

# DA PERSONALIDADE ELETRÔNICA À CLASSIFICAÇÃO DE RISCOS NA INTELIGÊNCIA ARTIFICIAL (IA)

Sergio Marcos Carvalho  
de Ávila Negri,  
Universidade Federal de  
Juiz de Fora

<https://orcid.org/0000-0003-2156-3518>

Giovana F. Peluso Lopes,  
Universidade de Bologna  
<https://orcid.org/0000-0003-4798-2542>

Data de submissão:

30/06/2021

Data de aceite:

22/11/2021

## From electronic personality to risk classification in artificial intelligence (ai)

### RESUMO

O presente artigo procura investigar a temática da criação de uma personalidade jurídica (eletrônica) para agentes autônomos de Inteligência Artificial, abordando as limitações desse tipo de iniciativa através da análise do instituto da pessoa jurídica. Procura-se, principalmente, confrontar o modelo da personalidade eletrônica com a nova proposta europeia de regulamentação da Inteligência Artificial, pautada na classificação dos riscos dos diversos sistemas. Para tanto, serão apresentadas as principais características desse tipo de tecnologia que justificariam a proposta de personificação, bem como os atributos necessários para o seu enquadramento como pessoas jurídicas.

**Palavras-chave:** Personalidade eletrônica. Pessoa jurídica. Inteligência Artificial.

### Abstract

This article seeks to investigate the theme of creating a legal personality (electronic) for autonomous artificial intelligence agents, addressing the limitations of this type of initiative through the analysis of the legal entity institute. It seeks to confront the electronic personality with the new European proposal for the regulation of Artificial Intelligence. Therefore, the main characteristics of this type of technology that would justify the proposal of personification will be analyzed, as well as the necessary attributes for its framing in the proposal.

**Keywords:** Eletronic personality. Legal entity. Artificial Intelligence (AI).



## INTRODUÇÃO

Em 16 de fevereiro de 2017, o Parlamento Europeu editou a Resolução 2015/2103-INL com recomendações à Comissão de Direito Civil sobre Robótica (UNIÃO EUROPEIA, 2017). O documento abordava variados temas relativos ao emprego crescente de tecnologias de Inteligência Artificial (IA) – como o seu impacto sobre a educação e o mercado de trabalho, o uso de veículos autônomos, propriedade intelectual, pesquisa e inovação. Especificamente no que tange à responsabilidade civil, o documento destacou que, na medida em que a autonomia de sistemas de Inteligência Artificial aumenta, não mais será possível considerá-los como meras ferramentas nas mãos de outros atores (como seu fabricante, operador, proprietário ou usuário).

Desde então, o tema da personificação de sistemas de IA autônomos tem ganhado destaque não apenas no meio acadêmico ou jurídico, mas tem também figurado em notícias e vem sendo debatido pela população em geral. A popularidade do tema não é de espantar, especialmente quando se tem em vista que a ideia de que seres humanos eventualmente seriam capazes de criar máquinas aptas a tomar decisões por si próprias e obter certa autonomia em relação a seus criadores sempre esteve presente no imaginário coletivo, servindo como pano de fundo para contos, mitos, obras literárias, teatrais e cinematográficas.<sup>1</sup>

Em 21 de abril de 2021, porém, a Comissão Europeia apresentou a Proposta de Regulamento sobre a Inteligência Artificial (“Artificial Intelligence Act”)<sup>2</sup>, que busca estabelecer um quadro jurídico uniforme para o desenvolvimento, a comercialização e a utilização da Inteligência Artificial no âmbito da União Europeia. A atual proposta se afasta da criação de uma personalidade jurídica eletrônica. O texto está amparado em uma abordagem baseada no risco, modulando o conteúdo das normas de acordo com a intensidade dos riscos criados pelos sistemas de IA.

<sup>1</sup> Cita-se, por exemplo, o mito judaico do golem; o clássico romance “Frankenstein ou o Moderno Prometeu” de Mary Shelly; as obras distópicas de Isaac Asimov como “Eu, Robô”; o popular desenho animado The Jetsons, produzido pela empresa Hanna-Barbera; e a personagem HAL 9000, principal antagonista de 2001: Uma Odisseia no Espaço, de Arthur C. Clarke.

<sup>2</sup> Disponível em: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206>.

**No lugar de se reconhecerem as particularidades das diferentes áreas de atuação de artefatos com Inteligência Artificial, unificam-se essas diferentes relações em modelo jurídico único, pautado exclusivamente na figura de um sujeito abstrato ou em outras generalizações. Esse é um problema frequente quando o Direito tenta se aproximar de novas tecnologias.**

Nesse contexto, o presente trabalho, a partir de uma abordagem exploratória, tem como objetivo investigar a temática da criação de uma personalidade jurídica para agentes autônomos de Inteligência Artificial, confrontando a proposta inicial, associada à personalidade eletrônica, com o novo modelo pautado na classificação dos riscos. Parte-se da hipótese de que a criação de uma personalidade eletrônica pode repetir os mesmos problemas presentes no unitarismo da pessoa jurídica. No lugar de se reconhecerem as particularidades das diferentes áreas de atuação de artefatos com Inteligência Artificial, unificam-se essas diferentes relações em modelo jurídico único, pautado exclusivamente na figura de um sujeito abstrato ou em outras generalizações. Esse é um problema frequente quando o Direito tenta se aproximar de novas tecnologias.

A ausência da referência à personalidade eletrônica na nova proposta não significa, contudo, que o debate sobre a subjetividade jurídica tenha desaparecido por completo. Mesmo sem a referência à personalidade, observa-se, ainda, em outras iniciativas a utilização de modelos jurídicos abstratos e gerais, indiferentes aos variados usos dos sistemas de IA. É o caso do Projeto de Lei 21/2020, que busca criar um marco legal para a Inteligência Artificial no Brasil, estabelecendo fundamentos, princípios e diretrizes para o seu desenvolvimento e aplicação, tanto pelo Poder Público quanto pela iniciativa privada<sup>3</sup>.

Embora a iniciativa constitua importante passo para a regulamentação do tema no ordenamento jurídico brasileiro, ela comete o mesmo equívoco presente na Resolução original do Parlamento Europeu ao estabelecer o mesmo tratamento unitário aos mais diversos tipos de tecnologias de IA, sem levar em consideração variáveis importantes, como o contexto de utilização e o nível de autonomia presente. Conforme buscar-se-á demonstrar ao longo deste artigo, a concessão de um tratamento jurídico unitário a entidades substancialmente diversas (como a equiparação da pessoa natural à pessoa jurídica e, finalmente, à pessoa eletrônica; ou, ainda, a equiparação de IAs com níveis de risco distintos para fins de regulação e responsabilização), é problemática.

Para o desenvolvimento do trabalho, serão inicialmente apresentadas as principais características desse tipo de tecnologia que

<sup>3</sup> Disponível em: <https://www.camara.leg.br/propostas-legislativas/2236340>.

justificariam a proposta de personificação, bem como os atributos necessários para o seu enquadramento como pessoas jurídicas. Em seguida, será feita a análise da personalidade jurídica (eletrônica) tendo como base as categorias de “sujeitos de direito” artificiais, apontando-se os riscos associados a tal retórica. Para se evitar a descontextualização da IA, o trabalho procura dialogar com as novas diretrizes da Robótica apresentadas por Frank Pasquale<sup>4</sup>, as quais, ao contrário das propostas por Asimov, são destinadas aos desenvolvedores de sistemas de IA, estes sim responsáveis pela criação de sistemas éticos e de confiança, e não aos robôs fictícios das obras literárias. Por fim, pretende-se analisar como a nova proposta europeia de regulamentação, ao se afastar do unitarismo da personalidade eletrônica, e ao adotar uma abordagem baseada no risco e na autorregulação, indica um caminho interessante para se repensar a relação entre Direito e Tecnologia, e para pautar futuras opções legislativas no contexto brasileiro. Metodologicamente, trata-se de pesquisa essencialmente bibliográfica, de caráter exploratório, valendo-se de fontes documentais diretas e indiretas para uma revisão compreensiva do tema.

