>35</sup> Ver, por exemplo, discussões feitas no âmbito do painel “#349 A Multistakeholder Approach to HRIAs: Lessons from ICANN” no Internet Governance Forum (IGF) em 2018. Disponível em: <<http://www.intgovforum.org/multilingual/content/igf-2018-ws-349-a-multistakeholder-approach-to-hrias-lessons-from-icann?fbclid=IwAR1wU0pKVLaznpPiWnaQQ9Ij4COaanBr3DG37iqKezMQ2bx2QZPAJrn1Bj4>>. Acesso: 25 de janeiro de 2019.

<sup>36</sup> Questionário disponível em: <[https://icannhumanrights.net/wp-content/uploads/2018/05/DRAFT-ICANN-HRIA-questionnaire-model.pdf?fbclid=IwAR1wITDPMYJ\\_i6POKAAik1jPN3\\_GbDFIGrcEbRyZfy38DMKLPZnf3xVMTc](https://icannhumanrights.net/wp-content/uploads/2018/05/DRAFT-ICANN-HRIA-questionnaire-model.pdf?fbclid=IwAR1wITDPMYJ_i6POKAAik1jPN3_GbDFIGrcEbRyZfy38DMKLPZnf3xVMTc)>. Acesso em: 25 de janeiro de 2019.

<sup>37</sup> Cf. Lorna McGregor e Vivian Ng, “Google’s new principles on AI need to be better at protecting human rights”, The Human Rights, Big Data and Technology Project, 2018. Disponível em: <<https://hrbdt.ac.uk/googles-new-principles-on-ai-need-to-be-better-at-protecting-human-rights/>>.

compromete a evitar criar ou reforçar preconceitos, incorporar princípios de privacidade *by design*, e se compromete a não buscar tecnologias que causam ou provavelmente causarão danos gerais à sociedade.

### **Diretrizes éticas para inteligência artificial**

Além de uma abordagem baseada em direitos humanos, a criação e o desenvolvimento de inteligência artificial devem ser pautados por diretrizes éticas. Os direitos humanos em si representam uma fonte para o desenvolvimento da ética de inteligência artificial.<sup>38</sup> Tais diretrizes podem servir como uma forma de esclarecer quais são as expectativas e quais são os valores que determinada sociedade tem em seu núcleo, a serem respeitadas no processo de desenvolvimento de novas tecnologias. Yuan Stevens sugere que normas e valores que membros da sociedade desejam manter devem ser incorporadas nas avaliações tecnológicas e que deve ser analisado se os usos em potencial de determinada tecnologia vão transformar ou violar tais normas e valores.<sup>39</sup> Desta forma, a criação de mecanismos de controle e monitoramento não serão exclusivos de juristas ou gestores de políticas públicas, mas terão que ser feitos em conjunto com outros setores da sociedade, a fim de espelhar os valores e normas esperadas no desenvolvimento de IA.<sup>40</sup>

Nesse sentido são os esforços de alguns países que conceberam estratégias nacionais para o desenvolvimento de inteligência artificial e incluíram uma parte especial sobre diretrizes éticas.<sup>41</sup> Um dos maiores exemplos deste

---

Acesso: 25 de janeiro de 2019, e Google, “AI at Google: our principles”, 2018. Disponível em: <<https://blog.google/technology/ai/ai-principles/>>. Acesso: 25 de janeiro de 2019.

<sup>38</sup> Cf. Urs Gasser, *AI and the Law: Setting the Stage*, 2017.

<sup>39</sup> Yuan Stevens, *The Promises and Perils of Artificial Intelligence: Why Human Rights and the Rule of Law Matter*, 2017. Disponível em: <<https://medium.com/@ystvns/the-promises-and-perils-of-artificial-intelligence-why-human-rights-norms-and-the-rule-of-law-40c57338e806>>. Acesso: 25 de janeiro de 2019.

