

Leis da robótica: da ficção de Isaac Asimov ao debate sobre sistemas de inteligência artificial

Laws of robotics: from Isaac Asimov's fiction to discussion about artificial intelligence systems

Artigo recebido em 30/06/2023 e aprovado em 01/08/2023.

Sergio Marcos Carvalho de Ávila Negri

Doutor e mestre em direito civil pela Universidade do Estado do Rio de Janeiro (UERJ). Especialização em direito civil pela Università degli Studi di Camerino – Itália. Professor adjunto do Departamento de Direito Privado da Faculdade de Direito da Universidade Federal de Juiz de Fora e do Corpo Permanente do Programa de Pós-graduação estrito senso em Direito e Inovação da Faculdade de Direito da UFJF.

Carolina Fiorini Ramos Giovanini

Mestranda no Programa de Pós-graduação estrito senso em Direito e Inovação da Faculdade de Direito da Universidade Federal de Juiz de Fora (UFJF) e graduada em direito pela mesma instituição. Pós-graduada em direito digital pelo CEPED/UERJ (Universidade do Estado do Rio de Janeiro) em parceria com o ITS (Instituto de Tecnologia e Sociedade do Rio).

Resumo

O presente trabalho busca analisar o papel da obra de Isaac Asimov na disseminação do debate sobre os impactos éticos, sociais e jurídicos da robótica. Para tanto, os dilemas decorrentes do uso de robôs sociais e robôs cuidadores são analisados a partir do conto “Robbie”, que integra o livro “Eu, Robô”. O trabalho também analisa as leis da robótica propostas por Asimov que, posteriormente, foram revisitadas e reformuladas por Frank Pasquale diante do debate sobre desenvolvimento e regulamentação de sistemas de inteligência artificial (IA). Desse modo, o trabalho demonstra que a obra de Asimov exerceu um papel relevante na reflexão sobre dilemas que, atualmente, seguem sendo discutidos em escala global sob a perspectiva de regulamentação do tema.

Palavras-chaves: direito; impacto social; inteligência artificial; legislação; obra literária; robô.

Abstract

This paper seeks to analyze the role of Isaac Asimov's work in the dissemination of the debate about the ethical, social and legal impacts of robotics. To this end, the dilemmas arising from the use of social robots and robot caregivers are analyzed based on the short story “Robbie”, which is part of the book “I, Robot”. The work also analyzes the laws of robotics proposed by Asimov, which were later revisited and reformulated by Frank Pasquale in the face of the debate over the development and regulation of artificial intelligence systems. Thus, the paper demonstrates that Asimov's work played a relevant role in the reflection on dilemmas that are currently being discussed on a global scale under the perspective of regulation of the theme.

Keywords: law; social impact; artificial intelligence; legislation; literature; robot.

1 Introdução

Há mais de duas décadas o Japão vem desenvolvendo robôs para cuidar de idosos, com crescente investimento público e privado no setor. No país, há um receio sobre o envelhecimento da população, uma vez que as taxas de natalidade estão abaixo dos níveis de reposição e, embora nos anos 2000 houvesse cerca de quatro adultos em

idade produtiva para cada pessoa com mais de 65 anos, em 2050 os dois grupos estarão próximos da paridade (WRIGHT, 2023).

Diante do cenário de aumento do número de idosos e dos custos decorrentes dos cuidados direcionados a essa parcela da população, os robôs sociais (ou robôs de cuidados) passam a ser vistos como uma possível solução. Os robôs podem ser fabricados com diferentes funcionalidades, formas e tamanhos. Alguns são voltados aos cuidados físicos, por exemplo, para ajudar a levantar pessoas idosas, como os robôs Robear¹ e Hug². Outros são voltados ao cuidado social e emocional, como fornecimento de companhia e terapia, por exemplo, no caso do robô Paro³, que reproduz a forma de foca e oferece uma modalidade “robótica” de terapia animal.

A pesquisa de etnografia de James Wright (2023) demonstra que a realidade de aplicação dos robôs é mais complexa, uma vez que as habilidades reais dos robôs ficam aquém das expectativas moldadas, por sua imagem exagerada. Wright demonstra que os próprios robôs de cuidado também exigem cuidados: precisam ser movimentados, limpos, inicializados, operados, repetidamente explicados, constantemente monitorados durante o uso e armazenados posteriormente.

A robótica e o imaginário projetado por esse campo, especialmente em razão de metáforas antropomórficas e zoomórficas, criam uma concepção de que humanos e robôs compartilharão lares e vidas. A indefinição ontológica e jurídica acerca dessa tecnologia emergente faz com que elementos da ficção científica sejam transportados para o debate jurídico, como é o caso das leis da robótica, propostas por Isaac Asimov em “Eu, Robô”.

Embora direito e literatura sejam campos distintos, é possível notar que as áreas se conectam e se influenciam mutuamente. Tanto o direito quanto a literatura exploram questões fundamentais sobre a natureza humana e a sociedade, buscando compreender e retratar as complexidades das relações humanas e as dimensões éticas e sociais do comportamento e de interações em geral.

Nesse contexto, o presente trabalho⁴ pretende abordar, a partir do conto “Robbie” – presente na obra “Eu, Robô” de Isaac Asimov – os dilemas éticos e sociais envolvendo robôs cuidadores. Além disso, analisa-se a aproximação entre direito e robótica a partir da passagem das Três Leis da Robótica de Asimov, propostas no âmbito da literatura de ficção, para as Novas Leis da Robótica de Frank Pasquale, elaboradas no contexto de crescente debate sobre a regulação de sistemas de inteligência artificial (IA). Para tanto, o tema será desenvolvido a partir de revisão bibliográfica das obras mencionadas e de trabalhos jurídicos que abordam a relação entre direito civil, robótica e inteligência artificial.