## 1 INTELIGÊNCIA ARTIFICIAL: AUTONOMIA, IMPREVISIBILIDADE E ININTELIGIBILIDADE

O objetivo central da Inteligência Artificial, como campo de pesquisa, é compreender os princípios que tornam possível o comportamento inteligente em sistemas, através da análise de agentes, naturais e artificiais, que assim se comportam, e da pesquisa e desenvolvimento de sistemas computacionais capazes de executar tais tarefas comumente vistas como exigindo inteligência.<sup>5</sup>

Embora exista uma miríade de definições propostas para a Inteligência Artificial, boa parte delas encontra-se alinhada a essa ideia central, como a proposição de Kurzweil de que a IA consistiria na arte de criar máquinas que executem funções que demandam inteligência quando realizadas por pessoas,<sup>6</sup> ou no desenvolvimento

<sup>4</sup> PASQUALE, Frank. *New laws of robotics: defending human expertise in the age of AI*. Cambridge, Massachusetts: The Belknap Press Of Harvard University Press, 2020.

<sup>5</sup> POOLE, David; MACKWORTH, Alan. *Artificial Intelligence: Foundations of Computational Agents*. Cambridge: Cambridge University Press, 2010.

<sup>6</sup> KURZWEIL, Ray. *The Singularity is Near: When humans transcend biology*. Nova York: Penguin Books, 2005.

**Entretanto, os principais avanços em IA nos últimos anos foram proporcionados por uma abordagem diversa, centrada na análise de dados – o aprendizado de máquina (ou machine learning). Este pode ser definido, em linhas gerais, como a ciência de fazer com que computadores desempenhem determinada tarefa sem serem explicitamente programados.**

de soluções automatizadas por sistemas ou agentes para problemas que requereriam a intervenção da inteligência se executados por humanos.<sup>7</sup>

É importante destacar que Inteligência Artificial é o gênero que abarca diversas metodologias, como procedimentos de programação clássica, concentrados na manipulação de símbolos de alto nível (tais como os sistemas especialistas) – comumente chamados, atualmente, de *good old fashioned artificial intelligence*. Este tipo de abordagem se faz presente no subcampo da IA denominado planejamento automático, cujo foco é o desenvolvimento de técnicas para a resolução de problemas que envolvam a formulação de uma série de etapas para atingir um objetivo desejado.<sup>8</sup>

Entretanto, os principais avanços em IA nos últimos anos foram proporcionados por uma abordagem diversa, centrada na análise de dados – o aprendizado de máquina (ou *machine learning*). Este pode ser definido, em linhas gerais, como a ciência de fazer com que computadores desempenhem determinada tarefa sem serem explicitamente programados.<sup>9</sup> Em linhas gerais, tem-se que programas de computador capazes de aprender são, na verdade, capazes de extrair padrões estatísticos e correlações ocultas em seus (volumosos) dados de entrada.

É importante notar que, embora os princípios básicos por trás do aprendizado de máquina não sejam necessariamente recentes, algumas alterações ocorridas nos últimos anos fizeram com que o campo de pesquisa experimentasse uma espécie de renascimento.

Dentre essas alterações, destaca-se a melhoria dos métodos utilizados – com a introdução, por exemplo, do algoritmo de retropropagação, que criou a possibilidade do treinamento multicamadas<sup>10</sup>;

<sup>7</sup> NEGNEVITSKY, Michael. *Artificial Intelligence: A guide to intelligent systems*, 2 Ed. Harlow: Addison-Wesley, 2004.

<sup>8</sup> Exemplos incluem: fornecer instruções de direção, empacotar caixas de tamanhos variados em um caminhão, provar teoremas matemáticos ou analisar contratos e regulamentos legais.

<sup>9</sup> SAMUEL, Arthur. L. Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, v. 3, n. 3, 1959, p. 210-229.

<sup>10</sup> Redes multicamadas possuem uma ou mais camadas intermediárias (ou ocultas) de neurônios (nós) entre aquelas de entrada e saída de dados, e podem aprender uma gama muito mais variada de funções do que suas antecessoras mais simples. De fato, a área de pesquisa denominada aprendizado profundo (*deep learning*), por vezes utilizada como sinônimo de aprendizado de máquina, refere-se ao uso de redes neurais artificiais que possuem um elevado número de camadas ocultas.

da disponibilidade crescente de enormes quantidades de dados e do aumento do poder computacional. Mais recentemente, tem-se ainda a transformação de certos espaços em ambientes mais receptivos à tecnologia de informação, como ocorre com as chamadas cidades inteligentes.<sup>11</sup>

Por fim, cumpre ressaltar que o conceito de Inteligência Artificial se aplica a sistemas que apresentam um comportamento inteligente, analisando o seu ambiente e tomando medidas, com um determinado nível de autonomia, para atingir objetivos específicos. Esses sistemas baseados em IA podem ser puramente confinados ao software, atuando no mundo virtual, ou podem ser integrados em dispositivos físicos – hipótese tratada pelo subcampo da robótica.

De fato, a Resolução do Parlamento Europeu de 16 de fevereiro de 2017 aborda, principalmente, a influência de artefatos robóticos inteligentes na sociedade. Para ser categorizado como tal, a Resolução estabeleceu os seguintes critérios para robôs: a) existência de sensores capazes de permitir a troca de dados com o ambiente; b) capacidade de aprendizado com a experiência e interação com o meio; c) existência de um suporte material; d) capacidade de adaptação; e e) ausência de vida na acepção biológica.<sup>12</sup>

Dentre essas características, apenas a existência de um suporte material diferencia artefatos robóticos de outros agentes de IA. Para Russell e Norvig, pode-se definir um agente como qualquer coisa que possa ser vista como percebendo seu ambiente por meio de sensores e atuando nesse ambiente através de atuadores.<sup>13</sup> Note-se que, segundo a definição por eles proposta, é possível abarcar dentro do conceito uma ampla variedade de entes – humanos, robóticos ou de software.

Assim, o ambiente de um agente pode variar desde o mundo real (no caso de um humano ou robô), até ambientes simulados no ciberespaço ou ambientes informacionais consistentes em bases de

<sup>11</sup> MITTELSTADT, Brent; FLORIDI, Luciano; ALLO, Patrick; TADDEO, Mariarosaria; WACHTER, Sandra. The ethics of algorithms: Mapping the debate. *Big Data & Society*, v. 3(2), 2016, p. 1-21.

<sup>12</sup> UNIÃO EUROPEIA. 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)*. Bruxelas: Comissão Europeia, 2018.

<sup>13</sup> RUSSELL, Stuart J.; NORVIG, Peter. *Artificial Intelligence: A Modern Approach*. 3. ed. New Jersey: Prentice-Hall, 2010.

dados, aplicações, serviços e redes (no caso de um agente de software). Cada um desses ambientes pode, por óbvio, incluir outros agentes, bem como ser modificado à medida em que o agente em questão prossegue com suas tarefas. Já a adaptabilidade de um agente pode ser avaliada pela sua capacidade de funcionar em ambientes especificados imprecisa ou incompletamente, de maneira dinâmica e sensível ao contexto em questão.<sup>14</sup>

### 1.1 Autonomia e imprevisibilidade

Com base no que foi exposto, pode-se dizer que um agente possui maior ou menor autonomia a depender de sua capacidade de planejamento e aprendizado. Agentes capazes de planejamento podem construir sequências de ações para realizar tarefas, utilizando-se de uma ampla variedade de técnicas algorítmicas para lidar com a incerteza em seus ambientes.

Possuem, portanto, uma natureza direcionada ao alcance de objetivos: um resultado final pode ser especificado e, dado o conhecimento do agente sobre as ações necessárias para alcançá-lo, ele poderá realizá-las considerando os recursos existentes e as características ambientais, selecionando dentre as diversas opções disponíveis de forma a maximizar sua performance e obter o melhor resultado possível.<sup>15</sup>

A capacidade de aprendizado consiste na aptidão para modificar a si mesmo, seu comportamento e suas respostas diante de suas experiências com o meio ambiente. Desse modo, conforme o agente dependa do conhecimento prévio de seu programador – no caso de um agente artificial - e não se baseie em suas próprias percepções, é possível afirmar que lhe falta autonomia.<sup>16</sup>

Um agente artificial autônomo, por sua vez, será considerado como tal se agir independentemente de instrução ou supervisão humana direta, com base nas informações que ele próprio adquire

<sup>14</sup> CHOPRA, Samir; LAWRENCE, White. *A Legal Theory for Artificial Autonomous Agents*. Ann Arbor: University of Michigan Press, 2011.

<sup>15</sup> CHOPRA, Samir; LAWRENCE, White. *A Legal Theory for Artificial Autonomous Agents*. Ann Arbor: University of Michigan Press, 2011.