<sup>40</sup> Yuan Stevens, *The Promises and Perils of Artificial Intelligence: Why Human Rights and the Rule of Law Matter*, 2017. Em sentido semelhante, ver também Iyad Rahwan, *Society-in-the-Loop: Programming the Algorithmic Social Contract*, 2016. Disponível em: <<https://medium.com/mit-media-lab/society-in-the-loop-54ffd71cd802>>. Acesso: 25 de janeiro de 2019.

<sup>41</sup> Ver panorama sobre estratégias nacionais em: Tim Dutton, “An Overview of National AI Strategies”, 2018. Disponível em: <<https://medium.com/politics-ai/an-overview-of-national-ai-strategies->

tipo de iniciativa diz respeito à União Europeia, que elabora no momento diretrizes éticas por meio de um grupo de trabalho multisetorial.<sup>42</sup> Códigos de conduta e comissões de ética podem ser criados nacionalmente a fim de assegurar o desenvolvimento de inteligência artificial com respeito a direitos humanos.

Alguns pontos devem ser observados e problematizados na elaboração de diretrizes éticas para inteligência artificial, como, por exemplo:

- Como conciliar ética com inovação;
- Quais princípios éticos devem pautar a criação de sistemas de IA, levando em consideração também direitos humanos;
- Diversas formas de pautar o papel do Estado na regulação tecnológica, incluindo a possibilidade de parcerias entre atores privados e públicos para regulação na área de inteligência artificial;
- Regulação baseada em transparência e ferramentas de *accountability*;
- Preocupação com segurança;
- Criação de órgão externo regulador multisetorial;
- Pensar medidas para o sistema educacional e treinamentos relacionados a tecnologias e pensar nas consequências para empregabilidade no futuro;
- Assegurar diversidade na criação por meio de incentivos à pesquisa.

### **Estudo de caso: União Europeia**

A Comissão Europeia publicou em 2018 a comunicação sobre “Inteligência Artificial para Europa”.<sup>43</sup> Este documento estabeleceu a iniciativa europeia em inteligência artificial, que tem como objetivos aumentar a capacidade tecnológica e industrial da UE e a aceitação da IA em toda a economia, preparar para as mudanças socioeconômicas provocadas pela inteligência artificial, e garantir diretrizes éticas e jurídicas apropriadas, baseadas nos valores da UE e de acordo com a Carta de Direitos Fundamentais da UE.<sup>44</sup>

---

2a70ec6edfd>. Acesso: 26 de janeiro de 2019, e Luciano Floridi: Artificial Intelligence and the 'Good Society': the US, EU, and UK approach, *Science and engineering ethics* 24 (2), 2018, p. 505–528.

<sup>42</sup> Ver próximo tópico.

<sup>43</sup> COM(2018) 237, disponível em: < <https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe>>. Acesso: 25 de janeiro de 2019.

<sup>44</sup> COM(2018) 237, p. 4.

Para atingir o último objetivo citado, a comunicação estabelece cooperação com os principais atores interessados em inteligência artificial por meio de uma aliança europeia de inteligência artificial (“*European AI Alliance*”) para o desenvolvimento de diretrizes éticas para inteligência artificial.<sup>45</sup>

O primeiro esboço destas diretrizes éticas foi divulgado em 18 dezembro de 2018, proposto pelo Grupo de Especialistas de Alto Nível em Inteligência Artificial (*High-Level Expert Group on Artificial Intelligence - AI HLEG*).<sup>46</sup> Este grupo é constituído de 52 especialistas, que representam academia, sociedade civil e setor privado. Para o desenvolvimento do esboço, os atores envolvidos examinaram os impactos da inteligência artificial em direitos fundamentais como privacidade, dignidade, proteção do consumidor e não discriminação. Questões como o futuro do mercado de trabalho, segurança, *fairness*, inclusão social e transparência algorítmica foram analisadas.