2 Direito e robótica

A robótica desempenha um papel fundamental em diversos setores da economia, por exemplo, no campo da indústria e automação, a robótica já desempenha, há alguns anos, funções relevantes em linhas de produção, montagem, soldagem, embalagem etc. A grande novidade diz respeito à aceitação, nos últimos anos, da ideia de que robôs podem e devem compartilhar o mesmo ambiente que seres humanos. Atualmente, é possível observar uma expansão deste campo em diversos outros setores, como saúde (robôs que auxiliam em procedimentos cirúrgicos, reabilitação, cuidado etc.), serviços (robôs de limpeza, entrega, assistentes virtuais etc.) e educação (robôs que auxiliam em sala de aula, promovem atividades avaliativas etc.).

A robótica está cada vez mais interligada com o campo da inteligência artificial, pois a implementação de algoritmos em artefatos robóticos possibilita o desenvolvimento de sistemas que sejam capazes de realizar tarefas mais complexas. Por exemplo, a visão computacional é um subcampo de estudo da área de inteligência artificial

¹ THE GUARDIAN. Robear: the bear-shaped nursing robot who'll look after you when you get old. Disponível em: <https://www.theguardian.com/technology/2015/feb/27/robear-bear-shaped-nursing-care-robot>. Acesso em: 29 jun. 2023.

² FUJI CORPORATION. Hub. Disponível em: <https://www.fuji.co.jp/en/about/hug/>. Acesso em: 29 jun. 2023.

³ PARO ROBOTS. Paro. Disponível em: <http://www.parorobots.com/>. Acesso em: 29 jun. 2023.

⁴ O trabalho apresenta resultados parciais da pesquisa realizada no âmbito do Projeto “Inovação e Direito na Inteligência Artificial: mapeamento normativo e análise de impacto para o exercício de direitos fundamentais” (CNPq Universal).

que permite que sistemas robóticos identifiquem objetos, reconheçam pessoas, naveguem em ambientes físicos complexos e realizem tarefas com base em informações visuais, contribuindo para o aumento da capacidade que esses artefatos possuem para interagir com o mundo físico ao seu redor.

Desse modo, nota-se que as decisões relacionadas ao desenvolvimento de artefatos robóticos podem influenciar diretamente a sociedade. Lessig (2006, p. 24) já destacava que no ciberespaço a regulação é imposta principalmente por meio de códigos, que podem ser compreendidos como arquiteturas de controle. Assim, o código incorpora valores, escolhas, e determina o que as pessoas podem ou não fazer em tais espaços (LESSIG, 2006, p. 114).

Na mesma direção, Jack Balkin (2015, p.47) aponta que, assim como a lei, a tecnologia também atua como intermediária nas relações sociais entre seres humanos. Para Balkin (2015, p. 48-49), o debate sobre tecnologia relaciona-se, na verdade, com a discussão sobre (i) como as pessoas interagem com novas tecnologias; e (ii) como as pessoas interagem com outras pessoas usando novas tecnologias.

Inclusive, por vezes, as regras incorporadas e propagadas por tecnologias terão mais impacto para permitir ou inibir determinados comportamentos do que as leis e normas que compõem o ordenamento jurídico. Em decorrência do desenvolvimento e da crescente utilização de sistemas de inteligência artificial, alguns países já avançam nas discussões sobre regulação do tema na esfera legislativa⁵.

Ainda que robótica e inteligência artificial não se confundam, o debate sobre regulação de sistemas de IA interfere na robótica na medida em que tais campos estão cada vez mais conectados. Nesse contexto, em 16 de fevereiro de 2017, o Parlamento Europeu publicou a Resolução 2015/2103-INL com recomendações à Comissão de Direito Civil sobre Robótica (UNIÃO EUROPEIA, 2017), que aborda diferentes tópicos relacionados aos impactos do uso de inteligência artificial em setores como educação, mercado de trabalho, transporte, propriedade intelectual, pesquisa e inovação.

A resolução enfatiza a influência de artefatos robóticos inteligentes na sociedade e estabelece os seguintes parâmetros para que um artefato seja categorizado como robô: (i) existência de sensores capazes de permitir a troca de dados com o ambiente; (ii) capacidade de aprendizado com a experiência e interação com o meio; (iii) existência de um suporte material; (iv) capacidade de adaptação; e (v) ausência de vida na acepção biológica.

Ao tratar do regime de responsabilidade civil, a resolução destaca que, em razão do aumento do nível de autonomia de tais artefatos, não seria possível considerá-los como simples ferramentas manuseadas por fabricantes, usuários ou proprietários. O debate acerca da personificação da inteligência artificial, por vezes, surge no contexto de concepção de sistemas de IA como agentes artificiais autônomos.

A partir dessa concepção, a resolução sustenta que, em razão da possibilidade de ações imprevisíveis, o ordenamento jurídico deveria reconhecer personalidade jurídica a esses artefatos, fazendo surgir uma nova subjetividade e constituindo um centro unitário de imputação. Ao apresentar a atribuição de personalidade eletrônica para os problemas de responsabilização civil, nota-se que a resolução trata a personalidade eletrônica como uma consequência automática do debate sobre responsabilidade.

É importante ressaltar, contudo, que a concepção de autonomia usualmente atribuída a robôs e sistemas de inteligência artificial está relacionada à capacidade de realizar tarefas sem a direção ou supervisão humana direta e à capacidade de aprendizado, não se confundindo com a autonomia em sentido forte, correspondente à existência de consciência ou autoconsciência e a consequente identificação de um agente moral (BERTOLINI, 2013, p. 221).