<sup>16</sup> RUSSELL, Stuart J.; NORVIG, Peter. *Artificial Intelligence: A Modern Approach*. 3. ed. New Jersey: Prentice-Hall, 2010.

**A ideia de comportamento emergente está associada a uma perspectiva holística, na qual o comportamento de um agente de IA não se confunde com a simples soma das suas partes, criando, em certas situações, a sensação de que ele realizou um comportamento inesperado, não programado.**

e analisa,<sup>17</sup> sendo capaz de direcionar suas próprias atividades em um ambiente imprevisível ou em constante modificação.

É importante ressaltar que a autonomia é aqui referida em um sentido fraco, como a capacidade de realizar tarefas sem a direção/supervisão humana direta e a capacidade de aprendizado. Diferente, assim, do sentido forte associado ao termo, como a existência de consciência ou autoconsciência e a consequente identificação de um agente moral.<sup>18</sup>

No que diz respeito à autonomia, é possível apontar uma das primeiras falhas no debate acerca da personificação desses agentes, qual seja, a falta de determinação do termo. Confunde-se autonomia com a imprevisibilidade do resultado, ignorando o fato de que máquinas operadas a partir de controle humano direto podem, igualmente, ocasionar resultados imprevisíveis.<sup>19</sup>

Segundo Richards e Smart, a Inteligência Artificial é analisada a partir de uma sensação de agência, que não se confunde com o seu sentido forte. Desse modo, trata-se de um sistema construído que exhibe, ainda que apenas aparentemente, uma agência física e mental, não estando vivo no sentido biológico, sendo um ente fabricado, que se move pelo ambiente (físico ou virtual), e apenas aparenta tomar decisões racionais (autonomia fraca) sobre o que fazer.<sup>20</sup>

Em vista da ambiguidade advinda do termo, Calo opta por substituir “autonomia” por “comportamento emergente”, tipicamente encontrado em sistemas adaptativos complexos nos quais há um comportamento global resultante da interação individual.<sup>21</sup>

A ideia de comportamento emergente está associada a uma perspectiva holística, na qual o comportamento de um agente de IA não se confunde com a simples soma das suas partes, criando, em certas situações, a sensação de que ele realizou um comportamento

<sup>17</sup> VLADECK, David C. Machines Without Principles: Liability Rules and Artificial Intelligence. *Washington Law Review*, v. 89, n. 01, 2014, p. 117-150.

<sup>18</sup> BERTOLINI, Andrea. Robots as products: the case for a realistic analysis of robotic applications and liability rules. *Law, Innovation and Technology*, v. 5, n. 2, 2013, p. 214-247.

<sup>19</sup> NEGRI, Sérgio. Robôs como pessoas: a personalidade eletrônica na Robótica e na inteligência artificial. *Pensar*, v. 25, n. 3, 2020, p. 1-14.

<sup>20</sup> RICHARDS, Neil M.; SMART, William D. How should the law think about robots? In: CALO, Ryan; FROOMKIN, A. Michael; KERR, Ian (orgs.) *Robot Law*. Cheltenham: Edward Elgar Publishing, 2016.

<sup>21</sup> CALO, Ryan. Robotics and the lessons of cyberlaw. *California Law Review*, v. 103, n. 3, 2015, p. 513-563.

inesperado, não programado. Esse tipo de surpresa pode depender da expectativa subjetiva do observador, não obstante, ainda que se adote a perspectiva do programador, é impossível estabelecer de antemão todos os comportamentos que emergem da interação que ocorre apenas em determinado tempo e espaço da execução.<sup>22</sup>

Note-se que, ainda que as ações desses agentes tenham sido inesperadas – certamente para observadores externos e talvez até para os seus programadores –, a sua capacidade de produzir ações imprevisíveis muito frequentemente resulta do planejamento dos seus desenvolvedores. Afinal, existe cada vez mais um considerável incentivo econômico para criar sistemas de IA capazes de gerar soluções inovadoras. Nesse sentido, Calo relembra que o comportamento emergente é um objetivo claramente declarado da robótica e da Inteligência Artificial.<sup>23</sup>

Poole e MackWorth relembra que dois agentes com o mesmo conhecimento prévio, história, habilidades e objetivos deverão, necessariamente, realizar a mesma ação; conseqüentemente, a alteração de qualquer uma dessas variáveis poderá resultar em ações diferentes. No caso de sistemas de IA, trata-se de agentes determinísticos, cujos *inputs* e demais ajustes de variáveis produzirão *outputs* previsíveis – mesmo em um ambiente controlado.<sup>24</sup>

O que ocorre é que, na medida em que a influência do programador sobre a máquina diminui, a influência do ambiente operacional aumenta. Essencialmente, o programador transfere parte de seu controle sobre o produto para o ambiente, sobretudo em se tratando de máquinas que continuam aprendendo e se adaptando em seu ambiente operacional final. Como nessa situação elas têm que interagir com um número potencialmente grande de pessoas e situações, via de regra não será possível prever a influência do ambiente operacional.

Seres humanos possuem uma propensão natural de antropomorfizar determinados entes cujo comportamento se assemelhe ao seu,

<sup>22</sup> NEGRI, Sérgio. Robôs como pessoas: a personalidade eletrônica na Robótica e na inteligência artificial. *Pensar*, v. 25, n. 3, 2020, p. 1-14.

<sup>23</sup> CALO, Ryan; FROOMKIN, A. Michael; KERR, Ian (orgs.) *Robot Law*. Cheltenham: Edward Elgar Publishing, 2016.

<sup>24</sup> POOLE, David; MACKWORTH, Alan. *Artificial Intelligence: Foundations of Computational Agents*. Cambridge: Cambridge University Press, 2010.

**Outro fator que contribui de maneira decisiva para a antropomorfização de agentes de IA é a dificuldade de se compreender como determinadas ações ou decisões foram alcançadas. Viu-se que, por trás de grande parte das aplicações da Inteligência Artificial, existe um número incomensurável de dados que viabilizam o seu funcionamento.**

e cujo funcionamento lhes seja incompreensível.<sup>25</sup> Essa projeção de atributos humanos é perigosa, porém, quando o objetivo é pensar em formas para regulamentar a Inteligência Artificial. Isso porque tais agentes são, e por muitos anos permanecerão sendo, apenas ferramentas, ainda que com elevado grau de sofisticação:

À medida que a autonomia do sistema aumenta, fica cada vez mais difícil definir a conexão entre os comandos de entrada (inputs) e o comportamento do robô (outputs), mas ele existe e é determinístico. O mesmo conjunto de entradas gerará o mesmo conjunto de saídas sempre. O problema, no entanto, é que o robô nunca verá exatamente a mesma entrada duas vezes. Duas imagens da câmera nunca são iguais, devido a mudanças sutis de iluminação e erros de medição na própria câmera. Os humanos podem não ser capazes de ver as diferenças, mas o software do robô vê. O problema é que esse comportamento diferente em situações aparentemente semelhantes pode ser interpretado como ‘livre-arbítrio’ ou agência por parte do robô.<sup>26</sup>

Assim, a metáfora antropomórfica, “robôs como pessoas”, esconde aspectos funcionais da Inteligência Artificial, fazendo com que essa retórica, que mimetiza qualidades e atributos humanos, possa comprometer o enfrentamento dos complexos desafios ético-regulatórios colocados pelas tecnologias emergentes.<sup>27</sup>

## 1.2 Ininteligibilidade

Outro fator que contribui de maneira decisiva para a antropomorfização de agentes de IA é a dificuldade de se compreender como determinadas ações ou decisões foram alcançadas. Viu-se que, por trás de grande parte das aplicações da Inteligência Artificial, existe um número incomensurável de dados que viabilizam o seu funcionamento.

Nos últimos anos, os notáveis progressos em Inteligência Artificial ocorreram na área de *deep learning*, através da utilização de redes neurais. Estas, contudo, não lidam com a incerteza de maneira

<sup>25</sup> DARLING, Kate. Extending Legal Rights to Social Robots. *We Robot Conference*, Miami, 2012, p. 1- 18.

<sup>26</sup> RICHARDS, Neil M.; SMART, William D. How should the law think about robots? In: CALO, Ryan; FROOMKIN, A. Michael; KERR, Ian (orgs.) *Robot Law*. Cheltenham: Edward Elgar Publishing, 2016, p. 18.