As diretrizes partem do princípio de que uma abordagem que tem o ser humano ao centro é necessária, pois o desenvolvimento e uso de inteligência artificial deve ter como objetivo aumentar o bem-estar humano, e não ser um fim em si mesmo.<sup>47</sup> Uma “inteligência artificial confiável” (“*Trustworthy AI*”) é o que as diretrizes têm como guia. O primeiro esboço das diretrizes estabelece que uma “inteligência artificial confiável” tem dois componentes: (1) “ela deve respeitar direitos fundamentais, a regulação aplicável e princípios e valores fundamentais, garantindo um 'propósito ético'”; e (2) deve ser tecnicamente estável e confiável, pois, mesmo com boas intenções, a falta de domínio tecnológico pode causar danos não intencionais”.<sup>48</sup> Como é possível observar, não basta o comprometimento com normas e valores fundamentais da sociedade, é necessário pensar no desenvolvimento da técnica em si.

---

<sup>45</sup> COM(2018) 237, p. 4. Sobre a aliança europeia de inteligência artificial, ver: <<https://ec.europa.eu/digital-single-market/en/european-ai-alliance>>. Acesso: 25 de janeiro de 2019.

<sup>46</sup> Sobre o grupo de especialistas, ver: <<https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>>. Acesso: 25 de janeiro de 2019.

<sup>47</sup> The European Commission’s High-Level Expert Group on Artificial Intelligence, Draft Ethics Guidelines for Trustworthy AI, 2018, p. i.

<sup>48</sup> The European Commission’s High-Level Expert Group on Artificial Intelligence, Draft Ethics Guidelines for Trustworthy AI, p. i.

Importante destacar que o esboço se preocupa em operacionalizar as diretrizes éticas. Após defini-las, são listados requisitos para alcançá-las e como operacionalizar estes requisitos por meio de uma lista de avaliação (*assessment list*), a fim de se oferecer um panorama dos métodos técnicos e não técnicos que podem ser usados para a implementação das diretrizes.<sup>49</sup>

Os principais direitos fundamentais elencados que devem ser observados são: respeito pela dignidade humana; liberdade individual; respeito pela democracia, justiça e Estado de Direito; igualdade, não-discriminação e solidariedade, incluindo os direitos das pessoas pertencentes a grupos minoritários; direitos dos cidadãos em relação à administração pública, como por exemplo direito à informação e acesso a documentos públicos.<sup>50</sup> Além destes direitos, que estão positivados na Carta de Direitos Fundamentais da UE, as diretrizes elencam também princípios éticos a serem seguidos. Eles são: o princípio da beneficência (“faça o bem”), no sentido de que sistemas de AI devem melhorar o bem-estar da sociedade; o princípio da não maleficência (“não cause danos”); o princípio da autonomia (“preservar a atuação humana”); o princípio da justiça (“seja justo”); e o princípio da explicabilidade (“opere de forma transparente”).<sup>51</sup>

Desenvolvimentos de inteligência artificial que causam preocupações também são levantados, como, por exemplo, a possibilidade de identificação de indivíduos sem consentimento, classificação em massa em cidadãos com fim de avaliar a “integridade moral” e desenvolvimento de sistemas de armas automatizadas.<sup>52</sup> Tendo analisado os possíveis cenários, as diretrizes estabelecem que, além do desenvolvimento dever ser feito tendo como base direitos fundamentais e princípios éticos, especial atenção deve ser dada a situações envolvendo grupos vulneráveis como crianças ou pessoas com

---

<sup>49</sup> The European Commission’s High-Level Expert Group on Artificial Intelligence, Draft Ethics Guidelines for Trustworthy AI, p. i.

<sup>50</sup> The European Commission’s High-Level Expert Group on Artificial Intelligence, Draft Ethics Guidelines for Trustworthy AI, p. 7.

<sup>51</sup> The European Commission’s High-Level Expert Group on Artificial Intelligence, Draft Ethics Guidelines for Trustworthy AI, p. 8-10.