Na mesma direção, Negri (2019, p. 4) ressalta a importância de que a concepção de autonomia não seja confundida com a ideia de imprevisibilidade do resultado, uma vez que máquinas operadas a partir de controle humano também podem ocasionar resultados imprevisíveis. Verifica-se que, embora exista uma tendência de antropomorfização de objetos que apresentam alguns comportamentos semelhantes aos de seres humanos, a ideia

⁵ Por exemplo, no caso dos Estados Unidos, é possível observar regulamentações setoriais específicas e abordagens de co-regulamentação, isto é, as agências reguladoras estabelecem diretrizes a serem implementadas pelos entes privados. Também é possível citar o caso da China e do Reino Unido, que adotam regulamentações sobre usos específicos de inteligência artificial, como é o caso das diretrizes sobre algoritmos. Por outro lado, é possível observar que Brasil e União Europeia adotam estratégias mais abrangentes para regulamentação da inteligência artificial, buscando estabelecer obrigações aplicáveis a todos os sistemas de IA.

de projeção de atributos humanos a ferramentas carrega o potencial de afetar a promoção de direitos fundamentais e não deve ser transposta para estratégias regulatórias.

Ainda que a tecnologia avance e alcance elevado grau de desenvolvimento, sistemas de inteligência artificial são ferramentas desprovidas de características inerentes aos seres humanos. Nesse ponto, Negri (2016, p. 2) ressalta que a armadilha das equiparações é fazer com que as diferenças sejam mascaradas, concretizando o processo denominado por Stefano Rodotà de “expropriação da subjetividade”, no qual, sob o pretexto de proteção de um sujeito abstrato, usurpam-se, no plano concreto, direitos inerentes ao ser humano.

Na mesma direção, Watson (2019, p. 434) entende que a antropomorfização de modelos estatísticos faz com que, implicitamente, ocorra uma concessão de certo grau de agência ao modelo. Para o autor, a antropomorfização faz com que as habilidades do modelo sejam exageradas e parte de nossa própria autonomia seja “apagada”. Trata-se de uma reflexão relevante porque determinados conceitos concebidos em uma lógica marcada pela antropomorfização de artefatos podem equiparar indistintamente pessoas naturais e robôs ou sistemas de inteligência artificial.

É importante notar que, especificamente em grupos marcados por vulnerabilidades, esse processo pode se manifestar de forma acentuada na usurpação de direitos inerentes ao ser humano, sob o pretexto de proteção de meras ferramentas e artefatos. Caso não haja uma reflexão sobre as definições que aproximam artefatos robóticos e sistemas de inteligência artificial de atributos inerentemente humanos, é possível vislumbrar um futuro cenário de restrição de direitos para seres humanos.

Por exemplo, a Suprema Corte dos Estados Unidos, no âmbito do caso *Hobby Lobby*, já reconheceu a liberdade religiosa de uma pessoa jurídica, autorizando que esta deixasse de fornecer aos empregados seguro de saúde que assegurava acesso a métodos anticoncepcionais. Da mesma forma, a criação de uma personalidade jurídica eletrônica para sistemas de inteligência artificial poderia gerar o reconhecimento de direitos para artefatos e, no plano concreto, impactar negativamente a promoção de direitos fundamentais para seres humanos.

Da mesma forma, caso não haja uma reflexão sobre as definições que aproximam artefatos robóticos e sistemas de inteligência artificial de atributos inerentemente humanos, é possível vislumbrar um futuro cenário de restrição de direitos para seres humanos.

Portanto, a proposta de criação de uma personalidade jurídica eletrônica, ao reconhecer uma nova subjetividade e conceder um tratamento jurídico unitário a aplicações de robótica e IA substancialmente diversas, é problemática, tendo o debate regulatório se afastado dessa estratégia⁶.

Em abril de 2021, a Comissão Europeia – instituição da União Europeia competente para propor normas regulatórias – apresentou a proposta de regulamento sobre inteligência artificial (COM/2021/206, 2021), pautada em uma abordagem regulatória baseada em risco, com vinculação da intensidade de intervenção legislativa ao nível de ameaça aos direitos fundamentais e à segurança de pessoas naturais.

Nesse sentido, o *Artificial Intelligence Act* adota uma estrutura de classificação de riscos dos sistemas de inteligência artificial para que as obrigações sejam moduladas de acordo com o risco de cada sistema. Referida classificação engloba os seguintes níveis de risco: (i) risco inaceitável, ou seja, sistemas cuja utilização é vedada; (ii) risco alto, isto é, sistemas autorizados a funcionar, mas sujeitos ao cumprimento de determinados requisitos obrigatórios; e (iii) risco baixo ou mínimo, compreendidos como sistemas de risco limitado ou mínimo (por exemplo, chatbots).

No Brasil, foi apresentado no Senado Federal o Projeto de Lei 2.338/2023, que busca regulamentar o uso de inteligência artificial no país e também apresenta uma proposta de classificação de riscos de sistemas de inteligência artificial a partir de uma avaliação preliminar a ser realizada pelo fornecedor. A classificação presente na proposta

⁶ É importante notar que o afastamento do debate sobre a criação de uma personalidade jurídica eletrônica ocorre na esfera regulatória, que caminha na direção de uma estrutura baseada na classificação de riscos dos sistemas de inteligência artificial. Isso não significa que o debate sobre novas subjetividades e personificação de entes abstratos esteja enfraquecido na doutrina nacional e internacional.

brasileira abrange os seguintes níveis de risco: (i) risco excessivo, classificação segundo a qual o sistema deve ser vedado (por exemplo, sistemas que explorem vulnerabilidades de grupos específicos para que se comportem de forma prejudicial/perigosa à sua saúde, segurança ou contra fundamentos da lei); (ii) alto risco, que demanda a implementação de medidas de governança específicas, como a realização de avaliação de impacto algorítmico (por exemplo, sistemas que envolvam avaliação de capacidade de endividamento de pessoa natural ou classificação de crédito); e (iii) demais classificações, que também ensejam aplicação de medidas de governança.