<sup>27</sup> NEGRI, Sérgio. Robôs como pessoas: a personalidade eletrônica na Robótica e na inteligência artificial. *Pensar*, v. 25, n. 3, 2020, p. 1-14.

rigorosa ou transparente, e menos ainda incorporam qualquer representação explícita do ambiente em que operam. Em vez disso, a arquitetura da rede fica livre para evoluir por conta própria e, ao final do treinamento, seu programador não tem ideia de quais cálculos estão sendo executados ou porque eles funcionam.<sup>28</sup>

Nesse mesmo sentido, Edwards e Veale ressaltam que esses algoritmos são gerados tendo como prioridade o desempenho preditivo, ao invés da interpretabilidade de suas ações.<sup>29</sup> A interpretabilidade pode ser definida como a medida do grau de um observador humano que tem de compreender as razões por trás de uma predição feita por um modelo de IA.<sup>30</sup> Todavia, otimizar um modelo de explicação para a interpretabilidade humana significa necessariamente diluir o desempenho preditivo para capturar apenas as principais lógicas do sistema. Como sistemas com mais variáveis geralmente desempenham melhor do que sistemas mais simples, acaba por existir um *trade off* entre performance e explicabilidade.

Desse modo, verifica-se que as ações e decisões tomadas por certos algoritmos de IA são, na maioria das vezes, dificilmente traduzidas em explicações que façam sentido para usuários humanos, não apenas o público em geral, mas incluindo-se aqui os próprios programadores. A ininteligibilidade de *outputs*, somada a certo grau de imprevisibilidade quanto aos resultados, resulta na presunção (equivocada) de que tais entes possuiriam autonomia, no sentido forte associado ao termo.

Em vista disso, e da indefinição ontológica e jurídica acerca dessa tecnologia emergente, o Direito se vê obrigado a recorrer a velhas figuras – metáforas já conhecidas – que auxiliam na aproximação, com certa familiaridade, daquilo que é novo e desconhecido.<sup>31</sup>

<sup>28</sup> PEARL, Judea; MACKENZIE, Dana. *The Book of Why: The new science of cause and effect*. New York: Basic Books, 1. Ed., 2018.

<sup>29</sup> EDWARDS, Lillian; VEALE, Michael. Slave to the Algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for. *Duke Law & Technology Review*, v. 16, n. 01, 2016, p. 18-84.

<sup>30</sup> É possível apontar três níveis distintos de interpretabilidade: alta, que inclui algoritmos tradicionais de regressão, árvores de decisão e classificadores baseados em regras (métodos de fácil compreensão); média, no caso de algoritmos um pouco mais avançados, tais como modelos gráficos; e baixa interpretabilidade, compreendendo as técnicas avançadas de machine learning, tais como SVM (Support Vector Machines), Ensemble Methods e Redes Neurais Profundas. Esses métodos podem, no máximo, esboçar ideias de importâncias das variáveis do problema como medida de interpretabilidade (SILVA, 2019).

<sup>31</sup> NEGRÍ, Sérgio. Robôs como pessoas: a personalidade eletrônica na Robótica e na inteligência artificial. *Pensar*, v. 25, n. 3, 2020, p. 1-14.

**A partir do momento em que se admite que os robôs atuais podem realizar comportamentos não previstos, restaria ao ordenamento jurídico, portanto, reconhecer-lhes a personalidade jurídica.**

Dentre tais metáforas, destaca-se a iniciativa de conferir personalidade à IA. Afinal, se para o Direito a personalidade jurídica já se encontra desvinculada ao substrato humano, ela poderia igualmente ser estendida à Inteligência Artificial, a despeito da inexistência de qualquer característica humana nesses artefatos.

## 2. PERSONALIDADE JURÍDICA (ELETRÔNICA)

No debate acerca da personificação da Inteligência Artificial, é comumente constatada a afirmação de que as normas legais existentes (a respeito, por exemplo, da responsabilidade civil em caso de danos causados por esses agentes) seriam incapazes de retratar e, conseqüentemente, disciplinar agentes artificiais autônomos. A partir do momento em que se admite que os robôs atuais podem realizar comportamentos não previstos, restaria ao ordenamento jurídico, portanto, reconhecer-lhes a personalidade jurídica.<sup>32</sup>

Esse tipo de raciocínio possui diversas falhas, sendo a primeira delas, conforme previamente apontado, a falta de determinação do significado de autonomia. Para além disso, a pessoa jurídica é apresentada como se não existissem problemas no processo de atribuição da personalidade jurídica às sociedades – razão pela qual a compreensão do instituto é necessária.

A partir da obra de Galgano, verifica-se que o recurso à personificação surge, na Modernidade, como um importante instrumento para a articulação patrimonial,<sup>33</sup> fazendo com que a separação de um patrimônio em virtude de uma determinada finalidade possa ser mais bem compreendida a partir da criação de um novo sujeito.

No discurso acerca da personalidade jurídica (eletrônica), é possível notar certa confusão entre atribuição de personalidade e separação patrimonial. A criação de um fundo próprio para eventuais danos ocasionados – garantindo, assim, a devida compensação para os indivíduos lesados, independe da criação de um novo sujeito, tendo em vista que a pessoa jurídica, ainda que associada à autonomia patrimonial, não detém o monopólio da destinação patrimonial.<sup>34</sup>

<sup>32</sup> NEGRI, Sérgio. Robôs como pessoas: a personalidade eletrônica na Robótica e na inteligência artificial. *Pensar*, v. 25, n. 3, 2020, p. 1-14.

<sup>33</sup> GALGANO, Francesco. *Lex Mercatoria*. Bologna: Il Mulino, 2001.

<sup>34</sup> NEGRI, Sérgio. Robôs como pessoas: a personalidade eletrônica na Robótica e na inteligência artificial. *Pensar*, v. 25, n. 3, 2020, p. 1-14.

Para além da alocação patrimonial, a personificação possibilita, ainda, a constituição de um sistema de imputação direta dos atos praticados pelos órgãos da pessoa jurídica, através da mediação orgânica, que possibilita conceber os órgãos do ente coletivo como parte do mesmo. Possibilita-se, desse modo, o reconhecimento, por parte do ordenamento jurídico, do ente personificado como verdadeiro autor dos atos realizados por seus agentes. O reconhecimento de uma nova subjetividade permite, ademais, simplificar situações jurídicas complexas com a constituição de um centro unitário de imputação, desempenhando a personalidade jurídica uma função heurística, isto é, um atalho mental. Isso porque não há nada que impeça, de antemão, que todos os integrantes de uma sociedade ou associação, por exemplo, relacionem-se conjuntamente com terceiros. Entretanto, a interposição de uma nova subjetividade permite que esse mesmo processo passe a ser descrito de forma unitária, sem a necessidade de menção, a cada ato praticado, de todos os sócios que se mostram retratados pelo termo pessoa jurídica.<sup>35</sup>

Nesse mesmo sentido, Bryson et al. afirmam que a criação de um discurso fictício segundo o qual determinados entes coletivos são pessoas constitui um atalho útil para lhes atribuir direitos e obrigações legais a princípio possuídos apenas por pessoas naturais dentro de um ordenamento jurídico – como o direito de vincular outras subjetividades através de um contrato, ou o dever de satisfazer as obrigações dele decorrentes.<sup>36</sup>

Cumprindo ainda ressaltar que conceder personalidade jurídica a entidades específicas revela um posicionamento acerca da avaliação social da entidade em questão. Para o jurista Stefano Rodotà, o Direito sempre contribuiu para a criação de antropologias, no sentido de que ele é responsável pela construção de figuras sociais.<sup>37</sup> Cada grande operação jurídica delineou um modelo de pessoa, que nunca foi o mero registro de uma natureza “humana”, mas um jogo de plenos e vazios, de seleção daquilo que podia encontrar acolhimento no espaço do Direito e aquilo que deveria ficar de fora.

<sup>35</sup> NEGRI, Sergio Marcos Carvalho de Ávila. As razões da pessoa jurídica e a expropriação da subjetividade. *Civilistica*, v. 5, n. 2, 2016, p. 1-18.

<sup>36</sup> BRYSON, Joanna J.; DIAMANTIS, Mihailis E.; GRANT, Thomas D. Of, for, and by the people: the legal lacuna of synthetic persons. *Artificial Intelligence Law*, n. 25, 2017, p. 237- 291.