<sup>52</sup> The European Commission’s High-Level Expert Group on Artificial Intelligence, Draft Ethics Guidelines for Trustworthy AI, p. 11-12

deficiência, e situações nas quais existam assimetrias de poder de informação, como por exemplo entre empregador e empregados ou empresas e consumidores. É importante ter em mente que a inteligência artificial pode ter impactos negativos na sociedade e, portanto, mostra-se essencial estar vigilante em áreas que levantam preocupações.<sup>53</sup>

Os requisitos de uma “inteligência artificial confiável” devem estar presentes desde o início da fase do desenvolvimento da tecnologia. Os requisitos elencados são: mecanismos de *accountability*, governança dos dados (qualidade dos dados, evitar dados enviesados), *design* para todos (inclusão no desenvolvimento da tecnologia), governança de IA por meio de supervisão humana, não discriminação, respeito pela autonomia humana, respeito pela privacidade, confiabilidade (incluindo a precisão do sistema e sua resistência a ataques), segurança e transparência.<sup>54</sup> Para a implementação destes requisitos, métodos técnicos e não técnicos devem ser observados. Estes requisitos devem ser levados em consideração em todos os momentos relacionados à tecnologia, como, por exemplo, ao montar a equipe que vai trabalhar no sistema, no desenvolvimento do sistema em si, nos testes do sistema e na aplicação do sistema.<sup>55</sup> Em suma, a implementação destes requisitos deve ser um processo contínuo.

Do ponto de vista técnico, é necessário fornecer informações a diversos atores sobre as capacidades e limitações de sistemas de inteligência artificial, para que seja possível estabelecer expectativas realistas em relação a estes sistemas. A rastreabilidade e a possibilidade de auditar os sistemas é essencial, a fim de se lidar com os desafios de desenvolver uma tecnologia baseada em transparência e explicabilidade.<sup>56</sup> Do ponto de vista não técnico, fazer uma “inteligência artificial confiável” parte da cultura das empresas e fornecer

---

<sup>53</sup> The European Commission’s High-Level Expert Group on Artificial Intelligence, Draft Ethics Guidelines for Trustworthy AI, p. 13.

<sup>54</sup> The European Commission’s High-Level Expert Group on Artificial Intelligence, Draft Ethics Guidelines for Trustworthy AI, p. 14-18.

<sup>55</sup> The European Commission’s High-Level Expert Group on Artificial Intelligence, Draft Ethics Guidelines for Trustworthy AI, p. 18.

<sup>56</sup> The European Commission’s High-Level Expert Group on Artificial Intelligence, Draft Ethics Guidelines for Trustworthy AI, p. 20



informações às partes interessadas sobre o que é feito para se obter essa “inteligência artificial confiável” são o primeiro passo. Além disso, para assegurar diversidade, mostra-se relevante garantir a participação e inclusão das partes interessadas no *design* e desenvolvimento dos sistemas de inteligência artificial.<sup>57</sup> Mais diversidade nas equipes responsáveis pelo *design* das tecnologias pode garantir um produto mais inclusivo.<sup>58</sup> Outra medida diz respeito à educação e à conscientização dos atores envolvidos a respeito desta mentalidade ética que está sendo desenvolvida.<sup>59</sup> Por fim, deve-se ter em mente que podem existir tensões entre os objetivos: “transparência pode ser uma abertura ao mau uso; identificar e corrigir viés (*bias*) pode estar em oposição à proteção da privacidade”.<sup>60</sup> Portanto, estes dilemas, quando presentes, devem ser identificados e comunicados às partes interessadas.