Para além das propostas legislativas em tramitação no Brasil, o Poder Judiciário brasileiro já conta com iniciativas que exploram o uso da inteligência artificial para aprimorar suas operações e processos de tomada de decisão. Por exemplo, para otimizar e agilizar a gestão de casos por meio de sistemas que realizam classificações dos casos e suas respectivas prioridades. Nesse contexto, o Conselho Nacional de Justiça (CNJ) realizou mapeamento dos projetos que envolviam uso de inteligência artificial nos tribunais e, em 2021, foram identificadas 41 iniciativas em andamento.

O CNJ atua no tema de regulamentação do uso de inteligência artificial por meio do Sinapses, plataforma virtual que centraliza as iniciativas da tecnologia dos tribunais brasileiros. Além disso, a Resolução 332/2020 do CNJ estabelece que a inteligência artificial, no âmbito do Poder Judiciário, visa promover o bem-estar dos jurisdicionados e a prestação equitativa da jurisdição, bem como descobrir métodos e práticas que possibilitem a consecução desses objetivos.

De modo geral, o afastamento - em âmbito regulatório na esfera legislativa - da estratégia de criação de uma personalidade jurídica eletrônica e o desenvolvimento de um modelo de classificação de riscos é um avanço relevante na medida em que afasta uma solução jurídica geral e abstrata. No entanto, é importante notar que a interação entre humanos e máquinas possui um impacto significativo no imaginário coletivo e, conseqüentemente, no debate jurídico, influenciando a forma como compreendemos a atuação de artefatos robóticos e imaginamos o futuro da tecnologia.

3 As Leis da Robótica de Isaac Asimov

O livro "Eu, Robô" reúne nove contos de ficção científica de Isaac Asimov publicados entre os anos 1940 e 1950. O livro aborda temas como inteligência artificial, automação, ética e responsabilidade humana, sendo estruturado como uma série de depoimentos da personagem Susan Calvin, uma psicóloga roboticista que relata suas experiências com robôs ao longo de sua carreira e envolvimento com a empresa U.S. Robôs.

Os contos se passam em um mundo no qual robôs são comuns e desempenham papéis relevantes na sociedade, ou seja, trata-se de um futuro permeado pela convivência entre humanos e máquinas. Por exemplo, o primeiro conto da obra de Asimov trata de Robbie, um robô babá de funcionamento semelhante aos atuais robôs cuidadores, o que mostra que, desde a ficção, robôs passam a ser vistos como parte da sociedade.

Robbie é babá de Glória, uma criança de 8 anos que se apega de forma afetiva ao robô. No conto, a família se preocupa com o efeito que o robô poderia gerar no desenvolvimento da criança e acaba devolvendo Robbie à fábrica. Glória passa a sentir falta da robô, que considera como sua melhor amiga, e deixa de aproveitar sua infância. Na tentativa de fazer com que Glória supere a perda de seu robô babá, a família tem a ideia de visitar uma fábrica de robôs, na qual Glória poderia então ver os robôs como objetos inanimados, não como "pessoas".

Fora do âmbito da ficção científica, Darling (2016, p. 214) aponta que as pessoas são propensas ao antropomorfismo, ou seja, as pessoas projetam qualidades inerentes a seres humanos em outras entidades para fazê-las parecer mais humanas. Esse fenômeno pode ser observado em robôs sociais, definidos por Darling (2016, p. 215) como artefatos fisicamente corporificados que se comunicam e interagem com os humanos em um nível social. A autora mostra que as relações entre humanos e robôs sociais são mais fortes do que as relações entre humanos e outros objetos devido a três fatores: fisicalidade, movimento autônomo percebido e comportamento social.

A fisicalidade está relacionada ao fato de que humanos são criaturas físicas e podem ter reações distintas quando estão em contato com objetos no seu espaço físico ou objetos virtuais. Nesse sentido, o movimento autônomo percebido também está relacionado ao espaço físico, trata-se do movimento físico que não pode ser antecipado. Por exemplo, o robô-aspirador Roomba segue um algoritmo para limpar o piso de um determinado

espaço, ou seja, não é programado para interagir com humanos e não possui nenhuma habilidade social, mas apenas pelo fato de se mover por conta própria em um espaço físico, as pessoas criam nomes, falam com ele e, até mesmo, se sentem mal quando o robô fica preso (DARLING, 2016, p. 217).

Por fim, o comportamento social amplia a antropomorfização na medida em que robôs sociais não são feitos apenas para parecerem adoráveis, mas também são capazes de simular o comportamento real, o que faz com que seres humanos associem suas ações a estados mentais, sentimentos e emoções (DARLING, 2016, p. 218). Para Darling (2016, p. 219), o antropomorfismo é ainda mais plausível quando as pessoas têm pouca noção de como um robô complexo funciona e, por isso, acabam atribuindo autonomia, intenção ou sentimentos a ações que resultam de algoritmos. Nesse sentido, a autora destaca a relação de crianças e idosos que interagem com robôs sociais, que podem confundir o real e o não real.