<sup>37</sup> RODOTÀ, Stefano. A antropologia do homo dignus. Trad. Maria Celina Bodin de Moraes. *Civilistica*, v. 6, n. 2, 2017, p. 1-17. Disponível em: <http://civilistica.com/aantropologia-do-homo-dignus/>. Acesso em: 10 jun. 2021.

## No debate sobre a personalidade eletrônica, o processo de atribuição de personalidade jurídica às sociedades é apresentado como um modelo que justificaria o reconhecimento da subjetividade jurídica para robôs com Inteligência Artificial.

Se compreendida como um jogo de soma zero, em que as decisões sobre a personalidade influenciam também outros interesses além dos da entidade em questão, a personificação de entidades não humanas pode representar prejuízos à personalidade de pessoas naturais. Perder de vista as razões específicas que informam as pessoas jurídicas, equiparando-as indistintamente às pessoas naturais, pode conduzir ao processo denominado por Rodotà de “expropriação da subjetividade”: a usurpação, no plano concreto, de direitos inerentes ao ser humano, sob o pretexto de proteção do sujeito abstrato.<sup>38</sup>

Tal fenômeno é identificado no discurso dos direitos fundamentais e da personalidade da pessoa jurídica, no qual se constata uma inversão: ao invés de se analisar a pertinência do conteúdo do direito em questão, por meio de uma valoração contextualizada, que leve em conta as particularidades de cada tipo de pessoa jurídica, procura-se conferir um tratamento generalizante ao problema. Assim, o debate se concentra na simples análise da subjetividade jurídica. Ou seja, a atribuição de personalidade jurídica, ao promover o desenvolvimento de um novo sujeito, já justificaria por si só o reconhecimento dos mencionados direitos (NEGRI, 2016).

Nesse mesmo sentido, Galgano aponta para as diversas desvantagens no processo de atribuição de personalidade jurídica às sociedades, como a utilização do termo, tanto por tribunais como juristas, de maneira como se existisse um ente único a ser protegido por trás do rótulo da pessoa jurídica – culminando em um tratamento unitário que, além de distorcer a função do instituto, mascara a diversidade de fenômenos que se articulavam em torno daquele termo.<sup>39</sup>

No debate sobre a personalidade eletrônica, o processo de atribuição de personalidade jurídica às sociedades é apresentado como um modelo que justificaria o reconhecimento da subjetividade jurídica para robôs com Inteligência Artificial.<sup>40</sup> Por exemplo, Jacob Turner sustenta que eventuais abusos, como a não responsabilização de programadores e engenheiros, poderiam ser combatidos por meio da desconsideração

<sup>38</sup> RODOTÀ, Stefano. *La vita e le regole: Tra diritto e non diritto*. Milano: Fetrinelli Editore, 2007.

<sup>39</sup> GALGANO, Francesco. Il costo della persona giuridica. *Rivista delle società*, v.13, n. 2, 1968, p. 1-16.

<sup>40</sup> NEGRI, Sérgio. Robôs como pessoas: a personalidade eletrônica na Robótica e na inteligência artificial. *Pensar*, v. 25, n. 3, 2020, p. 1-14.

da personalidade jurídica<sup>41</sup> – argumento que demonstra como a analogia com o direito societário e com as sociedades personificadas é mobilizada sem que, para tanto, sejam apontados os problemas presentes no modelo da sociedade empresária personificada.

O modelo em questão contribui para a indevida compreensão da limitação da responsabilidade dos sócios, ao ocultar a transferência desigual do risco empresarial para terceiros. Se, por um lado, existem credores que podem proteger os seus próprios interesses, renegociando o risco com a sociedade; há, por outro, credores que se mostram impossibilitados de fazê-lo, restando desamparados os indivíduos afetados pelo ente personificado.<sup>42</sup>

A responsabilidade centrada na personalidade desses novos sujeitos (artificiais) pode ocultar os verdadeiros responsáveis pelos danos, transferindo os riscos da atividade desenvolvida por programadores e engenheiros de computação para terceiros que compartilham os mesmos espaços com eles. Ademais, a desconsideração (como proposta por Turner)<sup>43</sup> não representa um instrumento adequado para remediar esses problemas, representando uma técnica que é a principal materialização do unitarismo que marca todo o discurso da pessoa jurídica.

A proposta de criação de uma personalidade eletrônica se insere também em um debate mais amplo: o reconhecimento de novas subjetividades e, conseqüentemente, novos atores jurídicos. Gunther Teubner lembra que, em 1522, ratos foram submetidos a um julgamento no tribunal eclesiástico de Autun.<sup>44</sup> Influenciado pelo processo de racionalização da ciência e da natureza, o número de atores no mundo jurídico foi, como sustenta o autor alemão, reduzido drasticamente por um desdobramento do discurso filosófico da modernidade. Dialogando com a Teoria dos Sistemas de Luhmann e com as concepções de Latour, Teubner rechaça o antropocentrismo que estaria subjacente à análise psicológica e

<sup>41</sup> TURNER, Jacob. *Robot Rules: Regulating Artificial Intelligence*. Nova York: Palgrave Macmillan, 2019.

<sup>42</sup> NEGRI, Sérgio. Robôs como pessoas: a personalidade eletrônica na Robótica e na inteligência artificial. *Pensar*, v. 25, n. 3, 2020, p. 1-14.

<sup>43</sup> TURNER, Jacob. *Robot Rules: Regulating Artificial Intelligence*. Nova York: Palgrave Macmillan, 2019.

<sup>44</sup> TEUBNER, Gunther. Rights of Non-Humans ? Electronic Agents and Animals as New Actors. In: *Journal of Law and Society*, Vol. 33, pp. 497-521, 2006.

sociológica de uma ação intencional, na qual o único ator plausível seria o indivíduo humano.

No desenvolvimento dos seus trabalhos, Teubner <sup>45</sup>, partindo do reconhecimento das diversas interações sociais dos agentes artificiais, chega a defender uma terceira via, pautada na ideia de uma atribuição mitigada ou parcial de personalidade jurídica.<sup>46</sup> Trata-se de construção original que, a princípio, afastaria a personalidade dos agentes artificiais dos problemas gerados pelo discurso da pessoa jurídica na gramática do Direito. Mesmo que essa construção tenha grande importância ao se afastar do unitarismo da pessoa jurídica, “a nova personalidade jurídica” pode acabar aprisionada em uma velha gramática inspirada ainda em um modelo antropocêntrico, como as ideias de direito subjetivo e titularidade. O mesmo pode acontecer com a suposta personalidade eletrônica. Ainda que se evite a associação com a dicotomia pessoa natural e pessoa jurídica, os novos sujeitos se articulam por meio de velhos modelos, que reforcem a já clássica modulação subjetiva do discurso jurídico.

A criação de uma personalidade eletrônica pode acabar repetindo os mesmos problemas. No lugar de se reconhecer as particularidades das diferentes aplicações da Inteligência Artificial, unificar-se-iam essas diferentes relações em modelo jurídico único, pautado exclusivamente na figura de um sujeito abstrato. É fundamental pensar em mecanismos de responsabilidade diferenciados, sensíveis aos diferentes usos desse tipo de tecnologia e aos variados danos que possam ser eventualmente ocasionados.

Possíveis problemas ocasionados, por exemplo, por drones autônomos para fins militares, ou por sistemas de reconhecimento facial no âmbito da segurança nacional certamente são diversos daqueles

<sup>45</sup> TEUBNER, Gunther. *Digital personhood? The status of autonomous software agents in private law*. Tradução de Jacob Watson. Ancilla Iuris, 2018.

<sup>46</sup> No Brasil, o tema foi analisado por José Luiz de Moura Faleiros Junior e Fabiano Menke. “O ponto fulcral dessa proposta, portanto, está na criação de novos elos para a cadeia causal: a noção de *Teilrechtsfähigkeit* significaria apenas que um agente inteligente deve ser tratado como um ente dotado de personalidade jurídica parcial até que se atinja o ponto de indistinção, em que seja capaz de romper tal cadeia causal, realizando, por si e de forma autônoma, um ato danoso. Dessa capacidade, é claro, não se segue que o agente inteligente sempre interrompa a cadeia e, por exemplo, o programador de software ou o passageiro de um veículo autônomo nunca sejam responsabilizados. A subjetividade jurídica parcial simplesmente fornece uma linha de argumento diferente para a interpretação da sucessão de eventos que conduz a averiguação do nexo de causalidade em etapas, permitindo a gradação da personalidade em ritmo consentâneo com a maturação computacional.” FALEIROS JUNIOR, José Luiz de Moura. MENKE, Fabiano. “*Teilrechtsfähigkeit*”: uma proposta alemã para a responsabilização civil na IA. In: *Migalhas de Responsabilidade Civil (Iberc)*. 29 de novembro, 2021.