Por último, o esboço fornece uma lista de avaliação com os requisitos examinados para uma “inteligência artificial confiável”.<sup>61</sup> O esboço esclarece que esta não deve ser vista como uma lista exaustiva, mas sim como um exercício que será constantemente atualizado. Os principais pontos de análise são: *accountability*; governança dos dados; *design* para todos; governando a inteligência artificial de forma autônoma; não discriminação, respeito pela privacidade; respeito pela autonomia humana, confiabilidade; segurança e transparência.<sup>62</sup>

O primeiro esboço esteve aberto para comentários das partes interessadas até o dia 1 de fevereiro de 2019. Em março, o grupo de

---

<sup>57</sup> The European Commission’s High-Level Expert Group on Artificial Intelligence, Draft Ethics Guidelines for Trustworthy AI, p. 22.

<sup>58</sup> The European Commission’s High-Level Expert Group on Artificial Intelligence, Draft Ethics Guidelines for Trustworthy AI, p. 22.

<sup>59</sup> The European Commission’s High-Level Expert Group on Artificial Intelligence, Draft Ethics Guidelines for Trustworthy AI, p. 22.

<sup>60</sup> The European Commission’s High-Level Expert Group on Artificial Intelligence, Draft Ethics Guidelines for Trustworthy AI, p. 23.

<sup>61</sup> The European Commission’s High-Level Expert Group on Artificial Intelligence, Draft Ethics Guidelines for Trustworthy AI, p. 24.

<sup>62</sup> The European Commission’s High-Level Expert Group on Artificial Intelligence, Draft Ethics Guidelines for Trustworthy AI, p. 24-27.

especialistas apresentará uma versão final à Comissão Europeia, que analisará as diretrizes e irá propor como avançar com este trabalho.<sup>63</sup> Este esboço representa uma tentativa bastante completa de mitigar riscos para os direitos humanos no âmbito da inteligência artificial. Desde o estabelecimento de um grupo de especialistas multisetorial, representando diferentes interesses da sociedade, considerando não somente os direitos fundamentais positivados no âmbito jurídico da UE, como levantando princípios éticos e examinando conjuntamente ações técnicas e não técnicas para que estes direitos sejam protegidos e os princípios éticos alcançados. Por fim, uma lista de avaliação torna operacional a possibilidade de se avaliar concretamente os requisitos traçados pelas diretrizes.

## CONSIDERAÇÕES FINAIS

Este artigo elencou e analisou os principais riscos que o desenvolvimento de inteligência artificial pode trazer para os direitos humanos. A fim de se mitigar estes riscos, algumas ações foram propostas. Entre elas, levar em consideração uma abordagem multisetorial na criação, desenvolvimento e aplicação de novas tecnologias, envolvendo diversas partes interessadas como poder público (legislativo, executivo e judiciário), setor privado, academia e sociedade civil. Também ter como base uma abordagem baseada em direitos humanos, utilizando-se de instrumentos já existentes, como a *due diligence* em direitos humanos, estabelecida pelos princípios orientadores sobre empresas e direitos humanos da ONU. Legislações internacionais e nacionais relativas à proteção de direito humanos devem continuar a ser aplicadas neste cenário de desenvolvimento de novas tecnologias. Contudo, novas formas de regulamentação devem ser pensadas. Dentre elas, a possibilidade de autorregulamentação por empresas não deve ser descartada. Além disso, a criação e o desenvolvimento de inteligência artificial devem ser pautados por

---

<sup>63</sup> Veja informações disponíveis no site da Comissão Europeia: <<https://ec.europa.eu/digital-single-market/en/news/have-your-say-european-expert-group-seeks-feedback-draft-ethics-guidelines-trustworthy>>. Acesso: 26 de janeiro de 2019.

diretrizes éticas. Para examinar com maiores detalhes o que diretrizes éticas podem oferecer neste sentido, o exemplo do desenvolvimento de diretrizes éticas no âmbito da União Europeia para uma “inteligência artificial confiável” foi analisado. O primeiro esboço feito por um grupo de especialistas multisetorial apresenta muito dos pontos problematizados aqui e pretende operacionalizar a garantia dos princípios éticos no desenvolvimento e aplicação da tecnologia por meio de uma lista de avaliação (“*assessment list*”) com diversos requisitos técnicos e não técnicos a serem observados.