Na mesma direção, Richards e Smart (2016, p. 16) alertam que a forma como pensamos sobre robôs e seus operadores humanos afetará as escolhas de *design*. A escolha metafórica de projetar uma tecnologia como uma nova versão de algo afeta a aplicação das leis e a formulação de instrumentos regulatórios. Os autores (2016, p. 18) apontam que a utilização de metáforas também é importante para compreender a interação entre humano e robô, pois uma pessoa pode reagir a uma tecnologia de maneira diferente de acordo com a forma de apresentação ou comercialização (um robô que representa uma criança ou um cachorro pode receber tratamento diferenciado de um robô que não simula aspectos antropomórficos ou zoomórficos, por exemplo).

A interação com robôs vem acompanhada de uma projeção de características antropomórficas. Richards e Smart (2016, p. 18) mostram que, até quando o robô não tem um formato integralmente humano, as características, intenções e motivações tipicamente humanas são ali projetadas (por exemplo, as câmeras são descritas como “olhos”), mas é necessário ter em mente que robôs são ferramentas, evitando-se a falácia androide.

Na mesma direção, Judith Donath (2020, p. 54) aponta que a linguagem utilizada nos debates sobre robôs sociais associa esses sistemas a sentimentos e intenções. Para a autora, essa linguagem torna obscura a distinção entre “X é um robô que sente” e “X é um robô projetado para parecer como se sente”. A autora (2020, p. 60) esclarece que, atualmente, nenhuma entidade artificial é consciente ou mesmo primitivamente senciente, porém, os seres humanos possuem a tendência de antropomorfização que contribui para a ilusão de que essas entidades estão alertas e conscientes.

Donath (2020, p. 68) chama atenção para o fato de que nem sempre o engano de identidade é acidental, uma vez que existem sistemas de inteligência artificial projetados para parecer, agir e/ou falar como se ali houvesse uma mente pensante, sentimental e motivada. Até mesmo a linguagem contribui para essa construção, por exemplo, se uma entidade declara “Eu sou um programa”, o uso da palavra “eu” já implica uma existência autoconsciente pensante. O tema ainda é objeto de discussão, especialmente diante das disputas envolvendo a definição do termo “inteligência artificial”.

Quaresma (2020, p. 22-23) entende que são inteligências artificiais todos os sistemas cibernético-informacionais e de computação que simulam ou tentam reconstituir artificialmente os diversos tipos de comportamentos das entidades biológicas humanas do mundo natural, no que se refere às ações e condutas consideradas por elas – entidades – como inteligentes. O autor ressalta que a própria expressão “inteligência artificial” contém um equívoco, qual seja, uma ambiguidade conceitual que decorre do fato de que esses sistemas não surgem artificialmente por si, mas sim da própria inteligência humana.

Nessa direção, Rodotà (2015) ensinava acerca da importância de se refletir sobre a forma como a tecnologia domina até mesmo o vocabulário de objetos de uso comum. Por exemplo, fala-se em “smartphones” e a palavra “inteligente” (*smart*) aparece e Rodotà (2015) já alertava para o fato de que isso não é um detalhe ou uma indicação trivial. Na verdade, trata-se da passagem de uma situação em que a inteligência era reconhecida apenas pelos humanos para outra situação na qual passa a ser apresentada como um atributo também das coisas.

Tais conceitos e definições que atribuem características humanas para objetos de uso comum devem ser questionados. Nessa direção, Kate Crawford (2021, p.19) afirma que a inteligência artificial não é nem artificial nem inteligente. Na verdade, a inteligência artificial é materializável, feita de recursos naturais, combustível, mão-de-obra humana, infraestruturas e logística. Além disso, os sistemas de IA não são autônomos, racionais ou capazes

de discernir nada sem treinamento extenso e computacionalmente intensivo com grandes conjuntos de dados ou regras e recompensas predefinidas.

Por essa razão, Crawford (2021, p. 21-22) aponta que a inteligência artificial como a conhecemos depende inteiramente de um conjunto muito mais amplo de estruturas políticas e sociais. Para a autora, a IA é uma prática técnica e social, que abrange instituições, infraestruturas, política e cultura. Nesse cenário, os sistemas de IA refletem e produzem relações sociais e entendimentos do mundo, consequentemente, o desenvolvimento e o uso de sistemas de inteligência artificial carregam o potencial de gerar riscos e impactos para a promoção dos direitos humanos.

Desse modo, é possível notar que a obra de Asimov⁷ (2014) - especialmente a partir do conto Robbie - já levantava debates acerca da antropomorfização de robôs e atribuição de características humanas a esses artefatos. Para além dos debates sociais e éticos, Isaac Asimov desenvolve as "Três Leis da Robótica", expostas a seguir:

1. *Um robô não pode ferir um ser humano ou, por omissão, permitir que um ser humano sofra algum mal.*
2. *Um robô deve obedecer às ordens dadas por seres humanos, exceto nos casos em que tais ordens entrem em conflito com a primeira lei.*
3. *Um robô deve proteger sua própria existência, desde que tal proteção não entre em conflito com a primeira ou a segunda lei*

A primeira lei estabelece que um robô não pode ferir um ser humano ou, por omissão, permitir que um ser humano sofra algum mal. Para Asimov, a principal diretriz de um robô é garantir a segurança e o bem-estar dos seres humanos. Trata-se de uma lei projetada para evitar que os robôs passem a representar uma forma de perigo para a humanidade e garantir que atuem como agentes protetores.

A segunda lei determina que um robô deve obedecer às ordens que lhe sejam dadas por seres humanos, exceto nos casos em que tais ordens contrariem a primeira lei. Nesse caso, é possível observar que Asimov enfatiza a importância de que robôs atuem seguindo as instruções humanas, ou seja, que atuem de forma secundária, como assistentes.