**Para Pasquale, há uma distinção importante entre “a humanização da tecnologia e a falsificação de características distintamente humanas”. Como afirma o próprio autor, a corrida para imitar a humanidade pode facilmente se tornar uma antessala para a sua substituição.**

ocasionados por robôs sociais que promovem assistência a indivíduos vulneráveis, e por veículos autônomos utilizados para o transporte de pessoas. Diferentes aplicações implicam, necessariamente, em diferentes tipos e graus de risco, demandando, como consequência, um tratamento ético-jurídico diferenciado.

### **3. NOVAS DIRETRIZES DA ROBÓTICA E A CLASSIFICAÇÃO DOS RISCOS NA INTELIGÊNCIA ARTIFICIAL**

#### **3.1 Novas Leis da Robótica**

A partir das Leis da Robótica formuladas na obra de ficção de Asimov, Frank Pasquale apresenta quatro “novas leis da robótica” que, ao contrário das propostas por Asimov, são destinadas aos desenvolvedores de sistemas de Inteligência Artificial e não aos robôs.<sup>47</sup> A primeira nova diretriz estabelece que os sistemas robóticos e a IA devem complementar os profissionais, não os substituir. A segunda diretriz determina que os sistemas robóticos e a IA não devem falsificar a humanidade. Sistemas de IA que criam imagens de “pessoas falsas” e vozes sintéticas têm se tornado cada vez mais comuns. Nesse ponto, devemos questionar se queremos viver em um mundo no qual os seres humanos não sabem se estão lidando com outro ser humano ou com uma máquina. Para Pasquale, há uma distinção importante entre “a humanização da tecnologia e a falsificação de características distintamente humanas”<sup>48</sup>. Como afirma o próprio autor, a corrida para imitar a humanidade pode facilmente se tornar uma antessala para a sua substituição.

O terceiro mandamento fixa que os sistemas robóticos e a IA não devem intensificar as corridas armamentistas de soma zero, isto é, quando o que uma parte ganha é precisamente o que a outra perde.<sup>49</sup> Não podemos esquecer que as tecnologias pioneiras dos exércitos também são utilizadas pela polícia (o uso de reconhecimento facial em multidões, por exemplo) e podem criar um estado de vigilância constante.

<sup>47</sup> PASQUALE, Frank. *New laws of robotics: defending human expertise in the age of AI*. Cambridge, Massachusetts: The Belknap Press Of Harvard University Press, 2020.

<sup>48</sup> PASQUALE, Frank. *New laws of robotics: defending human expertise in the age of AI*. Cambridge, Massachusetts: The Belknap Press Of Harvard University Press, 2020, p.13

<sup>49</sup> PASQUALE, Frank. *New laws of robotics: defending human expertise in the age of AI*. Cambridge, Massachusetts: The Belknap Press Of Harvard University Press, 2020.

Por fim, os sistemas robóticos e a IA devem sempre indicar a identidade de seu(s) criador(es), controlador(es) e proprietário(s). Apesar de a vanguarda dos campos de IA enfatizar a autonomia e uma ideia nebulosa de “robôs fora de controle”, Frank Pasquale entende que algumas pessoas ou organizações devem ser responsáveis por estes sistemas.<sup>50</sup> Mesmo diante de sistemas com aprendizado de máquina, Pasquale aponta que o criador original deve ser obrigado a incluir certas restrições na evolução do código para registrar as influências e evitar resultados ruins. Os reguladores precisam também exigir responsabilidade por *design*, complementando os modelos de segurança por *design* e privacidade por *design*. Como será abordado, algumas diretrizes apresentadas guardam relação com a nova proposta europeia de regulamentação da IA.

### 3.2 Classificação dos Riscos na Inteligência Artificial

Dando continuidade aos esforços de regulamentar o uso da Inteligência Artificial tanto por instituições públicas quanto privadas, a Comissão Europeia, instituição da União Europeia responsável pela realização de propostas normativas, lançou, em abril de 2021, sua proposta de Regulamento para estabelecer regras harmonizadas sobre o tema (*Artificial Intelligence Act – COM/2021/206*).<sup>51</sup>

O regulamento representa uma importante iniciativa legislativa que pode servir de modelo para medidas semelhantes em todo o mundo. Ele impõe amplas obrigações em relação a todos os sistemas de IA, que se aplicam desde os provedores até os usuários; proíbe inteiramente certas aplicações de IA; impõe obrigações especiais em relação a sistemas de IA de alto risco; e cria uma nova estrutura de regulação e processos de teste, monitoramento e conformidade. A ampla definição de sistemas de IA garante que o regulamento terá um impacto significativo em todos os setores da economia, não apenas no digital.<sup>52</sup>

Isso porque o regulamento define sistemas de IA como software que é desenvolvido por meio de determinadas abordagens e técnicas

<sup>50</sup> PASQUALE, Frank. *New laws of robotics: defending human expertise in the age of AI*. Cambridge, Massachusetts: The Belknap Press Of Harvard University Press, 2020.

<sup>51</sup> UNIÃO EUROPEIA. Proposal for a Regulation of the European Parliament and the Council Laying Down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (COM/2021/206). Bruxelas: Comissão Europeia, 2021.

<sup>52</sup> MODRALL, Jay. EU proposes new Artificial Intelligence Regulation. *Inside Tech Law*, April 22, 2021.

e pode, para um determinado conjunto de objetivos definidos por humanos, gerar resultados, como conteúdo, previsões, recomendações ou decisões que influenciam os ambientes com os quais eles interagem.<sup>53</sup> Tais técnicas e abordagens incluem: aprendizado de máquina (supervisionado, não supervisionado e por reforço, usando uma ampla variedade de métodos, incluindo aprendizado profundo); abordagens lógicas e baseadas em conhecimento (incluindo representação de conhecimento, programação indutiva/lógica, bases de conhecimento, mecanismos de inferência/dedução, raciocínio simbólico e sistemas especialistas); abordagens estatísticas, estimação bayesiana, métodos de busca e otimização.

Apesar da ampla definição adotada pela Comissão, com o intuito de abranger uma vasta gama de aplicações, incluindo não só sistemas de IA oferecidos como produtos de software autônomos, mas também produtos e serviços que dependem de serviços de IA direta ou indiretamente, são realizadas distinções, ao longo do regulamento, das normas aplicáveis aos diferentes tipos de sistema.

A Comissão, desse modo, afastou-se do tratamento unitário – que resultaria, por exemplo, da criação de uma personalidade jurídica eletrônica –, e optou por regular o tema de acordo com os diferentes usos e riscos associados a cada tipo de sistema. No documento elaborado, adota-se uma abordagem baseada no risco, diferenciando os usos de IA que criam (i) um risco inaceitável, (ii) um risco alto e (iii) um risco baixo ou mínimo.

Os sistemas de IA identificados como sendo de alto risco são aqueles destinados a serem usados como componentes de segurança de produtos<sup>54</sup>, ou que são eles próprios produtos, abrangidos pela legislação da UE listada no Anexo II do regulamento. Tem-se, por exemplo, máquinas, brinquedos, elevadores, equipamentos de rádio, equipamentos de pressão, equipamentos marítimos, cabos, aparelhos de queima de gás e dispositivos médicos.<sup>55</sup>

<sup>53</sup> UNIÃO EUROPEIA. Proposal for a Regulation of the European Parliament and the Council Laying Down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (COM/2021/206). Bruxelas: Comissão Europeia, 2021.

<sup>54</sup> Componentes de segurança são definidos como componentes de um produto ou sistema que cumprem uma função de segurança e cuja falha e/ou mau funcionamento põe em risco a saúde e a segurança de pessoas e bens (UNIÃO EUROPEIA, 2021).

<sup>55</sup> UNIÃO EUROPEIA. Proposal for a Regulation of the European Parliament and the Council Laying Down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (COM/2021/206). Bruxelas: Comissão Europeia, 2021.