## **BIBLIOGRAFIA**

Article 19, “Algorithms and automated decision-making in the context of crime prevention”. Disponível em <<https://www.article19.org/resources/algorithms-and-automated-decision-making-in-the-context-of-crime-prevention/>>. Acesso em: 25 de janeiro de 2019.

Bessi, Alessandro e Ferrara, Emilio, “Social bots distort the 2016 U.S. Presidential election online discussion”, First Monday 21, 2016.

Bickert, Monika, Fishman, Brian, Hard Questions: How We Counter Terrorism, 15 de Junho de 2017. Disponível em: <<https://newsroom.fb.com/news/2017/06/how-we-counter-terrorism/>>. Acesso: 25 de janeiro de 2019.

Bilbao Ubillos, Juan Maria, La eficacia de los derechos fundamentales frente a particulares: análisis de la jurisprudencia del Tribunal Constitucional. Madrid: Centro de Estudios Políticos y Constitucionales, 1997.

Buolamwini, Joy e Gebru, Timnit, Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification, Proceedings of Machine Learning Research 81, 2018, pp. 1-15.

Buolamwini, Joy, How I'm fighting bias in algorithms, 2017. Disponível em: <[https://www.ted.com/talks/joy\\_buolamwini\\_how\\_i\\_m\\_fighting\\_bias\\_in\\_algorithms](https://www.ted.com/talks/joy_buolamwini_how_i_m_fighting_bias_in_algorithms)>. Acesso: 26 de janeiro de 2019.

Conselho da Europa, Study on the Human Rights Dimensions of Automated Data Processing Techniques (in particular algorithms) and possible regulatory implications, 2017.

Conselho de Direitos Humanos da Organização das Nações Unidas (ONU), A/HRC/32/L.20, 27 de junho 2016.

Cullen, Daniel, Why Artificial Intelligence is Already a Human Rights Issue, Oxford Human Rights Hub, 2018. Disponível em: <<http://ohrh.law.ox.ac.uk/why-artificial-intelligence-is-already-a-human-rights-issue/>>. Acesso: 26 de janeiro de 2019.

Danieli, Oren, Hillis, Andrew e Luca, Michael, “How to Hire with Algorithms”, Harvard Business Review 2016, disponível em <[https://hbr.org/2016/10/how-to-hire-with-algorithms?referral=03758&cm\\_vc=rr\\_item\\_page.top\\_right](https://hbr.org/2016/10/how-to-hire-with-algorithms?referral=03758&cm_vc=rr_item_page.top_right)>. Acesso: 13 de março de 2018.

Dreyer, Stephan; Schulz, Wolfgang (2018): Was bringt die Datenschutz-Grundverordnung für automatisierte Entscheidungssysteme? Bertelsmann Stiftung. Gütersloh.

Dutton, Tim, “An Overview of National AI Strategies”, 2018. Disponível em: <<https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd>>. Acesso: 26 de janeiro de 2019.

European Commission, Artificial Intelligence: A European Perspective. Luxembourg: Publications Office of the European Union, 2018

European Commission, Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Artificial Intelligence for Europe, COM(2018), 237 final, Bruxelas, 2018.

European Union Agency for Fundamental Rights, #BigData: Discrimination in data-supported decision making, Luxemburgo: Publications Office, 2018.

Floridi, Luciano, Artificial Intelligence and the 'Good Society': the US, EU, and UK approach, Science and engineering ethics 24 (2), 2018, p. 505–528.

Gasser, Urs, AI and the Law: Setting the Stage, 2017. Disponível em: <<https://medium.com/berkman-klein-center/ai-and-the-law-setting-the-stage-48516fda1b11>>. Acesso: 28/05/2018.