De acordo com a terceira lei, um robô deve proteger sua própria existência, desde que tal proteção não entre em conflito com a primeira e a segunda leis. Nota-se que esta lei evidencia a autopreservação, isto é, a ideia de que os robôs não se coloquem voluntariamente em situações que poderiam resultar em sua destruição.

Por fim, há a lei zero, enunciada posteriormente e prevalecendo sobre as outras três: um robô não pode fazer mal à humanidade ou, por omissão, permitir que a humanidade sofra algum mal. Trata-se de um mandamento que prevalece sobre as necessidades individuais ou as ordens de seres humanos específicos. Ainda que uma ação aparentemente proteja um indivíduo ou um grupo de pessoas, caso represente uma ameaça à humanidade como um todo, a aplicação da lei zero deve ser priorizada.

Desse modo, a partir das leis da robótica é possível extrair o entendimento de que um robô deve ser programado para não só evitar prejudicar seres humanos, mas para proteger a vida e a segurança da humanidade. Por outro lado, os robôs devem se esforçar para preservar sua própria existência, tomando medidas para evitar danos a si mesmos, mas nunca podendo colocar seres humanos em perigo. Em geral, as leis da robótica são projetadas para criar uma estrutura de comportamento ético para máquinas inteligentes.

Evidentemente, as leis não possuem aplicação prática real, na verdade, são conceitos fictícios criados no contexto de obras literárias. Em que pese as três leis da robótica tenham sido criadas no âmbito da ficção científica, tiveram um impacto significativo na discussão sobre robótica e inteligência artificial no mundo real. As leis criadas por Asimov estabeleceram uma base para reflexões acerca do desenvolvimento ético da tecnologia, além de disseminarem essa discussão entre diferentes setores da sociedade.

A perspectiva de que artefatos robóticos não podem causar danos aos seres humanos influenciou o desenvolvimento de diretrizes e padrões que evidenciaram a importância da segurança e da responsabilidade no

⁷ O livro "Eu, Robô" foi originalmente publicado em 1950 e, posteriormente, republicado em edições mais recentes.

desenvolvimento de sistemas de inteligência artificial. Além disso, a obra de Asimov provoca reflexões acerca do nível de autonomia concedido a sistemas de IA, destacando a importância do controle humano sobre a tecnologia.

Evidentemente, as três leis de Asimov não podem ser encaradas como uma solução prática para a regulamentação da inteligência artificial, mas é importante reconhecer seu papel na abordagem de desafios éticos e sociais relacionados ao desenvolvimento e à implementação da inteligência artificial.

Posteriormente, as leis da robótica foram revisitadas por Frank Pasquale, que entende que o ponto central deste debate é a busca pelo equilíbrio entre um entusiasmo saudável pelos robôs e a consciência de sua alteridade. Para o autor, um robô pode observar o comportamento ao seu redor e implementar respostas a esse comportamento, mas não pode experimentar isso como uma fonte de significado ou motivação. As pessoas se relacionam com outras pessoas como sujeitos autônomos, mas uma máquina não experimenta tal autonomia como um agente.

4 As novas leis da robótica

A partir das leis da robótica formuladas na obra de ficção de Asimov (2014), Frank Pasquale (2020) 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. Pasquale esclarece que os debates sobre robôs geralmente tendem ao utópico ou ao distópico, porém, o futuro da automação dependerá de diversas pequenas decisões sobre como desenvolver a inteligência artificial.

A primeira nova lei estabelece que os sistemas robóticos e a inteligência artificial devem complementar os profissionais, em vez de substituí-los. De acordo com Frank Pasquale (2020, p. 10), o trabalho possui um valor intrínseco para as pessoas e para a governança das comunidades. Portanto, a abordagem humana para automação deve priorizar inovações que complementem os trabalhadores em seus empregos ou os substituam em ocupações perigosas ou degradantes, direcionando-os para outros papéis sociais. Pasquale (2020, p. 11) destaca que, mesmo em setores que aparentemente são suscetíveis à automação, como agricultura e logística, os trabalhadores desempenham um papel essencial na transição para sistemas de IA e robótica, seja coletando ou criando os dados necessários.

A segunda lei estipula que os sistemas robóticos e a inteligência artificial não devem falsificar a humanidade. Segundo Frank Pasquale (2020, p. 13), programas de aprendizado de máquina que geram imagens de “pessoas falsas” e vozes sintéticas podem se tornar cada vez mais comuns. O autor questiona se desejamos viver em um mundo em que os seres humanos não conseguem distinguir se estão interagindo com outro ser humano ou com uma máquina e ressalta a importância de distinguir entre a humanização da tecnologia e a falsificação de características genuinamente humanas.

Nesse ponto, destaca-se a popularização de *deepfakes*, que são vídeos ou imagens manipulados digitalmente que apresentam uma pessoa fazendo ou dizendo algo que ela não fez ou disse na realidade. De modo geral, as manipulações são feitas usando técnicas de inteligência artificial, especialmente o aprendizado de máquina e a síntese de imagens, para criar um conteúdo falso extremamente realista.

A terceira norma estabelece que os sistemas robóticos e a inteligência artificial não devem intensificar as corridas armamentistas de soma zero, em que o ganho de uma parte é exatamente a perda da outra, e vice-versa. Conforme esclarecido por Frank Pasquale (2020, p. 15-16), essa norma se aplica a várias situações além dos campos de batalha. Por exemplo, as tecnologias pioneiras desenvolvidas para uso militar também são utilizadas por departamentos de polícia, como o reconhecimento facial em multidões, podendo criar um estado de vigilância constante.