Também são considerados de alto risco os sistemas listados no Anexo III do documento, relacionados à identificação biométrica e categorização de pessoas naturais; gestão e operação de infraestrutura crítica; educação e formação profissional; emprego, gestão de obras e acesso ao trabalho autônomo; acesso e gozo de serviços privados essenciais e serviços públicos e benefícios; aplicação da lei; gestão de migração, asilo político e controle de fronteiras; e administração de justiça e processos democráticos.<sup>56</sup>

O regulamento proíbe, ainda, determinadas práticas para todos os sistemas de IA, independentemente da avaliação de risco, que violam os valores e direitos fundamentais da União Europeia. Isso inclui a disponibilização no mercado de produto ou serviço, ou o uso de sistemas de IA que: a) façam uso de técnicas subliminares, além da consciência de uma pessoa, para distorcer materialmente o seu comportamento de forma a poder ou efetivamente causar danos; b) explorem vulnerabilidades de um grupo devido à idade, deficiência física ou mental, para distorcer materialmente o comportamento de seus membros de uma maneira que cause ou possa causar danos; ou c) avaliem ou classifiquem a confiabilidade das pessoas físicas com base em seu comportamento social ou em características pessoais ou de personalidade conhecidas ou previstas, levando a um tratamento prejudicial ou desfavorável.<sup>57</sup>

A utilização de sistemas remotos de identificação biométrica em tempo real em espaços acessíveis ao público para a aplicação da lei é permitida apenas na medida em que seja estritamente necessário para os objetivos especificados, tendo em consideração determinados elementos. Tais utilizações também devem obedecer a salvaguardas e condições, em particular no que diz respeito a limitações temporais, geográficas e pessoais.<sup>58</sup>

Requisitos para sistemas de IA de alto risco incluem a implementação de um sistema de gerenciamento de risco por todo o ciclo de vida de do sistema, de forma a eliminar ou reduzir os riscos por meio de design e desenvolvimento adequados, implementação de medidas de mitigação e controle, fornecimento de informações,

<sup>56</sup> UNIÃO EUROPEIA. Proposal for a Regulation of the European Parliament and the Council Laying Down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (COM/2021/206). Bruxelas: Comissão Europeia, 2021.

<sup>57</sup> UNIÃO EUROPEIA. Proposal for a Regulation of the European Parliament and the Council Laying Down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (COM/2021/206). Bruxelas: Comissão Europeia, 2021.

<sup>58</sup> MODRALL, Jay. EU proposes new Artificial Intelligence Regulation. *Inside Tech Law*, April 22, 2021.

realização de treinamento e de testes. São, ainda, impostas vastas obrigações relacionadas à governança de dados; documentação técnica e manutenção de registros; transparência e prestação de contas; supervisão humana; precisão, robustez e segurança cibernética. Outro ponto de destaque no regulamento diz respeito ao registro de sistemas de IA de alto risco. Estes deverão, obrigatoriamente, ser registrados e a União Europeia manterá um banco de dados incluindo as informações fornecidas pelos fornecedores, que incluem: a) informações sobre o provedor, o nome comercial do sistema e qualquer referência adicional inequívoca que permita a sua identificação e rastreabilidade; uma descrição da finalidade pretendida do sistema; b) o status do sistema de IA (disponível no mercado, não mais disponível, dentre outros); c) o tipo, número e data de validade do certificado emitido pelo organismo notificado e o nome ou número de identificação desse organismo notificado; d) uma cópia do certificado de conformidade (quando necessário); os Estados-Membros nos quais o sistema IA será ou foi colocado no mercado; e) instruções eletrônicas de uso; e, se desejado, f) uma URL com informações adicionais sobre o sistema.<sup>59</sup> Finalmente, no que diz respeito à responsabilização, tem-se que o regulamento distingue as obrigações impostas entre os fornecedores de sistemas de IA, fabricantes, importadores, distribuidores, usuários e outros terceiros envolvidos na cadeia de desenvolvimento e aplicação da inteligência artificial.

Nesse contexto, os fornecedores de sistemas de IA de alto risco são responsáveis por garantir a conformidade de seus sistemas com o regulamento proposto, implementando um sistema de gestão de qualidade, elaborando a documentação técnica pertinente, mantendo os registros gerados por seus sistemas, cumprindo a avaliação de conformidade e obrigações de registro do sistema, tomando as ações corretivas conforme necessário e cooperando com as autoridades.

Os fabricantes de produtos abrangidos pela legislação da UE, incluindo sistemas de IA de alto risco, são responsáveis pela conformidade, por sua vez, como se fossem os fornecedores do sistema. Distribuidores, importadores, usuários e outros terceiros também estarão sujeitos às obrigações dos fornecedores se colocarem um

<sup>59</sup> UNIÃO EUROPEIA. Proposal for a Regulation of the European Parliament and the Council Laying Down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (COM/2021/206). Bruxelas: Comissão Europeia, 2021.

**Se nos modelos tradicionais, a preocupação se volta para a própria empresa e para os riscos do negócio, ainda que eventuais violações possam impactar nos riscos da própria atividade, o foco deve estar sempre nos detentores dos direitos que deverão ser respeitados.**

sistema de IA de alto risco no mercado ou lhe utilizarem para prestar serviços em seu nome ou marca registrada. Também serão responsabilizados se modificarem a finalidade pretendida do sistema, hipótese na qual o provedor original será isento de responsabilidade.

Usuários devem utilizar tais sistemas de acordo com as instruções de uso fornecidas, garantir que os dados de entrada alimentados ao sistema sejam relevantes, e monitorar a operação com base nas instruções dadas por fabricantes, fornecedores e outros.

No que diz respeito à responsabilidade por sistemas de Inteligência Artificial, portanto, a abordagem adotada pela Comissão revela-se apropriada na medida em que: a) adota uma abordagem baseada no risco, de forma a distinguir os diferentes tipos de aplicação de IA existentes, evitando, assim, limitar de maneira excessiva usos possivelmente benéficos e com potencial de risco baixo ou mínimo; e b) aloca a responsabilidade por eventuais danos causados aos diversos atores envolvidos na sua fabricação, venda e utilização – incluindo tanto pessoas jurídicas, como fabricantes e empresas que empreguem a tecnologia para prestar serviços, quanto pessoas naturais, como usuários que realizam usos inapropriados dos sistemas ou não se atentem às informações e orientações fornecidas.

É importante notar também que processos de gerenciamento de riscos já fazem parte da gramática difundida de *compliance* das próprias empresas. A confusão entre os modelos representa um grave problema, uma vez que, no caso de tecnologias que impactam diretamente direitos humanos, como se observa com o reconhecimento facial, esses processos devem seguir uma lógica própria, completamente diferente dos mecanismos tradicionais de gestão dos riscos de negócios relacionados à sociedade empresária.

Da mesma forma, grupos historicamente submetidos a opressões e violências podem se mostrar mais expostos a um risco maior de vigilância e de outros danos ocasionados pelo uso de novas tecnologias. Se nos modelos tradicionais, a preocupação se volta para a própria empresa e para os riscos do negócio, ainda que eventuais violações possam impactar nos riscos da própria atividade, o foco deve estar sempre nos detentores dos direitos que deverão ser respeitados. Nesse sentido, a efetiva construção de um modelo pautado na alocação diferencial pressupõe a percepção de que os riscos associados

aos usos de novas tecnologias podem não se distribuir de forma linear entre pessoas e grupos.<sup>60</sup>

## CONSIDERAÇÕES FINAIS

O debate sobre a personalidade eletrônica guarda relação direta com os problemas da responsabilidade civil, tendo em vista possíveis danos que possam ser causados pelos artefatos robóticos. Nota-se, como foi demonstrado, uma confusão nessa forma de abordagem: a atribuição de uma suposta personalidade jurídica para robôs é tratada como se fosse uma decorrência natural do debate sobre a responsabilidade. A criação de uma personalidade eletrônica pode, contudo, acabar repetindo os mesmos problemas presentes no unitarismo da pessoa jurídica, com a prevalência de generalizações abstratas que se mostram indiferentes aos variados usos dos sistemas de IA.

Em 21 de abril de 2021, a Comissão Europeia apresentou a Proposta de Regulamento sobre a Inteligência Artificial. O texto se distancia da ideia de uma personalidade eletrônica para adotar um modelo baseado no risco, que articula o conteúdo das disciplina aplicável de acordo com a intensidade dos riscos criados pelos sistemas de IA.

Com o gerenciamento de riscos, surgem obrigações relacionadas a *accountability* dos sistemas, exigindo a elaboração de documentações técnicas que informem características gerais, capacidades e limitações do sistema, manutenção de registros, implementação de medidas de transparência e adoção de rotinas e procedimentos internos que assegurem a governança de dados.