Gimpel, Lea, “Talking about Big Data: The Challenge of Privacy & Data Protection in International Development”. Disponível em: <<http://blogs.die-gdi.de/2017/09/08/talking-about-big-data/>>. Acesso: 13 de maio de 2018.

Google, “AI at Google: our principles”, 2018. Disponível em: <<https://blog.google/technology/ai/ai-principles/>>. Acesso: 25 de janeiro de 2019.

Information Commissioner's Office (ICO), Big data, artificial intelligence, machine learning and data protection, 2017.

Mannm Gideon e O'Neil, Cathy, “Hiring Algorithms Are Not Neutral” Harvard Business Review 2016. Disponível em: <<https://hbr.org/2016/12/hiring-algorithms-are-not-neutral>> . Acesso: 13 de março de 2018.

McGregor, Lorna e Ng, Vivian, “Google’s new principles on AI need to be better at protecting human rights”, The Human Rights, Big Data and Technology Project, 2018. Disponível em: <<https://hrbdt.ac.uk/googles-new-principles-on-ai-need-to-be-better-at-protecting-human-rights/>>. Acesso: 25 de janeiro de 2019.

McGregor, Lorna, Why We Need to Stop Talking About ‘Killer Robots’ and Address the AI Backlash, Blog of the European Journal of International Law, EJIL: Talk!, 9 de julho de 2018. Disponível em: <<https://www.ejiltalk.org/why-we-need-to-stop-talking-about-killer-robots-and-address-the-ai-backlash/>>. Acesso em: 24 de janeiro de 2019.

McKinsey Global Institute, Skill Shift Automation and the Future of the Workforce, Discussion Paper, 2018.

OHCHR, Corporate human rights due diligence – identifying and leveraging emerging practice, disponível em: <<https://www.ohchr.org/EN/Issues/Business/Pages/CorporateHRDueDiligence.aspx>>. Acesso: 25 de janeiro de 2019.

Princípios orientadores sobre empresas e direitos humanos da ONU, 2011. Disponível em:

<[https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR\\_EN.pdf](https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf)>. Acesso: 25 de janeiro de 2019.

Rahwan, Iyad, Society-in-the-Loop: Programming the Algorithmic Social Contract, 2016. Disponível em: <<https://medium.com/mit-media-lab/society-in-the-loop-54ffd71cd802>>. Acesso: 25 de janeiro de 2019.

Stevens, Yuan, The Promises and Perils of Artificial Intelligence: Why Human Rights and the Rule of Law Matter, 2017. Disponível em: <<https://medium.com/@ystvns/the-promises-and-perils-of-artificial-intelligence-why-human-rights-norms-and-the-rule-of-law-40c57338e806>>. Acesso: 25 de janeiro de 2019.

The European Commission's High-Level Expert Group on Artificial Intelligence, Draft Ethics Guidelines for Trustworthy AI, 2018.

The Guardian, "Elon Musk: regulate AI to combat 'existential threat' before it's too late", 17 julho de 2017. Disponível em: <<https://www.theguardian.com/technology/2017/jul/17/elon-musk-regulation-ai-combat-existential-threat-tesla-spacex-ceo>>. Acesso: 24 de janeiro de 2019.

The New York Times, "Sent to Prison by a Software Program's Secret Algorithms", 1 de maio de 2017. Disponível em <<https://www.nytimes.com/2017/05/01/us/politics/sent-to-prison-by-a-software-programs-secret-algorithms.html>>. (Acesso: 13 de março de 2018).

The Norwegian Data Protection Authority, Artificial intelligence and privacy, 2018.

United Nations Human Rights Regional Office for Europe, Making a Difference: An Introduction to Human Rights, 2018, p. 14. Disponível em: <[https://europe.ohchr.org/Documents/Publications/MakeADifference\\_EN.pdf](https://europe.ohchr.org/Documents/Publications/MakeADifference_EN.pdf)>. Acesso: 25 de janeiro de 2018.