Por fim, a quarta lei estabelece que os sistemas robóticos e a IA devem sempre indicar a identidade de seu(s) criador(es), controlador(es) e proprietário(s). Essa lei está diretamente relacionada ao fato de que as pessoas são sempre responsáveis por sistemas robóticos ou algorítmicos. Mesmo diante de sistemas sofisticados e de funcionamento complexo, é possível delimitar controles para que desenvolvedores incluam certas restrições durante o processo de desenvolvimento do código. Diante desse cenário, Pasquale (2020, p. 18) entende que os reguladores precisam exigir responsabilidade por *design*, complementando os modelos de segurança por *design* e privacidade por *design*.

Jack Balkin (2017, p.13) também revisitou as leis da robótica de Asimov realizando uma leitura de que, em nosso mundo real, as “leis da robótica” são os princípios jurídicos que devem estabelecer como os seres humanos usam robôs, algoritmos e agentes de inteligência artificial. Assim como Pasquale, Balkin também entende que, na sociedade atual, as leis da robótica deveriam ser leis dirigidas a seres humanos e organizações humanas, não aos próprios robôs, pois uma questão fundamental é entender como os seres humanos exercem poder sobre outros seres humanos a partir da mediação de novas tecnologias. O autor aponta que as obrigações de transparência, devido processo legal e prestação de contas decorrem das três leis reformuladas como requisitos substanciais.

A releitura de Balkin (2017, p.20-23) estabelece que a primeira lei é que os operadores de robôs, algoritmos e agentes de inteligência artificial são fiduciários de informações que têm deveres especiais de boa fé e negociação justa com seus usuários finais (clientes). Na mesma direção, a segunda lei prevê que as empresas privadas que não são fiduciárias de informações também têm deveres para com o público em geral.

Por fim, no âmbito da terceira lei, Balkin (2017, p.30) defende que o dever público central daqueles que usam robôs, algoritmos e agentes de inteligência artificial é não gerar “incômodos algorítmicos”, isto é, as organizações não podem aproveitar as assimetrias de informação, capacidade de monitoramento e poder computacional para externalizar os custos de suas atividades para o público em geral.

Como forma de impulsionar benefícios e gerenciar riscos gerados por sistemas de IA, entende-se que a promoção de estruturas de governança acerca do tema pode auxiliar no desafio de desenvolver e utilizar sistemas a partir de valores éticos e sociais. A governança em IA enfrenta o desafio de traduzir princípios éticos inicialmente vistos como amplos e abstratos em rotinas práticas e concretas. Além disso, a governança da IA precisa tratar de todo o ciclo de vida do sistema de IA, assegurando múltiplos alinhamentos, seja em operações internas, seja em relação a requisitos externos.

O desenvolvimento de sistemas de inteligência artificial exige que uma série de considerações sejam feitas no nível social, ético, técnico e jurídico. O gerenciamento de riscos decorrentes desse processo deve ser acompanhado de obrigações relacionadas à prestação de contas e à demonstração de evidências de conformidade, por exemplo, por meio de documentações técnicas, rotinas e procedimentos internos que farão parte de uma estrutura de governança.

Portanto, busca-se estabelecer parâmetros, rotinas e regras para que o desenvolvimento e o uso de sistemas de IA e artefatos robóticos se mostrem legítimos. Como condição para aceitação social de novas tecnologias, é sempre importante refletir acerca de garantias e modelos transparentes, passíveis de escrutínio e avaliação, capazes de gerar maior confiança para a sociedade, bem como efetivar a responsabilização e a prestação de contas.

5 Considerações finais

Para além dos avanços em termos de eficiência e de benefícios econômicos trazidos por artefatos robóticos e sistemas de inteligência artificial, é importante observar impactos éticos e jurídicos, uma vez que tais tecnologias atravessam interações humanas e passam a estar presentes em processos do dia a dia. Nesse contexto, foi possível identificar que a obra de Isaac Asimov no âmbito da ficção científica exerceu grande influência no debate sobre a robótica.

A partir das leis da robótica, Asimov estabeleceu um conjunto de diretrizes éticas para robôs, que se tornaram referências e pontos de partida para diversos debates e discussões sobre ética e inteligência artificial. Para além disso, os contos de Asimov despertaram reflexões sobre dilemas envolvendo a interação entre humanos e robôs, englobando questões como autonomia e dependência social.

Conforme buscou-se demonstrar, é necessário evitar o imaginário sobre robôs e se afastar da associação com artefatos antropomórficos, como os androides, que aparecem nos filmes e na literatura. A metáfora de enxergar robôs como seres autônomos e conscientes, projetando características humanas nesses sistemas, é particularmente perigosa para o direito.

Frequentemente, o debate sobre autonomia e consciência de artefatos robóticos e sistemas de IA leva à discussão sobre a possibilidade de atribuição de personalidade jurídica, o que também levanta problemas em

razão da criação de um centro unitário que pode desencadear generalizações e reconhecimento de direitos a entes abstratos, prejudicando a proteção da personalidade de pessoas naturais.

Desse modo, ainda que a obra de Asimov e a formulação das leis da robótica tenham tido um papel importante para a disseminação do debate, fato é que não existem leis direcionadas a artefatos robóticos, mas sim aos seres humanos responsáveis pelo desenvolvimento e pela operação de tais sistemas. Por essa razão, Frank Pasquale revisita as leis elaboradas por Asimov no âmbito da ficção científica e propõe mandamentos direcionados especificamente aos desenvolvedores de sistemas de IA, responsáveis por tomadas de decisão durante a estruturação de tais tecnologias.