Em síntese, a abordagem baseada em risco deve vir acompanhada de sólidas medidas de transparência, que permitam avaliar o nível de segurança, exatidão, solidez e desempenho dos sistemas de IA. A construção de um sistema sensível aos variados riscos depende, contudo, do reconhecimento do impacto não linear da tecnologia em relação a pessoas e grupos historicamente submetidos a processos de violência e exclusão.

<sup>60</sup> MACHADO, Joana de Souza; NEGRI, Sergio M. C. A.; GIOVANINI, Carolina F. R. Nem invisíveis, nem visados: inovação, direitos humanos e vulnerabilidade de grupos no contexto da COVID-19. *LIINC EM REVISTA*, v. 16, p. 1-21, 2020. Disponível em: <http://revista.ibict.br/liinc/article/view/5367>. Acesso em: 22/01/2021.

## REFERÊNCIAS

- BERTOLINI, Andrea. Robots as products: the case for a realistic analysis of robotic applications and liability rules. *Law, Innovation and Technology*, v. 5, n. 2, 2013, p. 24-247.
- BRYSON, Joanna J.; DIAMANTIS, Mihailis E.; GRANT, Thomas D. Of, for, and by the people: the legal lacuna of synthetic persons. *Artificial Intelligence Law*, n. 25, 2017, p. 237- 291.
- CALO, Ryan. Robotics and the lessons of cyberlaw. *California Law Review*, v. 103, n. 3, 2015, p. 513-563.
- CHOPRA, Samir; LAWRENCE, White. *A Legal Theory for Artificial Autonomous Agents*. Ann Arbor: University of Michigan Press, 2011.
- DARLING, Kate. Extending Legal Rights to Social Robots. *We Robot Conference*, Miami, 2012, p. 1- 18. Disponível em: [http://robots.law.miami.edu/wp-content/uploads/2012/04/Darling\\_Extending-Legal-Rights-to-Social-Robots-v2.pdf](http://robots.law.miami.edu/wp-content/uploads/2012/04/Darling_Extending-Legal-Rights-to-Social-Robots-v2.pdf). Acesso em: 06 jun. 2021.
- EDWARDS, Lilian; VEALE, Michael. Slave to the Algorithm? Why a ‘right to an explanation’ is probably not the remedy you are looking for. *Duke Law & Technology Review*, v. 16, n. 01, 2016, p. 18-84. Disponível em: <https://ssrn.com/abstract=2972855>. Acesso em: 05 jun. 2021.
- FALEIROS JUNIOR, José Luiz de Moura. MENKE, Fabiano. “Teilrechtsfähigkeit”: uma proposta alemã para a responsabilização civil na IA. In: *Migalhas de Responsabilidade Civil (Iberc)*. 29 de novembro, 2021.
- GALGANO, Francesco. *Lex Mercatoria*. Bologna: Il Mulino, 2001.
- GALGANO, Francesco. Il costo della persona giuridica. *Rivista delle società*, v.13, n. 2, 1968, p. 1-16.
- KURZWEIL, Ray. *The Singularity is Near: When humans transcend biology*. Nova York: Penguin Books, 2005.
- LOPES, Giovana. *Inteligência Artificial: Considerações sobre personalidade, agência e responsabilidade civil*, 1 Ed. São Paulo: Editora Dialética, 2021.
- MACHADO, Joana de Souza; NEGRI, Sergio M. C. A.; GIOVANINI, Carolina F. R. Nem invisíveis, nem visados: inovação, direitos humanos e vulnerabilidade de grupos no contexto da COVID-19. *LIINC EM REVISTA*, v. 16, p. 1-21, 2020. Disponível em: <http://revista.ibict.br/liinc/article/view/5367>. Acesso em: 22/01/2021.
- MATTHIAS, Andreas. The responsibility gap: ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*, v. 6, n. 3, 2004, p. 175-183.

- MITTELSTADT, Brent; FLORIDI, Luciano; ALLO, Patrick; TADDEO, Mariarosaria; WACHTER, Sandra. The ethics of algorithms: Mapping the debate. *Big Data & Society*, v. 3(2), 2016, p. 1-21.
- MODRALL, Jay. EU proposes new Artificial Intelligence Regulation. *Inside Tech Law*, April 22, 2021. Disponível em: <https://www.insidetechlaw.com/blog/eu-to-propose-new-artificial-intelligence-regulation>. Acesso em: 17 jun. 2021.
- NEGNEVITSKY, Michael. *Artificial Intelligence: A guide to intelligent systems*, 2 Ed. Harlow: Addison-Wesley, 2004.
- NEGRI, Sérgio. Robôs como pessoas: a personalidade eletrônica na Robótica e na inteligência artificial. *Pensar*, v. 25, n. 3, 2020, p. 1-14.
- NEGRI, Sergio Marcos Carvalho de Ávila. As razões da pessoa jurídica e a expropriação da subjetividade. *Civilistica*, v. 5, n. 2, 2016, p. 1-18.
- PASQUALE, Frank. *New laws of robotics: defending human expertise in the age of AI*. Cambridge, Massachusetts: The Belknap Press Of Harvard University Press, 2020.
- PEARL, Judea; MACKENZIE, Dana. *The Book of Why: The new science of cause and effect*. New York: Basic Books, 1. Ed., 2018.
- POOLE, David; MACKWORTH, Alan. *Artificial Intelligence: Foundations of Computational Agents*. Cambridge: Cambridge University Press, 2010.
- RICHARDS, Neil M.; SMART, William D. How should the law think about robots? In: CALO, Ryan; FROOMKIN, A. Michael; KERR, Ian (orgs.) *Robot Law*. Cheltenham: Edward Elgar Publishing, 2016.
- RODOTÀ, Stefano. A antropologia do homo dignus. Trad. Maria Celina Bodin de Moraes. *Civilistica*, v. 6, n. 2, 2017, p. 1-17. Disponível em: <http://civilistica.com/aantropologia-do-homo-dignus/>. Acesso em: 10 jun. 2021.
- RODOTÀ, Stefano. *La vita e le regole: Tra diritto e non diritto*. Milano: Fetrinelli Editore, 2007.
- RUSSELL, Stuart J.; NORVIG, Peter. *Artificial Intelligence: A Modern Approach*. 3. ed. New Jersey: Prentice-Hall, 2010.
- SAMUEL, Arthur. L. Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, v. 3, n. 3, 1959, p. 210-229.
- TEUBNER, Gunther. *Digital personhood? The status of autonomous software agents in private law*. Tradução de Jacob Watson. Ancilla Iuris, 2018.
- TEUBNER, Gunther. Rights of Non-Humans ? Electronic Agents and Animals as New Actors. In: *Journal of Law and Society*, Vol. 33, pp. 497-521, 2006.

TURNER, Jacob. *Robot Rules: Regulating Artificial Intelligence*. Nova York: Palgrave Macmillan, 2019.

UNIÃO EUROPEIA. 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). Bruxelas: Comissão Europeia, 2018. Disponível em: [Communication Artificial Intelligence for Europe | Shaping Europe's digital future \(europa.eu\)](#). Acesso em: 06 jun. 2021.

UNIÃO EUROPEIA. Proposal for a Regulation of the European Parliament and the Council Laying Down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (COM/2021/206). Bruxelas: Comissão Europeia, 2021. Disponível em: <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206>. Acesso em: 10 jun. 2021.

VLADECK, David C. Machines Without Principles: Liability Rules and Artificial Intelligence. *Washington Law Review*, v. 89, n. 01, 2014, p. 117-150.

## QUALIFICAÇÃO

**Sergio Marcos Carvalho de Ávila Negri** é doutor e mestre em Direito Civil pela Universidade do Estado do Rio de Janeiro. Professor Adjunto do Departamento de Direito Privado da Faculdade de Direito da Universidade Federal de Juiz de Fora e membro do corpo docente permanente do Programa de Pós-graduação *Stricto Sensu* da mesma instituição. Especialista em Direito Civil pela Università Degli Studi di Camerino (Itália). E-mail: [sergio.negri@uff.edu.br](mailto:sergio.negri@uff.edu.br).

**Giovana F. Peluso Lopes** é doutoranda no programa Law, Science and Technology da Universidade de Bologna (Itália). Mestre em Direito pela Universidade Federal de Minas Gerais. E-mail: [giovana.figueiredo@unibo.it](mailto:giovana.figueiredo@unibo.it).