6 Referências

ASIMOV, Isaac. *Eu, robô*. Curitiba: Aleph, 2014.

BALKIN, Jack. The path of robotics law. *California Law Review Circuit*, Berkeley, v. 06, p. 45-60, jun. 2015.

BALKIN, Jack. The three laws of robotics in the age of big data. *Ohio State Law Journal*, v. 78, (2017), Forthcoming *Yale Law School, Public Law Research Paper* n. 592, 2017. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2890965. Acesso em: 18 jul. 2023.

BERTOLINI, Andrea. Robots as products: the case for a realistic analysis of robotic applications and liability rules. *Law, Innovation and Technology*, Londres, v. 5, n. 2, p. 214-247, 30 dez. 2013.

Burwell v. Hobby Lobby Stores. Inc. Disponível em: www.supremecourt.gov/opinions/13pdf/13-354_olp1.pdf. Acesso em: 16 jul. 2023.

CONSELHO NACIONAL DE JUSTIÇA. Resolução n. 332 de 21/08/2020. Dispõe sobre a ética, a transparência e a governança na produção e no uso de inteligência artificial no poder judiciário e dá outras providências. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 16 jul. 2023.

CONSELHO NACIONAL DE JUSTIÇA. Pesquisa do CNJ mapeia iniciativas de inteligência artificial no judiciário. 2022. Disponível em: <https://www.cnj.jus.br/pesquisa-do-cnj-mapeia-iniciativas-de-inteligencia-artificial-no-judiciario/>. Acesso em: 16 jul. 2023.

CRAWFORD, Kate. *Atlas of AI: power, politics, and the planetary costs of artificial intelligence*. New Haven: Yale University Press, 2021.

DARLING, Kate. Extending legal protection to social robots: the effects of anthropomorphism, empathy, and violent behavior towards robotic objects. In: CALO, Ryan; FROOMKIN, A. Michael; KERR, Ian (ed.). *Robot law*. Cheltenham, 2016.

DONATH, Judith. Ethical Issues in Our relationship with artificial entities. In: DUBBER, Markus D.; PASQUALE, Frank; DAS, Sunit (ed.). *The Oxford handbook of ethics of AI*. Nova Iorque: Oxford University Press, 2020.

FUJI CORPORATION. *Hub*. Disponível em: <https://www.fuji.co.jp/en/about/hug/>. Acesso em: 29 jun. 2023.

LESSIG, Lawrence. *Code: version 2.0*. New York: Basic Books, 2006.

NEGRI, Sergio Marcos Carvalho de Ávila. As razões da pessoa jurídica e a expropriação da subjetividade. *Civillistica.com*, Rio de Janeiro, ano 5, n. 2, p. 1-18, 2016. Disponível em: <https://civillistica.emnuvens.com.br/redc/article/view/265/215>. Acesso em: 29 jun. 2023.

NEGRI, Sergio Marcos Carvalho de Ávila. Robôs como pessoas: a personalidade eletrônica na Robótica e na inteligência artificial. *Pensar*, v. 25, n. 3, 2020, p. 1-14. Disponível em: <https://ojs.unifor.br/rpen/article/view/10178>. Acesso em: 29 jun. 2023.

PARO ROBOTS. *Paro*. Disponível em: <http://www.parorobots.com/>. Acesso em: 29 jun. 2023.

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

QUARESMA, Alexandre. *Inteligência artificial e bioevolução: ensaio epistemológico sobre organismos e máquinas*. Dissertação. [Mestrado em tecnologias da inteligência e design digital] – Pontifícia Universidade Católica de São Paulo, São Paulo.

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

RODOTÀ, Stefano. L'uso umano degli esseri umani. *MicroMega*, Roma, ago. 2015. Disponível em: <http://temi.repubblica.it/micromega-online/addio-a-stefano-rodota-una-vita-per-la-costituzione-la-laicita-e-i-diritti/?printpage=undefined>. Acesso em: 23 ago. 2023.

ROBEAR: the bear-shaped nursing robot who'll look after you when you get old. *The Guardian*, United Kingdom, feb. 2015. Disponível em: <https://www.theguardian.com/technology/2015/feb/27/robear-bear-shaped-nursing-care-robot>. Acesso em: 29 jun. 2023.

THE GUARDIAN. Robear: the bear-shaped nursing robot who'll look after you when you get old. Disponível em: <https://www.theguardian.com/technology/2015/feb/27/robear-bear-shaped-nursing-care-robot>. Acesso em: 29 jun. 2023.

UNIÃO EUROPEIA. (COM/2021/206), 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. Bruxelas: Comissão Europeia, 2021. Disponível em: <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206>. Acesso em: 29 jun. 2023.

UNIÃO EUROPEIA. *Resolução do Parlamento Europeu, de 16 de fevereiro de 2017*. Com recomendações à Comissão Direito Civil sobre Robótica. Bruxelas: Parlamento Europeu, 2017. Disponível em: http://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_PT.html?redirect. Acesso em: 29 jun. 2023.

WATSON, D. The rhetoric and reality of anthropomorphism in artificial intelligence. *Minds & Machines*, Berlim, v. 29, n. 3, p. 417–440, 21 set. 2019. Disponível em: <https://link.springer.com/article/10.1007%2Fs11023-019-09506-6>.

WRIGHT, James. Inside Japan's long experiment in automating elder care. *MIT Technology Review*, 2023. Disponível em: <https://www.technologyreview.com/2023/01/09/1065135/japan-automating-eldercare-robots/>. Acesso em: 29 jun. 2023.